



HAL
open science

Using facial cues to produce social decisions. A cognitive and evolutionary approach

Lou Safra

► **To cite this version:**

Lou Safra. Using facial cues to produce social decisions. A cognitive and evolutionary approach. Cognitive Sciences. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT : 2017PA066317 . tel-01724105

HAL Id: tel-01724105

<https://theses.hal.science/tel-01724105>

Submitted on 6 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Pierre et Marie Curie

École Doctorale Cerveau Cognition Comportement

Laboratoire de Neurosciences Cognitives

Using facial cues to produce social decisions

A cognitive and evolutionary approach

Par Lou Safra

Thèse de doctorat de sciences cognitives

Dirigée par Dr. Coralie Chevallier

Présentée et soutenue publiquement le 5 septembre 2017

Devant un jury composé de :

Dr. Jean-François Bonnefon, Rapporteur

Dr. Hugo Mercier, Rapporteur

Dr. Mathias Pessiglione, Examineur

Dr. Coralie Chevallier, Directrice de thèse

Contents

ACKNOWLEDGEMENTS	4
ABSTRACT	7
RÉSUMÉ	9
FOREWORD	11
CHAPTER ONE. FACE EVALUATIONS AS AN ENTRY POINT INTO SOCIAL COGNITION	13
Importance of face evaluations	13
Functional bases of face evaluations	17
One face, multiple evaluations	21
CHAPTER TWO. INVESTIGATION OF AN INTERNAL VARIABLE: STUDY OF SOCIAL MOTIVATION	27
Social motivation as cooperative strategy	27
<i>Paper 1: Motivation for social bonding promotes high-stakes cooperative strategies, Safra, L., Wyart, V., Baumard, N., Chevallier, C.</i>	
Effect of social motivation across populations	64
<i>Paper 2: Distinct effects of social motivation on face evaluations in adolescents with and without autism, Safra, L., Ioannou, C., Amsellem, F., Delorme, R., Chevallier, C.</i>	
Social motivation as a high-stake cooperative strategy	95
Towards new investigations of cooperative strategies	96

CHAPTER THREE. INVESTIGATION OF THE COGNITIVE PROCESSES UNDERLYING SOCIAL CHOICES: STUDY OF LEADER CHOICES	99
An ecological approach to human behavior	99
Childhood environment and social preferences	103
<i>Paper 3: Childhood harshness predicts long-lasting leader preferences</i>	
<i>Safra, L., Algan, Y., Tecu, T., Grèzes, J., Baumard, N., Chevallier, C.</i>	
An evolutionary perspective on leader preferences	125
<i>Paper 4: Why would anyone elect a narcissistic untrustworthy leader?</i>	
<i>Safra, L., Baumard, N., Chevallier, C.</i>	
Political attitudes as behavioral responses to the environment	130
CHAPTER FOUR. GENERAL DISCUSSION	164
An original approach to study social behavior	164
Beyond faces and social interactions	167
From the observer to the target	172
Limitations of the cooperation-power model	175
EPILOGUE. FACE EVALUATIONS, AN EVOLUTIONARY PUZZLE	179
An evolutionary perspective on face evaluations	179
Puzzle 1. Face evaluations accuracy	181
Puzzle 2. Signaling untrustworthiness	183
GENERAL REFERENCES	189
APPENDIX	230
FORMATIONS SUIVIES	354

Acknowledgements

Bien que j'ai été prévenue de la difficulté de la tâche, je ne me doutais pas qu'il serait si compliqué d'écrire ces remerciements tant il est clair pour moi que je dois énormément aux personnes qui m'ont aidée, entourée et accompagnée durant toutes ces années, sans lesquelles ce travail n'aurait sûrement jamais vu le jour, et surtout sans lesquelles je ne serais sûrement pas la même personne aujourd'hui.

Je souhaiterais tout d'abord remercier tous les chercheurs qui ont accepté de m'accueillir dans leurs équipes et qui m'ont offert la chance de travailler sur leurs projets, depuis le tout début alors même que je ne connaissais rien, ou pas grand chose, à la recherche et aux sciences cognitives. Ainsi, je remercie de m'avoir donné ma chance, par ordre chronologique : Michaël Zugaro et Anne Cei, Mathias Pessiglione et Florent Meyniel, Julie Grèzes et Lucile Gamond, Natalie Sebanz, Günther Knoblich et Cordula Vesper, et

bien sûr Coralie Chevallier et Nicolas Baumard. Plus que tout ce que j'ai pu apprendre scientifiquement en travaillant avec vous, j'ai également pu découvrir qu'il était possible de faire de sa passion son travail ; je vous en suis extrêmement reconnaissante.

Je souhaite plus particulièrement remercier ma directrice de thèse, Coralie Chevallier, de m'avoir encadrée, formée, soutenue et, surtout, fait confiance. Merci notamment de m'avoir toujours encouragée à développer mes projets, aussi bien dans le monde de la recherche qu'à l'extérieur. Ces quatre dernières années ont été pour moi particulièrement enrichissantes, aussi bien sur le plan scientifique que personnel ; elles ne l'auraient certainement pas autant été avec un autre directeur de thèse.

Merci plus largement à tous les membres de l'équipe Evolution et Cognition Sociale, j'ai adoré faire partie d'une équipe dans laquelle les idées s'échangent aussi librement et avec autant de bienveillance. Merci notamment à Pierre pour m'avoir autant épaulée et à Hugo d'avoir supporté mes blagues pendant toute une année, merci à Christina de rester anarchiste dans un pays aussi capitaliste, à Katharina pour les musées d'art moderne de Paris et d'ailleurs.

Merci à mes voisins du LNC, notamment Gabriel, pour m'avoir toujours rappelé qu'il ne fallait jamais oublier d'aller skier, et le 'Social Group' : à Michele (à qui est dédié le troisième chapitre de cette thèse), Marwa et

Emma pour leurs rires qui réchauffent l'open space en hiver et y apportent un peu de fraîcheur en été et à Guillaume, même s'il reste un petit joueur à côté de Jésus. Merci aussi aux nouveaux Pls, Stefano et Valentin, pour leur aide et leur soutien et pour avoir pimenté mes réunions de travail. Enfin, merci à Margaux et Marine, aux goodies post-déjeuner et à la compagnie créole *en open space*.

Merci aux membres de *Scalp!* Lise, Henri, Tristan et Thibaud, grâce à qui j'ai pu découvrir de nouveaux horizons bien loin des laboratoires.

Et bien sûr, merci à tous ceux à qui je dois bien plus que cette thèse : ma maman, mon frère&Marie, Agathe&Félix, mon papa, Sarah, Jérôme, Amélie et Mélanie. Merci d'être toujours là pour moi, et surtout merci de rire de tout. La vie est beaucoup, beaucoup moins terne et beaucoup, beaucoup plus drôle grâce vous.

Je n'oublie bien sûr pas mes PPettes adorées: Camille, Élise, Lucile, MK, Clément et Paco, ma bouffée d'oxygène venue des quatre coins de la planète.

Merci aussi à Roland, Jessica, Tom et Adrian pour leur soutien sans faille.

Merci à Éric, pour tout.

*En bénissant les hasards qui nous firent naître ensemble, et pas d'un siècle d'écart,
qui nous firent naître ensemble, ni trop tôt et ni trop tard.*

Abstract

Face evaluations are a crucial component of social behavior, influencing a large range of social decisions from mating to political vote. Face evaluations are also susceptible to great individual differences. In this thesis, I propose that individual differences in face evaluations constitute a promising tool to investigate social behavior through the analysis of variations in the weight granted to different social signals, and notably cooperation- and power-related cues. I apply this approach in two ways. First, I examine the hypothesis that social motivation can be construed an adaptation to highly cooperative environments. Across six studies, I confirm a central prediction of this theory, by revealing that highly socially motivated individuals grant a higher importance to cooperation-related signals. Second, I investigate the cognitive mechanisms underlying political choices by examining their responses to different environmental signals. In a study on leader preferences in children and in adults, I show that early exposure to environmental harshness is associated with a preference for

stronger leaders. Building on these results, I then develop an original theory on political choices stating that leader preferences are biased towards the candidates perceived as the most competent for succeeding in the current context, independently of their leadership abilities. To summarize, my thesis puts forward a new framework to investigate social decisions based on individual variations in face evaluations and sheds light on the cognitive processes underlying social behavior as well as their evolutionary bases.

Résumé

Les évaluations faciales sont un élément central des comportements sociaux, influençant un large spectre de choix allant du choix du partenaire sexuel aux choix politiques, mais sont également le sujet de grandes variations individuelles. Dans cette thèse, je propose que l'analyse des différences individuelles dans le poids accordés à différents signaux sociaux, et notamment à ceux de coopération et de pouvoir, lors des évaluations faciales est un outil prometteur pour l'étude des comportements sociaux; j'applique cette approche de deux façons. Dans un premier temps, pour examiner l'hypothèse selon laquelle la motivation sociale correspond à une adaptation aux environnements coopératifs. À travers six études, je montre que la motivation est associée à une plus grande importance donnée aux signaux de coopération, confirmant ainsi une prédiction centrale de cette théorie. Dans un second temps, j'étudie les mécanismes cognitifs impliqués dans les choix politiques, en examinant leurs réponses à différents signaux de l'environnement. Dans une étude sur les préférences politiques des enfants et des adultes, je montre que l'expérience précoce d'un environnement difficile est associée à une

préférence pour les hommes forts. À partir de ces résultats, je développe une théorie originale sur les préférences politiques selon laquelle les choix politiques sont orientés vers les individus perçus comme les plus aptes à réussir dans le contexte présent, indépendamment de leurs qualités en tant que chef de groupe. En résumé, tout au long de cette thèse, je présente un nouveau cadre de travail pour étudier les décisions sociales, basé sur les variations individuelles dans les évaluations faciales, et je démontre comment ce cadre de travail peut éclairer les mécanismes cognitifs sous-tendant les comportements sociaux ainsi que leurs bases évolutives.

Foreword

As soon as they detect their conspecifics, human beings automatically attribute them personality traits. The question of the accuracy of these first impressions is historically ancient, with the first evidence of physiognomy dating back from the antiquity. This approach peaked in popularity during the 19th century, notably with the work of Cesare Lombroso who tried to identify the facial features that characterize criminals. Although the work of Lombroso is now outdated and physiognomy considered a pseudo-science, the accuracy of face evaluations is still a hot question in psychology. In the past few years, multiple scientific studies, especially in evolutionary psychology, have tried to answer this question and to understand what face evaluations actually reveal about the target of the evaluations. This line of research has revealed that there may be at least a kernel of truth in face

evaluations, with perceived dominance, trustworthiness or extraversion correlating with the individuals' actual behavior (Bonnenfon, Hopfensitz, & De Neys, 2013, 2017a; Bonnenfon, Hopfensitz, & Neys, 2015; Carré, McCormick, & Mondloch, 2009a; De Neys, Hopfensitz, & Bonnenfon, 2013; Kramer, King, & Ward, 2011; Tognetti, Berticat, Raymond, & Faurie, 2013).

However, what face evaluations reveal about the observer is far less investigated. While the first impressions individuals form about others can vary greatly from one observer to the next (Hehman, M, Flake, & Slepian, 2017), the mechanisms that are responsible for such variability are mostly underspecified. Yet, investigating inter-individual variability in face evaluations has the potential to provide unique insights into social decisions. On the one hand, studying the association between variations in individual characteristics (e.g., in personality traits) and in the reliance on specific facial features, i.e., on specific social signals, for social choices can shed a new light on the functional consequences of these differences and thus on their adaptive value in different environments. On the other hand, the analysis of these individual variations may also reveal important information on the way social decisions respond to external factors and thereby on the underlying cognitive processes and on their evolutionary roots. Based on these ideas, in the present thesis, we investigated individual differences in face evaluations as responses to internal and environmental variables to gain new insights on social behavior from an evolutionary perspective.

Chapter one. Face evaluations as an entry point into social cognition

Importance of face evaluations

Faces are among the first signals individuals have access to during a new encounter and they carry important information about the future of the upcoming social interaction. They provide a vehicle for emotion expressions signaling others' internal state and intentions as well as crucial features of the environment, such as the presence of an external threat (Dezecache, Mercier, & Scott-Phillips, 2013; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007). In addition to conveying information, faces are also used as social signals *per se*. Indeed, faces are consistently used to infer personality

traits and to form expectations about others' behavior (see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015 for a review on this topic).

These face evaluations are far from anecdotic: they heavily influence social interactions of high evolutionary relevance, such as cooperation and mating (e.g., Rezsescu, Duchaine, Olivola, & Chater, 2012; Valentine, Li, Penke, & Perrett, 2014). For instance, facial trustworthiness has been shown to strongly influence investments in economic games (Chang, Doll, van 't Wout, Frank, & Sanfey, 2010; Rezsescu et al., 2012; Tingley, 2014; van 't Wout & Sanfey, 2008). More precisely, van't Wout & Sanfey (2008) have shown that the more trustworthy participants perceived their partner's face the more they invested money in a trust game. Similarly, Tingley (2014) has found that partners represented by avatars with higher levels of trustworthiness received higher investments in a trust game.

Importantly, face evaluations not only impact the very first interactions with completely unknown individuals but their influence persists even after people have received information about others' actual behavior (Chang et al., 2010; Rezsescu et al., 2012). As an example, even when participants have access to their partners' cooperative history, they still cooperate more with a trustworthy-looking partner than with an untrustworthy-looking one (Rezsescu et al., 2012). Going one step further, Chang et al (2010) have revealed that partner's facial trustworthiness still impacts investment decisions in a trust game after 14 cooperative interactions. More strikingly, first impressions from faces also influence decisions for which multiple other

sources of information are accessible, such as judicial and political decisions (e.g., Porter, Brinke, & Gustaw, 2010; Todorov, Mandisodza, Goren, & Hall, 2005a; J. P. Wilson & Rule, 2015a). For instance, individuals require less evidence to condemn more untrustworthy-looking individuals and are more likely to condemn them to a death-sentence (Porter et al., 2010; J. P. Wilson & Rule, 2015a). Similarly, elections outcomes can be predicted with a power as high as 70% based on the evaluations of the candidates' faces (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Chen, Jing, & Lee, 2014; Laustsen, 2013; Olivola & Todorov, 2010; Sussman, Petkova, & Todorov, 2013; Todorov et al., 2005a).

These high-impact evaluations are not the result of an effortful and conscious process but rather of an unreflective mechanism. Indeed, ratings of competence made after a 100-millisecond exposure already hold a predictive power 70% on election outcome (Ballew & Todorov, 2007; Olivola & Todorov, 2010). More precisely, 50 to 100 milliseconds are enough to form stable evaluations of unknown faces on various personality traits, ranging from attractiveness to competence (Ballew & Todorov, 2007; Bar, Neta, & Linz, 2006; Borkenau, Brecke, Möttig, & Paelecke, 2009; Olson & Marshuetz, 2005; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Going one step further, the categorization of individuals based on their facial features seems to be automatic. Using fMRI, Engell et al. (2007) have shown that faces are automatically encoded as trustworthy or untrustworthy even when participants perform unrelated tasks. Similarly,

experiments on the preconscious processing of facial cues have shown that untrustworthy and dominant avatars are processed differently compared to trustworthy or submissive ones (Getov, Kanai, Bahrami, & Rees, 2015; Stewart et al., 2012).

Automatic social evaluations of neutral faces are not only present in adults but have been evidenced throughout the lifespan. From 6 months, infants process fixed facial features of attractiveness, dominance and trustworthiness (Jessen & Grossmann, 2016; Ramsey, Langlois, Hoss, Rubenstein, & Griffin, 2004) and from 6 years old, children have first impressions of unknown faces similar to those of adults (Cogsdill, Todorov, Spelke, & Banaji, 2014). Going one step further, it has been shown that children are not only able to form adult-like impressions from others' faces but that they also use these face evaluations to guide their social decisions. For instance, 10 year-old children use facial cues of trustworthiness to decide who to cooperate with in a trust game (Ewing, Caulfield, Read, & Rhodes, 2015). In summary, face evaluations are both automatically triggered by the detection of a new face and orient social decisions in many domains throughout the lifespan, placing them as a central component of human social behavior.

Functional bases of face evaluations

The importance of face evaluations for social interactions has made first impressions a major area of research in cognitive science. Functional models of first impressions have been put forward to characterize the cognitive and evolutionary bases of face evaluations. In particular, in 2008, Oosterhof & Todorov (2008) have proposed a functional model of face evaluations using a data-driven approach. More precisely, in order to identify the dimensions on which face evaluations are built, they applied a principal components analysis on 15 traits evaluated on 66 faces by 327 participants. Thereby, they demonstrated the existence of two dimensions underlying face evaluations, respectively explaining 63% and 18% of the total variance. The first dimension correlated strongly with perceived caring and sociability. It was construed as representing the valence of the face and approximated as a 'trustworthiness dimension' due to its strong correlation with this trait ($r = 0.94$). The second one was construed as a 'dominance dimension' (correlation with dominance evaluations: $r = 0.93$).

This functional model states that any face evaluation can be decomposed in the two-dimensional space of dominance and trustworthiness. In other words, it implies that any social judgment made from a face can be construed as a weighted combination of perceived trustworthiness and perceived dominance. For instance, in their seminal paper (Oosterhof & Todorov, 2008), the authors reconstructed a "threat" gradient corresponding to a 45° rotation of the trustworthiness and dominance axes

such that less trustworthy and more dominant faces are rated as most threatening while more trustworthy and less dominant faces are perceived as least threatening (Figure 1).

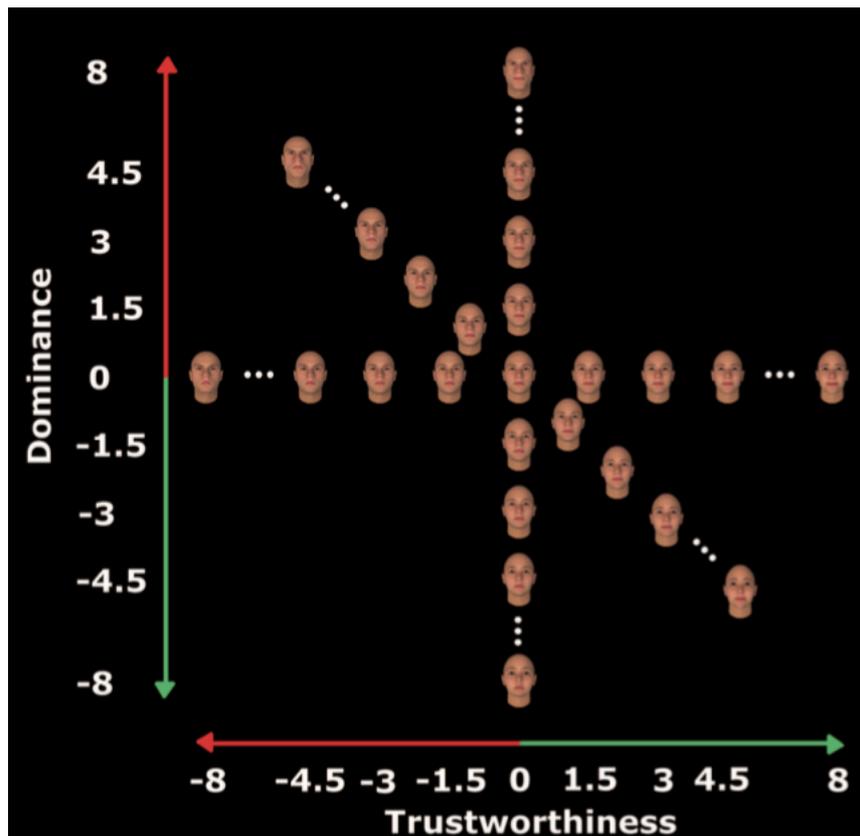


Figure 1. The threat gradient (diagonal line) modeled by Oosterhof & Todorov (2008) based on the dominance and trustworthiness dimensions

The model of face evaluations developed by Oosterhof & Todorov (2008) provides new conceptual tools for the understanding of social choices. Especially, it opens the path to more precise characterizations of the impact of multidimensional social traits, such as masculinity, on social decisions.

Masculinity is an important cue for social choice (Little, DeBruine, & Jones, 2010; Spisak, Homan, Grabo, & Van Vugt, 2012) and appears to be associated with multiple – sometimes contradictory – traits, such as higher and lower levels of prosociality (Boksem et al., 2013; Bos, Terburg, & Honk, 2010; Carré et al., 2009a; van Honk, Montoya, Bos, Vugt, & Terburg, 2012; Wibrál, Dohmen, Klingmüller, Weber, & Falk, 2012). This multidimensionality makes difficult to clearly identify the behavior that is expected from masculine partners. On the contrary, in the model developed by Oosterhof and Todorov (2008), masculinity can be decomposed as a weighted combination of dominance and trustworthiness, two traits that are associated with clear behavior.

Indeed, both trustworthiness and dominance are associated with precise behaviors that are closely associated with important dimensions of human social interaction. On the one hand, trustworthiness relates to the probability of cooperating (Chang et al., 2010), and cooperation is acknowledged to be a major feature of human ecological niche as well as to have acted as an important pressure in the evolutionary history of the human species (Kaplan, Hooper, & Gurven, 2009; Stevens & Hauser, 2004). On the other hand, dominance is associated with another fundamental aspect of human social organization: power (Kaplan et al., 2009; Todorov, Said, Engell, & Oosterhof, 2008). Indeed, across taxa, facial features of dominance, such as the width-to-height ratio, have been quite systematically associated with physical and social power, indicating an

evolutionary old relation between facial cues of dominance and power-related behavior (Carmen Emilia Lefevre et al., 2014, 2014; G. J. Lewis, Lefevre, & Bates, 2012; V. Wilson et al., 2014). The framework developed by Oosterhof and Todorov (2008) thus offers the possibility to analyze face evaluations and social decisions from faces in terms of weighting of two highly relevant signals for human social organization: cooperation and power.

Before exposing some of the lines of research opened by this framework, it is important to note that this model has been replicated for face evaluations using different samples of faces (Sutherland et al., 2013; Vernon, Sutherland, Young, & Hartley, 2014) as well as for first impressions from voices (McAlear, Todorov, & Belin, 2014). Going one step further, these two dimensions of trustworthiness-cooperation and dominance-power closely match more general models of social interactions, such as the warmth-competence model developed by Fiske et al. (2007). Indeed, trustworthiness signals the cooperative intentions of the agent (her warmth) while dominance gives an indication of her ability to implement these intentions (her competence). In line with the high relevance of cooperation- and power-related signals for humans' social life, the correspondence between these two frameworks suggests that the functional model developed by Oosterhof and Todorov (2008) can be extended to a more general cooperation-power framework for analyzing social information processing. As a consequence, the investigation of the weighting of

cooperation- and power-related signals in face evaluations can provide more general insights on their importance for social behavior. In particular, variations in social choices from faces can reveal more global variations in the importance granted to cooperation and power for social decisions.

One face, multiple evaluations

The model of face evaluations developed by Oosterhof and Todorov (2008) provides a new conceptual framework to understand the diversity of social choices both in terms of contextual variations and individual differences. Indeed, face-based social choices are highly sensitive to external signals such as pathogens, resource scarcity or social threats (DeBruine, Jones, Crawford, Welling, & Little, 2010; Little, Roberts, Jones, & DeBruine, 2012; Watkins, DeBruine, Little, Feinberg, & Jones, 2012; Watkins & Jones, 2012). For instance, placing participants in a social threat context systematically increases their preference for masculine leaders (Laustsen & Petersen, 2015; Little, Burriss, Jones, & Roberts, 2007a; Little et al., 2012; Spisak, Dekker, Krüger, & Vugt, 2012). The Oosterhof & Todorov's model (2008) offers the possibility to interpret these contextual variations as differences in the weight granted to cooperation- and power-related signals for leader choice in different contexts. In particular, as perceived masculinity is associated with both an increased perceived dominance and a decreased perceived trustworthiness (Oosterhof & Todorov, 2008), this model allows

to disentangle which of the dominance-power or trustworthiness-cooperation dimension drives the preference for more masculine leaders in times of social threat. In other words, it can inform us about the relevance of cooperation- and power-related signals for leader choices in these contexts. Therefore, applying this cooperation-power framework can provide new insights on the functional consequences of the variability of social decisions across contexts, and thus can shed a new light on the underlying cognitive mechanisms.

In addition to this individual flexibility, the cooperation-power framework can also be used to investigate a crucial aspect of social choices: inter-individual variability. Indeed, social choices often vary from one individual to the next, as regularly revealed by political elections outcomes. Similarly, individuals often differ in their evaluation of others' faces. As an example, while inter-rater reliability can be as high as 0.90 for trustworthiness and attractiveness evaluations, it can be as low as 0.60 for boringness evaluations and raters' characteristics can explain up to 25% of the evaluations variance (Engell et al., 2007; Hehman et al., 2017; Oosterhof & Todorov, 2008; Rule, Krendl, Ivcevic, & Ambady, 2013).

More than just the descriptive interest of inter-individual differences in face evaluations, the study of these differences can be especially fruitful when investigated in the Oosterhof & Todorov's model (2008). Indeed, by applying the same reasoning on inter-individual variations as on inter-context variations, one can hypothesize that individual differences in face

evaluations reveal individual differences in the relevance of cooperation- and power-related signals. This method can thus provide insights into the influence of specific individual variables, such as personality traits or culture, on the importance granted to these two dimensions of social interactions. For instance, Chen et al. (2016) have revealed that social competence was more correlated with trustworthiness in American participants than in Chinese participants, suggesting that American participants granted more weight to cooperation-related signals, and thus to cooperative motives, for evaluating social competence than Chinese participants. Going one step further, this analysis can reveal some the functional consequences of specific individual characteristics, and help to discuss their adaptive value in different environments. In summary, applying the cooperation-power framework to individual differences in face evaluation can help to gain a better understanding of human social behavior in an evolutionary perspective.

In the present thesis, we precisely apply this approach by using individual differences in the weight granted to cooperation- and power-related cues for face evaluations to investigate both the internal variables that influence social decision-making and the cognitive processes underlying social choices.

In the second chapter of this thesis, we investigate social motivation, an individual variable that is also a major drive of human behavior (Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012). More precisely, we use individual

variability in face evaluations to test a central prediction of a recent evolutionary hypothesis of social motivation stating that social motivation can be construed as an adaptation to highly cooperative environments (Chevallier, Kohls, et al., 2012). In a total sample of 410 participants and across 5 experiments, we show, in line with this hypothesis, that highly socially motivated individuals grant more relative weight to cooperation-related cues, i.e., trustworthiness cues. This weighting difference results in the widening of the cooperation circle to more powerful partners in highly socially motivated individuals, which may be identified as a high-stake cooperative strategy.

Building on these results, we then examine whether these differences would extend to younger individuals as well as to individuals with autism, a neuropsychological condition notably characterized by a decrease in social motivation. More precisely, we apply the dimensional approach to neuropsychological conditions that has received an increased interest in the past few years (Cuthbert, 2014; Insel et al., 2010; National Institute of Mental Health, 2015; Sanislow et al., 2010) to test whether social motivation has the same effects in adolescents with and without autism (each $N = 20$) as in adults. Confirming the evolutionary model of social motivation (Chevallier, Kohls, et al., 2012) and the results obtained in adults, we show that highly socially motivated adolescents without autism grant more weight to cooperation-related cues. However, our results reveal that social motivation has a different effect in individuals with autism, which questions

the insights that can be derived from the dimensional approach of neuropsychological conditions. As a conclusion, the investigation of the face evaluations differences associated with social motivation provides new insights into key aspects of this specific type of motivation from both an evolutionary and a clinical perspective.

Internal variables, such as personality traits, are not the only factors influencing behavior and social cognition. In particular, an increasing number of findings have shown the impact of environmental factors, such as poverty or pathogens prevalence, on social behavior and on face evaluations (Little et al., 2010, 2012; Watkins et al., 2012; White, Kenrick, & Neuberg, 2013). For instance, Little et al. (2012) have found that women's preference for more masculine men decreases as the general health status in their country increases. In addition, a large body of evidence in ecology suggests that the behavior is influenced not only by the current environment of the individuals but also by the environment experienced during development (Roff, 2002; Stearns, 2000). Indeed, it has been shown that differences in the environmental signals perceived during childhood account for important variations in behavior (Griskevicius et al., 2013; Mittal, Griskevicius, Simpson, Sung, & Young, 2015a; Nettle, Frankenhuys, & Rickard, 2013).

By revealing how specific behaviors respond to external pressures, these individual differences provide unique insights on the underlying cognitive mechanisms and on their development. Building on this literature, in a third

chapter, we adopt an ecological approach, and investigate political behavior by analyzing its variations in different environments. More precisely, we further examine the way political preferences respond to environmental harshness by testing the effect of the early experience of poverty on the relative importance of cooperation- and power-related signals for leader choices ($N = 40$ children, $N = 1000$ adults). We show a both immediate and long-lasting effect of childhood poverty on the preference for strong leaders, i.e., more dominant and less trustworthy ones, providing new evidence of the deep roots of the political responses to external threats. We then discuss this experimental work in the context of an extensive review of the literature on variability of leader preferences in different contexts that is at the basis of a theoretical work on the cognitive mechanisms underlying leader choice.

In summary, throughout this thesis, we demonstrate how the investigation of individual differences in the weighting of two highly ecological social signals (i.e., power-dominance and cooperation-trustworthiness signals) can provide new insights into the cognitive mechanisms underlying social behavior and on their evolutionary bases.

Chapter two. Investigation of an internal variable: study of social motivation

Social motivation as cooperative strategy

Human beings' social motivation is self-evident in cognitive science and countless papers state in their opening lines that '*humans are a highly social species*'. While many non-human animals live in groups and interact with conspecifics, *Homo sapiens* are indeed exceptional in the variety of social interactions they pursue (Kaplan et al., 2009). For humans, social interactions are indeed frequent, diverse and rewarding (Hayden, Parikh, Deaner, & Platt, 2007; Izuma, Saito, & Sadato, 2008; O'Doherty et al., 2003; Pfeiffer et al., 2014; Rademacher et al., 2010; Ruff & Fehr, 2014). Social

stimuli such as faces and speech are granted special attention from birth (Salva, Farroni, Regolin, Vallortigara, & Johnson, 2011; Vouloumanos, Hauser, Werker, & Martin, 2010) and interactive activities are consistently favored over solitary ones by children as young as three (Rekers, Haun, & Tomasello, 2011). In adults, social cues receive attentional priority (Langton, Law, Burton, & Schweinberger, 2008; Lavie, Ro, & Russell, 2003; Ro, Friggel, & Lavie, 2007) and positive social feedback reinforces learning (Chevallier et al., 2016; Jones et al., 2011; A. Lin, Adolphs, & Rangel, 2012- see Appendix). By contrast, ostracism and social rejection are experienced as painful (DeWall & Bushman, 2011; Eisenberger, Lieberman, & Williams, 2003; Kross, Berman, Mischel, Smith, & Wager, 2011) and can trigger intense negative psychological states (Cacioppo & Hawkley, 2009). Strikingly, Epley and Schroeder (2014) demonstrated that even when participants predict that they would prefer social isolation over social interaction, the precise opposite happens. Even when forced then, social interactions are rewarding.

Recently, it has been proposed that social motivation can be construed as an adaptation to collaborative environments, which are characterized by a strong reliance on cooperation to access resources (Chevallier, Kohls, et al., 2012). In this perspective, social motivation would be a coordinated strategy to motivate individuals to take part in collaborative interactions and to allocate enough attention and resources to prevent the fitness costs associated with cooperation, such as being exploited or left alone. This

theory predicts that the more people are socially motivated, the more they should care about others' reliability - to avoid being exploited - and the more they should care about their reputation - to avoid being shunned. In line with this idea, individuals who score high on a measure of social motivation (the social desirability scale) are also more concerned with their reputation and give more to charities as a result (Satow, 1975). On the opposite side of the spectrum, individuals who have diminished social motivation, such as people on the Autism Spectrum Disorder, are less concerned with their reputation (Begeer et al., 2008; Cage, Pellicano, Shah, & Bird, 2013; Chevallier et al., 2014; Chevallier, Grèzes, Molesworth, Berthoz, & Happé, 2012; Izuma, Matsumoto, Camerer, & Adolphs, 2011; Scheeren, Begeer, Banerjee, Terwogt, & Koot, 2010). In the same way, this theory predicts that social motivation should also be associated with differences in approach-avoidance behavior. More precisely, according to this theory, highly socially motivated individuals should grant more importance to cooperation-related than to power-related signals to decide who to approach and who to avoid.

To test this hypothesis, we investigated individual differences in the weight granted to cooperation- and power-related signals in face evaluation in a total sample of 410 participants tested across three experiments with direct replications. We found that social motivation is robustly associated with specific differences in the way power- and cooperation-related signals are combined to produce approach-avoidance judgments. We then modeled

participants' decisions in a trustworthiness-dominance two-dimensional space and found that highly socially motivated individuals are ready to approach a wider spectrum of trustworthy partners, which may be identified as a high-stake cooperative strategy relying on a widening of the cooperation circle to more powerful partners. Our results thus confirmed the link between higher social motivation and higher importance granted to cooperation-related cues predicted by the evolutionary model of social motivation and refined the hypothesis about the cooperative strategy associated with social motivation (Chevallier, Kohls, et al., 2012).

My contribution to this work was as follows: design of the experiment, collection and analysis of the data, and writing of the paper. The manuscript in which the results are reported has been submitted to Nature Human Behavior.

Motivation for social bonding promotes high-stakes cooperative strategies

Safra, L.¹, Wyart, V.¹, Baumard, N.², Chevallier, C.¹

¹ *Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, Inserm, U960, Laboratoire de Neurosciences Cognitives (LNC), F-75005 Paris, France*

² *Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, CNRS, UMR8129, Institut Jean-Nicod, F-75005 Paris, France*

Submitted to *Proceedings of the National Academy Sciences*

Abstract

That humans belong to a highly social species is hardly debated and it is now well-established that being motivated to form solid social bonds enhances fitness across multiple mammal species. Yet, the proximate mechanisms by which such social motivation promotes success in the human ecological niche are mostly underspecified. Here, we demonstrate across five experiments that social motivation is robustly associated with a targeted increase in the weight granted to cooperation-related cues to produce approach-avoidance decisions. Modeling participants' decisions further revealed that motivation for social bonding results in a high-risk / high-reward cooperative strategy relying on a widening of the cooperation circle to include more powerful partners.

Significance statement

The motivation to form social bonds is widely regarded as a fitness enhancing feature in social species. Yet, initiating social interactions also brings about a range of risks. Here we test the hypothesis that increased motivation for social bonding comes with a targeted increase in people's sensitivity to cooperation-related cues, which protects individuals from exploitation risks. High social motivation does not lead individuals to approach more partners indiscriminately; rather, motivation for social bonding is associated with a high-risk / high-reward cooperation strategy involving interactions with more powerful partners.

Main text

While many non-human animals live in groups and interact with conspecifics, *Homo sapiens* are exceptional in the variety of social interactions they pursue¹. For humans, social interactions are indeed frequent, diverse and rewarding²⁻⁸. Individuals are biased to seek positive social interactions and maintain social bonds⁹⁻¹³. There is now a consensus on the fact that the motivation to form social bonds plays a key role in human behaviour and enhances individual fitness^{8,14,15}. However, beyond the fitness advantage conferred by successful and stable social bonds¹⁶, the mechanisms by which social motivation ultimately promote success in the human ecological niche are still underspecified. In particular, all social interactions are not beneficial and individuals constantly face a trade-off between maximizing the number of cooperative interactions and minimizing exploitation risk. Therefore, motivation for social bonding should not operate indiscriminately but rather shape approach-avoidance decisions to maximize the number of *successful* interactions. According to this hypothesis then, social motivation should enhance responsiveness to cues that are relevant to assess the success of cooperative interactions. Highly socially motivated individuals should thus grant more importance to cooperation-related signals - such as trustworthiness, than to power-related cues - such as dominance.

In the present paper, we tested this hypothesis by investigating the weight individuals grant to cooperation-related cues during social decision-making. More precisely, based on previous work showing that approach-

avoidance decisions can be decomposed into cooperation and power evaluations¹⁷⁻²⁰, we compared the weight granted to cooperation-related cues to that granted to power-related cues. Our analysis relied on face evaluations, a major determinant of social interactions²¹. Face evaluations indeed predict who participants trust in economic games, who will be elected in political elections and who juries are most likely to condemn to death sentences²²⁻²⁴. Importantly, studies on face evaluations have shown that the weight granted to cooperation- and power-related cues, i.e. trustworthiness and dominance cues, varies across social contexts, which suggests that people adjust their social decisions flexibly according to the relevance of these cues. For instance, leader preferences are more driven by dominance in war-time, when physical strength is important, than in peace-time^{25,26}. Going one step further, we hypothesized that this flexibility also operates across individuals and that motivation for social bonding is associated with an increase in the importance granted to trustworthiness to produce social decisions.

To test the hypothesis that the motivation to form social bonds is associated with a higher weight granted to cooperation-related cues, we first examined participants' reliance on cooperation- and power-related cues to produce judgments of threat, which are key to produce avoidance decisions. More precisely, we asked 60 participants to rate 40 faces on threat, trustworthiness and dominance (Figure 1A). We reconstructed their two-dimensional threat space, by modelling threat evaluations as a function of perceived dominance and perceived trustworthiness. Following the analysis

plan used in Todorov et al.'s paper²⁷, our model of threat evaluations included linear, quadratic and interaction effects of perceived trustworthiness and perceived dominance.

Results

As can be seen in Figure 1B, faces perceived as less trustworthy and more dominant were rated as more threatening ($b_T = -0.35 \pm 0.05$, $t(2316) = -15.25$, $p < .001$; $b_D = 0.35 \pm 0.04$, $t(2316) = 17.32$, $p < .001$; Figure 1B). In addition, a quadratic effect of dominance indicated that threat evaluations were more sensitive to higher than to lower levels of dominance ($b_{D^2} = 0.21 \pm 0.07$, $t(2316) = 6.18$, $p < .001 = .024$).

In line with our hypothesis, highly socially motivated participants granted more importance to perceived trustworthiness, but only to evaluate faces perceived as dominant ($b_{T*D*SocMot} = -0.17 \pm 0.11$, $t(2316) = -2.97$, $p = .042$; no other significant effect of social motivation was found: all $ps > .119$; Figure 1B). This result thus confirms and refines the hypothesized effect of the motivation for social bonding on approach-avoidance behaviour by revealing that the association between higher levels of social motivation and a higher weight granted to cooperation-related cues might be specific to the evaluation of powerful individuals.

We tested the robustness of this finding in two ways²⁸. Firstly, by replicating the same experiment on 30 participants tested in the lab using avatar faces varying in dominance and trustworthiness, and then by extending this result to the opposite end of the approach-avoidance continuum through the

investigation of likeability evaluations (60 participants).

As expected, the association between the motivation to form social bonds and an increased weight granted to trustworthiness to evaluate dominant faces was confirmed in the experiment on threat evaluation conducted in lab ($b_{T^*D^*SocMot} = -0.31 \pm 0.14$, $t(2359) = -4.26$, $p = .006$). The reconstruction of the two-dimensional likeability space replicated the general approach-avoidance pattern found in the threat evaluation studies, with more trustworthy and less dominant faces perceived as more likeable ($b_T = 0.38 \pm 0.03$, $t(2317) = 21.87$, $p < .001$; $b_D = -0.34 \pm 0.03$, $t(2317) = -22.22$, $p < .001$) and likeability ratings being more sensitive to higher than to lower levels of dominance ($b_{D^2} = -0.15 \pm 0.08$, $t(2317) = -5.40$, $p < .001$; Figure 1C). Even more importantly, the effect of the motivation for social bonding on the combination of perceived trustworthiness and perceived dominance was also evidenced in likeability evaluations ($b_{T^*D^*SocMot} = 0.19 \pm 0.09$, $t(2317) = 3.93$, $p < .001$; Figure 1C). Finally, a meta-analysis conducted on all three social evaluation experiments further confirmed the association between high levels of social motivation and a higher weight granted to trustworthiness for the evaluation of dominant faces ($b_{T^*D^*SocMot} = -0.21 \pm 0.07$, $z = 6.33$, $p < .001$).

In order to better understand the functional consequences of these weighting differences on approach-avoidance behaviour, we ran post-hoc analyses on the predictions of the meta-analytic model. We found that social motivation was associated with an increase in the range of trustworthy faces rated as approachable ($b_{SocMot} = 0.78 \pm 0.15$, $z = 10.12$, $p < .001$),

which was driven by a higher propensity to approach dominant and trustworthy individuals ($b_{\text{SocMot}} = 0.95 \pm 0.17$, $z = 11.01$, $p < .001$; all the predicted proportions of submissive and trustworthy faces rated as approachable were equal to one, model p-value: $p > .250$; Figure 1C).

To further examine these functional consequences, we directly measured participants' preference for dominant and trustworthy faces. To do so, we asked 60 participants to choose their preferred face in pairs of avatars parametrically varying in dominance and trustworthiness²⁹ (Figure 1D). Using a logistic regression on participants' responses, we modelled the probability of choosing a more trustworthy and more dominant face over a more untrustworthy and more submissive one^{30,31}.

This experiment confirmed that highly socially motivated participants had a higher probability of preferring a more dominant and more trustworthy face to a less dominant and less trustworthy face ($b_{\text{SocMot}} = 0.01 \pm 0.01$, $z = 2.67$, $p = .001$; Figure 1E). We assessed the robustness of this result by conducting a replication of this task in a larger sample of participants ($N = 200$). As in the original study, motivation for social bonding was associated with a higher probability of preferring the more trustworthy and more dominant face in this replication study ($b_{\text{SocMot}} = 0.01 \pm 0.01$, $z = 3.28$, $p < .001$) as well as in the meta-analysis conducted on these two experiments ($b_{\text{SocMot}} = 0.01 \pm 0.01$, $z = 4.11$, $p < .001$; Figure 1E-F).

Discussion

Across five studies, we confirmed that motivation for social bonding is associated with a higher weight granted to cooperation-related cues, i.e., trustworthiness cues, during social-decision making. Our results thus provide support for the idea that social motivation is associated with specific behavioural differences that favour successful cooperative interactions⁸.

Importantly, this weighting difference results in a specific distortion of subjects' approach-avoidance two-dimensional space, such that highly socially motivated participants are more likely to approach partners that are perceived as both dominant and trustworthy. This implies that highly socially motivated individuals are more likely to approach a larger range of trustworthy partners (i.e. both submissive, lowly powerful, and dominant, highly powerful ones). One interpretation of this finding is that in doing so, highly socially motivated individuals maximize the number of successful cooperative activities they engage in and, thereby, their fitness in cooperative environments.

However, this widening of the cooperative circle to individuals who are both dominant and trustworthy raises the question of the value of these individuals as cooperative partners. Indeed, interactions with dominant individuals, i.e., physically and socially powerful ones³²⁻³⁵, are not neutral compared to interactions with less dominant partners. More precisely, interacting with socially powerful individuals can provide immaterial benefits, such as social status^{36,37}, and physical power sometimes constitutes

an important lever for resource acquisition³⁸⁻⁴⁰. Crucially however, cooperation with dominant individuals also brings about undeniable risks. For one thing, retaliation against a highly powerful partner is more costly than retaliation against someone who does not hold any power^{41,42}. Importantly, unlike trustworthiness, dominance is not reliably associated with particular cooperative tendencies⁴³⁻⁴⁷. Therefore, interactions with dominant individuals bear higher stakes, but not lower probabilities of success, than interactions with submissive individuals.

The approach behaviour of highly socially motivated individuals towards partners that are perceived as both cooperative and powerful may thus be construed as a high-stake cooperation strategy. Importantly, this strategy may be particularly successful to obtain larger amounts of resources through cooperation. Precise investigations of this hypothesis would help to further understand the adaptive value of the motivation to form social bonds in cooperative environments. In addition, and more generally, our results open the path to the analysis of the adaptive value of different social strategies, depending on the reliance on cooperation for resource acquisition in the environment as well as on the level of resources individuals can invest in cooperation. Indeed, while cooperation and cooperative tendencies have often been investigated as a trait (notably by self-reports of trust or economic games with unknown partners⁴⁸⁻⁵⁰), our results suggest that measuring cooperation strategies through partner choice as a combination of potential stakes, i.e., power-related cues, and

probability of success, i.e., cooperation-related cues, can provide new insights into individuals' cooperative strategies.

Material and methods

All the presented studies received ethics approval from the local ethics committee and each participant received a description of the study and signed an informed consent before starting the experiment.

Social evaluations studies

Experimental Procedure

Following Oosterhof and Todorov's methodology, the questions bearing on the three traits of interest, i.e. trustworthiness, dominance or threat/likeability, were presented in different blocks simultaneously with the face¹⁷. Participants had to answer: "How [trait] is this person?" using a cursor on a 9-point scale ranging from 1 "not at all [trait]" to 9 "extremely [trait]" (recoded from -1 to +1 for the analyses). Depending on the block, [trait] was replaced with "trustworthy", "dominant" or 'threatening'/'likeable'. Participants were instructed to answer following their first impression and they were told that there was no right or wrong answer. The cursor was initially set to the middle in order to reinforce the salience of the positive and the negative sides of the scale. If the participant wanted to select the middle of the scale, she still had to click on the scale to generate a valid response. The name of the dimension was displayed in each trial. At the end of the experiment, participants completed the Social

Anhedonia Scale^{51,52}, a 40-item self-report questionnaire measuring lack of pleasure during social interactions. The Social Anhedonia Scale has been widely used and is considered an adequate tool to measure individual differences in the motivation to form social bonds in the general population^{8,53}. The internal consistency of the Social Anhedonia Questionnaire was high for all the three studies (all α s > .81⁵⁴). The entire procedure lasted approximately 20 minutes.

The online experiments were programmed on Qualtrics (<http://www.qualtrics.com>) and the experiment ran in the lab was programmed on ePrime (Psychology Software Tools, 2002). In the first threat evaluation experiment, participants rated 40 faces (20 woman faces) from the Karolinska database⁵⁵. 80 faces generated with FaceGen 3.1 (<http://www.facegen.com>)¹⁷ replaced these photographs in the threat replication experiment. The number of faces was doubled in this experiment to obtain the same total number of trials as in the original threat evaluation experiment. 20 maximally distinct source identities were randomly selected from the face database available on Todorov's website (<http://tlab.princeton.edu>). Four variations of each identity were used: extremely dominant, extremely submissive, extremely trustworthy and extremely untrustworthy¹⁷. Because we took advantage of an ongoing study, a memorization task using the same faces preceded this last evaluation experiment. Finally, 40 faces varying parametrically in dominance and trustworthiness generated using FaceGen 3.1 were used in the likeability evaluations experiment. Previous work has demonstrated that

these faces elicit dominance and trustworthiness judgments both at the explicit and the implicit level^{56,57}.

Analyses

To investigate the impact of social motivation on the composition of threat/likeability judgments, we ran a mixed linear regression on threat/likeability evaluations, taking social motivation, ratings of dominance and ratings of trustworthiness as predictors and participants' ID as a random factor. Following Todorov et al.²⁷, this model included linear and quadratic interaction effects of perceived trustworthiness and perceived dominance as well as interaction terms between these two factors. The p-values reported for the social motivation effects are exact p-values were obtained via a permutation test ran over 1000 random reassignments of the motivation for social bonding scores. The three-way interaction between social motivation, trustworthiness and dominance was meta-analysed using a fixed effect model⁵⁸.

Finally, as a post-hoc analysis, we estimated the influence of social motivation on the proportion of trustworthy faces rated as approachable (likeable or not threatening) using the parameter estimates of the meta-analytic model (Table S1). To take into account the uncertainty of the coefficients estimates, for each participant, the proportion of trustworthy faces rated as approachable was computed using coefficients randomly drawn from normal distributions around the model's coefficients estimates, with deviations equal to the coefficients estimates' standard deviations. The

impact of social motivation on the proportion of trustworthy faces rated as approachable was then assessed using a beta regression. Because this simulation method may produce different results depending on the coefficients used for computing the probabilities, we repeated this procedure 100 times to ensure that our overall results would not be influenced by the characteristics of each simulation. We averaged the results of the beta regressions (the impact of social motivation on the proportion of faces rated as approachable and the associated z statistic) across the 100 simulations.

Participants

Threat Evaluation Original Study: 60 American participants (28 women), aged 23 to 65 years old (34.92 ± 2.57 years), were recruited via Amazon Mechanical Turk (MTurk, <http://www.mturk.com>), which offers a large diversity of participants^{59,60}. The number of participants was fixed *a priori* based on our past experience with online and social cognition experiments. Each participant received 2\$ for completing the 20-minute task (which corresponds to the prorated average 6\$/hour compensation commonly used on MTurk). Reaction times inferior to 200 ms were discarded for not reflecting a complete processing of the faces (percentage of excluded trials: 0.58%; mean reaction time before data cleaning: 4.62 ± 0.46 s). Analyses were performed on at least 75% of the faces for all the participants. Finally, the motivation for social bonding scores were scaled between -1 and +1.

Threat Evaluation Replication Study: 30 French participants (18 women), aged 18 to 35 years old (25.43 ± 0.10 years) were recruited via an ad posted on a university mailing list. The number of participants was reduced for this study because of material constraints. Participants received 10€ for completing this study, which is the standard payment for studies conducted at the École Normale Supérieure. As in the original study, trials with a reaction time inferior to 200 ms were discarded (percentage of excluded trials: 0.01%; mean reaction time before data cleaning: 4.6 ± 0.59 s). After data cleaning, all participants had analysable responses on at least 99% of the faces. Following the same procedure as for the original study, the motivation for social bonding scores were scaled between -1 and +1.

Likeability Evaluation Study: The number of subjects for this experiment was fixed *a priori* to 60, based on the online threat evaluations experiment. 60 American participants (25 women), aged 19 to 66 years old (34.52 ± 2.63 years), participated in this online study via Amazon Mechanical Turk. Each participant received 2\$ for completing the 20-minute task. As in the threat evaluation studies, reaction times below 200 ms were discarded (percentage of excluded trials: 0.54%; mean reaction time before data cleaning: 5.05 ± 0.71 s). After data cleaning, all participants had analysable responses on at least 92% of the faces. As in the previous studies, the motivation for social bonding scores were scaled between -1 and +1.

Social preferences studies

Experimental Procedure

The experiment was programmed on Qualtrics. 16 avatar faces varying parametrically in dominance and trustworthiness similar to those used in the likeability evaluations studies were used for this experiment. In the original experiment, the 16 faces corresponded to all the possible combinations of dominance and trustworthiness in a range of -3 to +3 points with an increment of 2 points (1 point corresponds to 1 standard deviation in Oosterhof and Todorov's model¹⁷). The faces were regularly spaced on both the dominance and the trustworthiness dimensions. In each presented pair, the faces were 2 to 6 points different from each other on at least one dimension. This resulted in 120 pairs of faces. In the replication study, only 8 avatar faces were used, corresponding to all the possible combinations of dominance and trustworthiness in a range of -2 to +2 points with an increment of 2 points, resulting in a total of 36 pairs of faces.

Each trial began with a central fixation cross presented for 300 ms, then the two faces were presented simultaneously. Participants had to select their preferred face and had up to 2 seconds to answer by pressing "e" for the face on the right and "p" for the face on the left (Figure 1D). Each trial was followed by a blank page presented for about 500 ms. If they failed to answer within 2 seconds, the next trial was automatically presented. All the possible pairs of faces were presented in a random order. The presentation position (right or left) of the faces was randomized between participants. In

the original experiment, the trials were separated into three blocks of 40 trials each.

As in the social evaluation studies, participants completed the Social Anhedonia Scale⁵¹ at the end of the experiment (high internal consistency: original study: $\alpha = .94$; replication study: $\alpha = .94$ ⁵⁴).

Analyses

As recommended by McFadden³⁰, choices were analysed using a mixed logit logistic regression, taking subject ID, trial number and face position as random factors. As each combination of dominance and trustworthiness was presented only once, no quadratic effect of trustworthiness and dominance and no interaction between trustworthiness and dominance was added to this model, such that the probability of choosing one face in a pair was equal to 1 minus the probability to choose the other face of the pair. Levels of trustworthiness and dominance as well as participants' social motivation were used as regressors in the logistic model (Table S2). Based on these coefficients, we estimated the probability of choosing a more trustworthy but more dominant face for each level of social motivation⁶¹. As for the estimation of the proportion of trustworthy faces rated as approachable, we ran 100 simulations using, for each participant, coefficients randomly drawn from normal distributions around the model's coefficients estimates, with deviations equal to the coefficients estimates' standard deviations. For each simulation, the effect of social motivation on choice probability was assessed using a beta-regression. The presented results correspond to the

coefficient estimate and the associated z-value averaged across the 100 simulations. Finally, a meta-analysis on the association between the motivation to form social bonds and the probability of choosing the more dominant and more trustworthy face was conducted using a weighted z test using squared roots of sample sizes as weights⁶².

Participants

Original Study: 60 American participants (28 women) aged 19 to 67 years old (31.80 ± 2.49 years), were recruited via Amazon Mechanical Turk. Each participant received 2\$ for completing the 20-minute task.

Two participants were removed from the analyses for pressing the same key in more than 90% of the trials. As previously, reaction times below 200 ms were discarded (percentage of excluded trials: 18.18%; mean reaction time of the included participants before filtering: 1.07 ± 0.84 s; mean percentage of analysed trials for the analysed participants: 83%). The analyses were performed on 58 participants. Finally, the motivation for social bonding scores were scaled to range between -1 and +1.

Replication Study: 200 American participants (95 women), aged 19 to 68 years old (34.27 ± 1.51 years), were recruited using Amazon Mechanical Turk. This sample size was defined based on previous work on the stability of behavioural differences associated with personality traits⁶³. Participants received 1.5\$ for completing the 5-minute task.

Four participants were removed from the analyses for suspicion of having already completed the original preference study based on their IP address. Nine participants were removed from the analyses for pressing the same key more than 90% of the trials. As in the original study, reaction times were filtered to leave out reaction times shorter than 200 ms (percentage of excluded trials: 18.03%; mean RT before filtering: 1.05 ± 0.39 s) and the motivation for social bonding scores were scaled to range between -1 and +1. The analyses were performed on 187 participants.

Acknowledgments

This study was supported by the Institut d'Études Cognitives (ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL). We are also grateful to Dr Bahador Bahrami for providing us the stimuli.

References

1. Kaplan, H. S., Hooper, P. L. & Gurven, M. The evolutionary and ecological roots of human social organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3289–3299 (2009).
2. Hayden, B. Y., Parikh, P. C., Deaner, R. O. & Platt, M. L. Economic principles motivating social attention in humans. *Proc. R. Soc. Lond. B Biol. Sci.* 274, 1751–1756 (2007).
3. Izuma, K., Saito, D. N. & Sadato, N. Processing of Social and Monetary Rewards in the Human Striatum. *Neuron* 58, 284–294 (2008).
4. O’Doherty, J. et al. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41, 147–155 (2003).
5. Rademacher, L. et al. Dissociation of neural networks for anticipation and consumption of monetary and social rewards. *NeuroImage* 49, 3276–3285 (2010).
6. Pfeiffer, U. J. et al. Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage* 101, 124–137 (2014).
7. Ruff, C. C. & Fehr, E. The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* 15, 549–562 (2014).
8. Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S. & Schultz, R. T. The social motivation theory of autism. *Trends Cogn. Sci.* 16, 231–239 (2012).
9. Chevallier, C. et al. Measuring Social Motivation Using Signal Detection and Reward Responsiveness. *PLOS ONE* 11, e0167024 (2016).

10. Pizzagalli, D. A., Jahn, A. L. & O'Shea, J. P. Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biol. Psychiatry* 57, 319–327 (2005).
11. Lin, A., Adolphs, R. & Rangel, A. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7, 274–281 (2012).
12. Jones, R. M. et al. Behavioral and Neural Properties of Social Reinforcement Learning. *J. Neurosci.* 31, 13039–13045 (2011).
13. Satow, K. L. Social approval and helping. *J. Exp. Soc. Psychol.* 11, 501–509 (1975).
14. Platt, M. L., Seyfarth, R. M. & Cheney, D. L. Adaptations for social cognition in the primate brain. *Phil Trans R Soc B* 371, 20150096 (2016).
15. Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. In Search of the Uniquely Human. *Behav. Brain Sci.* 28, 721–727 (2005).
16. Seyfarth, R. M. & Cheney, D. L. The Evolutionary Origins of Friendship. *Annu. Rev. Psychol.* 63, 153–177 (2011).
17. Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proc. Natl. Acad. Sci.* 105, 11087–11092 (2008).
18. Todorov, A., Said, C. P., Engell, A. D. & Oosterhof, N. N. Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* 12, 455–460 (2008).
19. Sutherland, C. A. M. et al. Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition* 127, 105–118 (2013).

20. McAleer, P., Todorov, A. & Belin, P. How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE* 9, e90779 (2014).
21. Todorov, A., Olivola, C. Y., Dotsch, R. & Mende-Siedlecki, P. Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annu. Rev. Psychol.* 66, 519–545 (2015).
22. Bonnefon, J.-F., Hopfensitz, A. & De Neys, W. The modular nature of trustworthiness detection. *J. Exp. Psychol. Gen.* 142, 143–150 (2013).
23. Todorov, A., Mandisodza, A. N., Goren, A. & Hall, C. C. Inferences of Competence from Faces Predict Election Outcomes. *Science* 308, 1623–1626 (2005).
24. Wilson, J. P. & Rule, N. O. Facial Trustworthiness Predicts Extreme Criminal-Sentencing Outcomes. *Psychol. Sci.* 26, 1325–1331 (2015).
25. Little, A. C., Burriss, R. P., Jones, B. C. & Roberts, S. C. Facial appearance affects voting decisions. *Evol. Hum. Behav.* 28, 18–27 (2007).
26. Spisak, B. R., Homan, A. C., Grabo, A. & Van Vugt, M. Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *Leadersh. Q.* 23, 273–280 (2012).
27. Todorov, A., Said, C. P., Oosterhof, N. N. & Engell, A. D. Task-invariant Brain Responses to the Social Value of Faces. *J. Cogn. Neurosci.* 23, 2766–2781 (2011).
28. Collaboration, O. S. Estimating the reproducibility of psychological science. *Science* 349, aac4716 (2015).

29. Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N. & Falvello, V. B. Validation of data-driven computational models of social perception of faces. *Emotion* 13, 724–738 (2013).
30. McFadden, D. Econometric Models for Probabilistic Choice Among Products. *J. Bus.* 53, S13–S29 (1980).
31. Safra, L. et al. Childhood harshness predicts long-lasting leader preferences. *Evol. Hum. Behav.* (2017). doi:<https://osf.io/v7k73/>
32. Buss, D. M. Evolutionary Psychology: A New Paradigm for Psychological Science. *Psychol. Inq.* 6, 1–30 (1995).
33. Lefevre, C. E. et al. Facial Width-To-Height Ratio Relates to Alpha Status and Assertive Personality in Capuchin Monkeys. *PLOS ONE* 9, e93369 (2014).
34. Lewis, G. J., Lefevre, C. E. & Bates, T. C. Facial width-to-height ratio predicts achievement drive in US presidents. *Personal. Individ. Differ.* 52, 855–857 (2012).
35. Wilson, V. et al. Personality and facial morphology: Links to assertiveness and neuroticism in capuchins (*Sapajus [Cebus] apella*). *Personal. Individ. Differ.* 58, 89–94 (2014).
36. Lin, N. Social Networks and Status Attainment. *Annu. Rev. Sociol.* 25, 467–487 (1999).
37. von Rueden, C. The Roots and Fruits of Social Status in Small-Scale Human Societies. in *The Psychology of Social Status* (eds. Cheng, J. T., Tracy, J. L. & Anderson, C.) 179–200 (Springer New York, 2014). doi:10.1007/978-1-4939-0867-7_9

38. Kaufmann, J. H. On the Definitions and Functions of Dominance and Territoriality. *Biol. Rev.* 58, 1–20 (1983).
39. Stockley, P. & Bro-Jørgensen, J. Female competition and its evolutionary consequences in mammals. *Biol. Rev.* 86, 341–366 (2011).
40. von Rueden, C., Gurven, M. & Kaplan, H. Why do men seek status? Fitness payoffs to dominance and prestige. *Proc. R. Soc. Lond. B Biol. Sci.* rspb20102145 (2010). doi:10.1098/rspb.2010.2145
41. Carré, J. M., McCormick, C. M. & Mondloch, C. J. Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychol. Sci.* 20, 1194–1198 (2009).
42. Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P. & Penton-Voak, I. S. Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biol. Lett.* 10, 20140729 (2014).
43. Boksem, M. A. S. et al. Testosterone Inhibits Trust but Promotes Reciprocity. *Psychol. Sci.* 0956797613495063 (2013). doi:10.1177/0956797613495063
44. Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M. & Fehr, E. Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature* 463, 356–359 (2009).
45. Stirrat, M. & Perrett, D. I. Valid Facial Cues to Cooperation and Trust Male Facial Width and Trustworthiness. *Psychol. Sci.* 21, 349–354 (2010).
46. van Honk, J., Montoya, E., Bos, P., Vugt, M. van & Terburg, D. New evidence on testosterone and cooperation. 485 (2012).

47. Wibral, M., Dohmen, T., Klingmüller, D., Weber, B. & Falk, A. Testosterone Administration Reduces Lying in Men. *PLOS ONE* 7, e46774 (2012).
48. McCullough, M. E., Pedersen, E. J., Schroder, J. M., Tabak, B. A. & Carver, C. S. Harsh childhood environmental characteristics predict exploitation and retaliation in humans. *Proc. R. Soc. B Biol. Sci.* 280, 20122104 (2013).
49. Nettle, D., Colléony, A. & Cockerill, M. Variation in Cooperative Behaviour within a Single City. *PLOS ONE* 6, e26922 (2011).
50. Petersen, M. B. & Aarøe, L. Birth Weight and Social Trust in Adulthood: Evidence for Early Calibration of Social Cognition. *Psychol. Sci.* 26, 1681–1692 (2015).
51. Chapman, L. J., Chapman, J. P. & Raulin, M. L. Scales for physical and social anhedonia. *J. Abnorm. Psychol.* 85, 374–382 (1976).
52. Assouly-Besse, F., Dollfus, S. & Petit, M. Traduction française des questionnaires d'anhédonie sociale et physique de Chapman : validation de la traduction française à partir de témoins et de patients schizophrènes. in *L'Encéphale* 21, 273–284 (Elsevier Masson, 1995).
53. Berthoz, S., Lalanne, C., Crane, L. & Hill, E. L. Investigating emotional impairments in adults with autism spectrum disorders and the broader autism phenotype. *Psychiatry Res.* 208, 257–264 (2013).
54. Santos, J. Cronbach's alpha: A tool for assessing the reliability of scales. *J. Ext.* 37, 1–5 (1999).

55. Lundqvist, D., Flykt, A. & Öhman, A. The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.* (1998).
56. Stewart, L. H. et al. Unconscious evaluation of faces on social dimensions. *J. Exp. Psychol. Gen.* 141, 715–727 (2012).
57. Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N. & Falvello, V. B. Validation of data-driven computational models of social perception of faces. *Emot. Wash. DC* 13, 724–738 (2013).
58. Berkey, C. S., Hoaglin, D. C., Mosteller, F. & Colditz, G. A. A random-effects regression model for meta-analysis. *Stat. Med.* 14, 395–411 (1995).
59. Medin, D., Ojalehto, B., Marin, A. & Bang, M. Systems of (non-)diversity. *Nat. Hum. Behav.* 1, 0088 (2017).
60. Buhrmester, M., Kwang, T. & Gosling, S. D. Amazon's Mechanical Turk. *Perspect. Psychol. Sci.* 6, 3–5 (2011).
61. Long, J. S. *Group comparisons in logit and probit using predicted probabilities 1.* (2009).
62. Zaykin, D. V. Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J. Evol. Biol.* 24, 1836–1841 (2011).
63. Schönbrodt, F. D. & Perugini, M. At what sample size do correlations stabilize? *J. Res. Personal.* 47, 609–612 (2013).

Figure

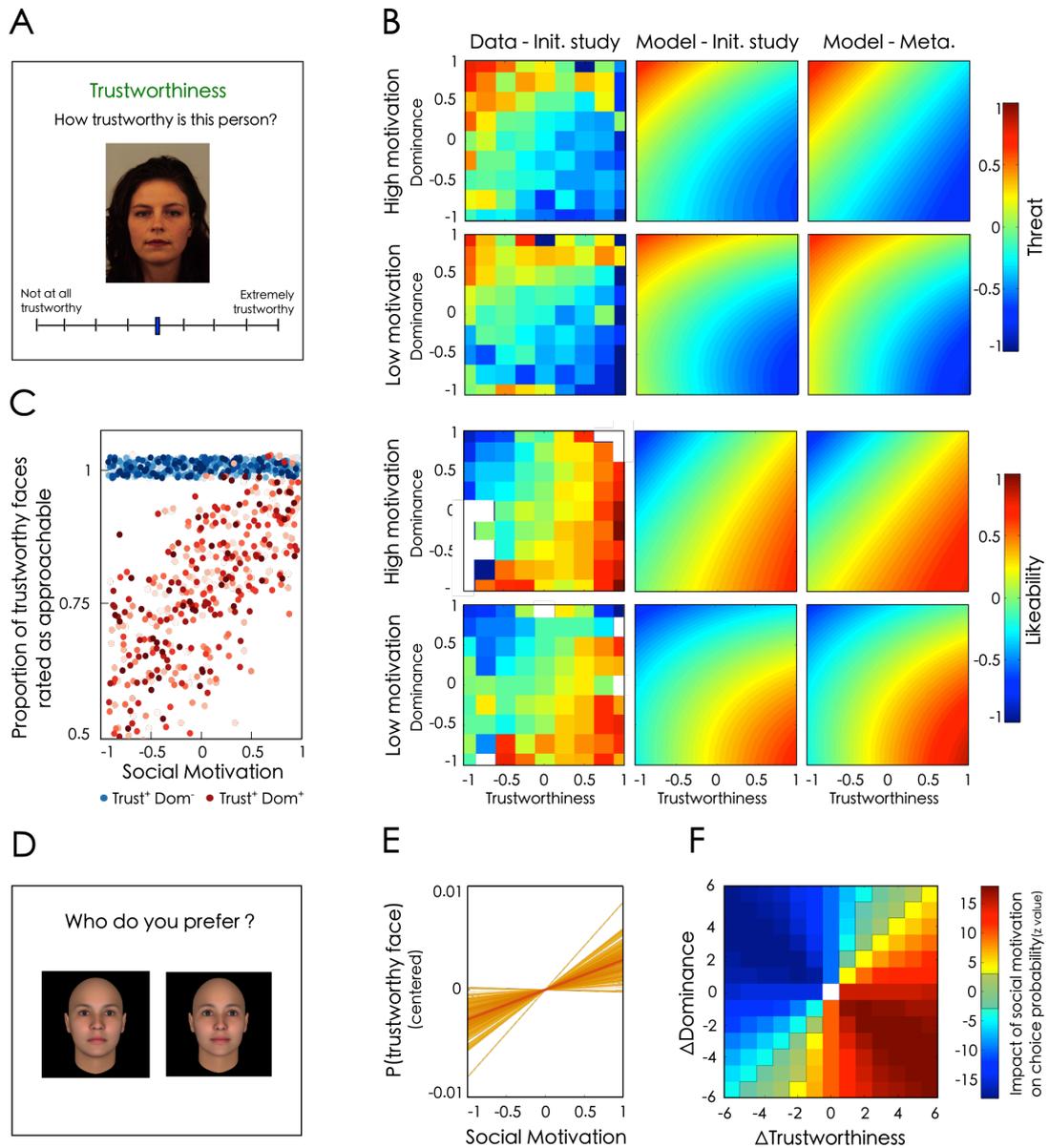


Figure 1 – Motivation for social bonding is associated with an increase in the relative weight granted to trustworthiness during social decisions

(A) Example of an evaluation trial. Participants had to rate each face by moving a cursor, initially positioned in the middle of the scale. **(B) Threat (top) and**

Likeability ratings (bottom) as a function of trustworthiness (x axis) and dominance ratings (y axis). Rating intensity is represented on a scale ranging from blue for lower ratings to red for higher ratings. Pixelized figures correspond to averaged data in the initial study for the most (upper row) and least (lower row) socially motivated participants (median split). Smoothed figures represent the predictions of the regression models obtained in the initial study (Model–Init study) and in the meta-analysis (Model–Meta.). Faces perceived as both dominant and trustworthy were rated as more approachable by highly socially motivated participants. **(C) Effect of the motivation for social bonding on the evaluation of trustworthy faces.** Social motivation was associated with an increase in the proportion of dominant and trustworthy ($\text{Trust}^+\text{Dom}^*$, in red) but not of submissive and trustworthy faces ($\text{Trust}^+\text{Dom}^-$, in blue) perceived as approachable. **(D) Example of a preference trial. Both faces appeared simultaneously on the screen.** Participants had up to 2 seconds to answer by pressing “e” to select the face on the left and “p” to select the face on the right. **(E) Computed probability of choosing a 1-point more trustworthy and 1-point more dominant face as a function of the motivation for social bonding.** 100 simulations (light lines) revealed that the probability of choosing a more trustworthy but more dominant face increased with social motivation in both the original study and its replication (the bold regression line corresponds to the meta-analysis). **(F) Impact of the motivation for social bonding on choice probability for every combination of differences in trustworthiness (x axis) and dominance (y axis) in the meta-analysis.** Meta-analytic z values are represented on a scale ranging from blue for negative values to red for positive values. Shaded areas correspond to z-values below the 5% threshold of statistical significance after applying a Bonferroni

correction for 169 comparisons. Social motivation is associated with a higher probability of preferring the most dominant and trustworthy faces.

Supplementary analyses

Cues decoding:

In order to assess whether participants actually computed the presented stimuli, we checked whether their ratings were in accordance to the avatars' levels of trustworthiness and dominance (replication study on threat evaluations and the study on likeability evaluations). To measure participants' ability to decode trustworthiness and dominance cues, we ran mixed linear regressions on trustworthiness and dominance ratings, taking avatars' levels of trustworthiness / dominance as regressors and participants' ID as a random factor. As expected, intensity levels of both trustworthiness and dominance cues were successfully perceived in the two experiments (trustworthiness: all $ps < .001$; dominance: all $ps < .001$). In addition, highly socially motivated participants did not differ in their ability to detect these cues (meta-analytic effect across the two studies: both $ps > .250$; p-values obtained via a permutation test) and were not biased to rate faces as more trustworthy (meta-analytic effect across the two studies: both $ps > .250$).

Reaction times:

To check that differences in cue combination was not due to differences in cue processing, we measured the influence of social motivation on reaction times in the Preference studies. Reaction times were analysed to assess if social motivation affected the cognitive processes underlying the combination of social cues. This analysis was performed using a mixed

linear regressions on the population-level scaled reaction times, taking as regressors: social motivation, the absolute value of the difference in trustworthiness between the faces of each pair ($|\Delta\text{Trustworthiness}| = |\text{Trustworthiness}_{\text{Left}} - \text{Trustworthiness}_{\text{Right}}|$) and the absolute value of the difference in dominance between the faces of each pair ($|\Delta\text{Dominance}| = |\text{Dominance}_{\text{Left}} - \text{Dominance}_{\text{Right}}|$). This analysis revealed that social motivation was not associated with any difference in the speed at which social cues are processed (meta-analytic effect across the two studies: all p s > .192; p -values obtained via a permutation test), which suggests that the effect of social motivation was circumscribed to the way social cues are used to produce social judgments.

Supplementary figures

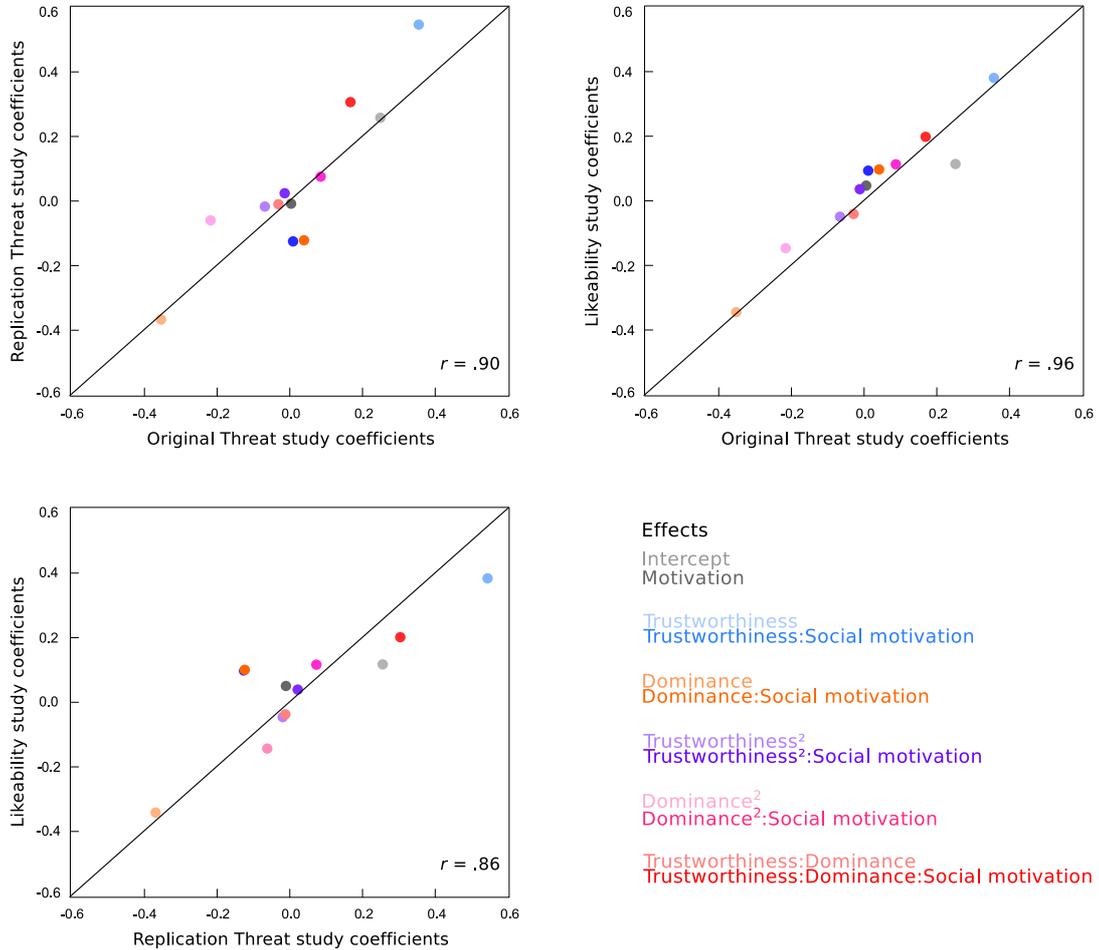


Figure S1 – Replication of the approachability evaluation models

Parameters from the different evaluation studies significantly correlate with each others (threat original – threat replication: $r = .90$, $N = 12$, $p < .001$; threat original - likeability: $r = .96$, $N = 12$, $p < .001$; threat replication – likeability: $r = .86$, $N = 12$, $p < .001$; the black line represents perfect equality between the two studies' coefficients), revealing the robustness of the results.

Supplementary Tables

Table S1 – Meta-analytic value of the coefficient parameters of the approachability evaluations models

Model parameters	Coefficients estimates
Intercept	0.20 ± 0.10
Trustworthiness	0.43 ± 0.12
Dominance	-0.35 ± 0.02
Trustworthiness ²	-0.05 ± 0.04
Dominance ²	-0.15 ± 0.08
Social Motivation	0.03 ± 0.06
Trustworthiness:Dominance	-0.03 ± 0.04
Trustworthiness:Social Motivation	-0.01 ± 0.13
Dominance:Social Motivation	0.01 ± 0.13
Trustworthiness ² :Social Motivation	0.02 ± 0.07
Dominance ² :Social Motivation	0.10 ± 0.06
Trustworthiness:Dominance:Social Motivation	0.21 ± 0.07

Meta-analytic value of the regression coefficients and their 95% confidence interval.

Table S2 – Logistic regression coefficients obtained in the preference studies

Parameter	Original study (N = 58)	Replication study (N = 187)
Trustworthiness	0.31 ± 0.02	0.38 ± 0.03
Dominance	-0.19 ± 0.02	-0.18 ± 0.03
Trustworthiness:Social motivation	-0.02 ± 0.04	0.05 ± 0.04
Dominance:Social motivation	0.03 ± 0.04	-0.04 ± 0.04

Coefficient estimates of the logistic regressions used to compute the probability of choosing a more dominant and more trustworthy face and their 95% confidence interval.

Effect of social motivation across populations

These results provide the first evidence of the effect of social motivation on the importance of cooperation-related cues for social decisions. However, the evolutionary model of social motivation predicts that social motivation influences social behavior starting from the age at which resource acquisition starts to rely on cooperation with peers, i.e., from childhood (Sheskin, Chevallier, Lambert, & Baumard, 2014). The effects of social motivation found on adult participants should thus also be present in younger individuals.

In addition, one can wonder whether the weighting differences of cooperation-related cues associated with social motivation can provide insights into the behavioral atypicalities of clinical populations characterized by diminished social motivation such as autism spectrum disorders (Chevallier, Kohls, et al., 2012; Klin, Jones, Schultz, & Volkmar, 2003; Mundy, 2003). Indeed, it has recently been argued that dimensional approaches, investigating relevant biological traits across clinical and non-clinical populations, and notably social motivation (National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria, 2016), could be particularly fruitful for understanding neuropsychological conditions (Cuthbert, 2014; Insel et al., 2010; National Institute of Mental Health, 2015; Sanislow et al., 2010).

Using the same protocol as in the previous studies, we investigated these two points by testing whether the effects of social motivation evidenced in adult participants could also be found in 20 typically developed adolescents and in 20 adolescents with autism spectrum disorders. Our results confirmed that social motivation is associated with an increased weight granted to trustworthiness in typically developing adolescents, providing evidence of the robustness of this effect across the lifespan. However, this association was not present in adolescents with autism spectrum disorders. These results question the possibility to directly map effects of psychological traits, such as social motivation, found in the general population to clinical ones (Harvey, Bodnar, Sergerie, Armony, & Lepage, 2009; Olsen, Bjorkquist, Bodapati, Shankman, & Herbener, 2015). In particular, it invites to investigate the mechanisms that could account for the observed discrepancies, such as the interaction of decreased social motivation with comorbid traits (Matson & Nebel-Schwalm, 2007; Simonoff et al., 2008), retroactive loops on social behavior (Chamberlain, Kasari, & Rotheram-Fuller, 2006; Kasari, Locke, Gulsrud, & Rotheram-Fuller, 2010; Locke, Ishijima, Kasari, & London, 2010; Shattuck, Orsmond, Wagner, & Cooper, 2011) or the intrinsic value of clinical cut-offs, in order to better understand the functional bases of neuropsychological conditions.

My contribution to this work was as follows: design of the experiment, analysis of the data and writing of the paper. The attached manuscript has been submitted to Scientific Reports.

Distinct effects of social motivation on face evaluations in adolescents with and without autism

Safra, L.¹, Ioannou, C.¹, Amsellem, F.^{2,3}, Delorme, R.^{2,3}, Chevallier, C.¹

¹ *Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, Inserm, U960, Laboratoire de Neurosciences Cognitives (LNC), F-75005 Paris, France*

² *Service de Psychiatrie de l'Enfant et de l'Adolescent, Hôpital Universitaire Robert Debré, Paris, 75019, France*

³ *Génétique Humaine et Fonction Cognitive, Institut Pasteur, Paris, 75015, France*

Submitted to *Scientific Reports*

Abstract

Individual differences in social motivation have an influence on many behaviours in both clinical and non-clinical populations. As such, social motivation has been identified as a biological trait that is particularly well-suited for dimensional approaches cutting across neuropsychological conditions. In the present paper, we tested whether social motivation had a similar impact in the general population and in a neuropsychological condition characterized by diminished social motivation: Autism Spectrum Disorders (ASD). More precisely, we evaluated the effect of social motivation on face evaluations in 20 adolescents with ASD and 20 matched controls using avatars parametrically varying in dominance and trustworthiness. In line with previous research, we found that social motivation is associated with a larger weight granted to perceived trustworthiness when producing likeability judgments in the control group. However, this pattern was not found in the ASD group. Social motivation thus appears to have a different effect in ASD and control populations, which raises questions about the relevance of subclinical or non-clinical populations to understand ASD.

Main text

Compared to many other animals, humans stand out when it comes to the variety of social interactions they pursue and the importance of social activities in their ecological niche ¹. The willingness to be included in social interactions and the propensity to preferentially attend to the social world is present early on in development and remains a driving force throughout the lifespan ²⁻⁴. Social stimuli such as faces and speech are granted special attention from birth ^{4,5} and interactive activities are consistently favored over solitary ones by children as young as three ⁶. In adults, social cues receive attentional priority ⁷⁻⁹ and positive social feedback reinforces learning ¹⁰⁻¹². Yet, individuals vary in the degree to which they are socially motivated ¹³⁻¹⁸ and atypicalities in social motivation are found in many clinical conditions ^{17,18}. Atypical social motivation is indeed an important characteristic of multiple psychiatric conditions, including Autism Spectrum Disorders (ASD)¹⁹⁻²¹, anorexia nervosa, schizophrenia and major depressive disorder ^{22,23}. Attention to social stimuli is also thought to provide a starting point for the development of social abilities, such as face processing ^{24,25}, and lack of social motivation might thus have cascade effects on other areas of social cognition. In ASD in particular, it has been argued that early deficits in social motivation and social reward responsiveness might have a long lasting impact on social skills ^{17,26,27}.

Recent developments in psychiatry (the Research Domain Criteria Framework) have emphasised the need to investigate variations in relevant biological traits across clinical and non-clinical populations ²⁸⁻³¹. Given its

central role across conditions, social motivation has been identified as a relevant biological trait to investigate in a dimensional framework ³². For instance, Parish-Morris et al. ³³ have shown using such a dimensional approach that individual differences in social attention is a better predictor of face processing skills than being diagnosed with ASD. Importantly however, the Research Domain Criteria Framework emphasizes that a given trait may have non-linear effects on behaviour, as is for example the case with the classic U-shaped curve relating stress and performance. In addition, variations in single traits do not necessarily have a uniform impact when they are taken in isolation and when they are combined in the context of psychiatric conditions ²⁸⁻³¹. Social anhedonia for instance, has a different effect on social cognition in patients with schizophrenia, patients with major depressive disorder and healthy controls ^{18,34}. In the case of ASD, stereotypical interests, anxiety, sensory peculiarities or number of co-frequent morbidities, such as anxiety, might also influence the way diminished social motivation alters individual behaviour ^{35,36}. It is therefore crucial to identify potential points of disjunction at which variations on a given traits affect cognitive functioning differently

The aim of this paper is to apply the insights of dimensional approaches to the study of social motivation in ASD by testing whether social motivation has a uniform effect on individuals with and without ASD ³⁰. To investigate this question, we focused on face evaluation, which is key for social decision making ³⁷, and which is sensitive to variations in social motivation. Specifically, Safra et al. ³⁸ have shown that highly socially motivated adults

place more weight on perceived trustworthiness when producing likeability evaluations. Here, we asked 40 adolescent participants (20 typically developing (TD) and 20 with ASD) to rate faces on likeability using well-controlled stimuli varying parametrically in dominance and trustworthiness^{39,40} (Figure 1A). Our prediction was that higher levels of social motivation would increase the weight granted to trustworthiness during face evaluations in both TD and ASD populations.

Results

Cue processing in the TD and ASD groups

As our measure relied on the processing of facial features, we first checked that both groups were able to accurately detect and combine facial cues. Replicating previous findings^{41,42}, TD adolescents successfully detected trustworthiness ($b = 0.19 \pm 0.03$, $t(579) = 13.947014$, $p < .001$; results are given in the standard form: mean \pm 95% confidence intervals) and dominance cues ($b = 0.16 \pm 0.03$, $t(579) = 10.93$, $p < .001$), giving higher trustworthiness/dominance ratings to avatars presenting higher levels of trustworthiness/dominance. Similarly, in line with previous studies⁴²⁻⁴⁴ adolescents with ASD successfully gave ratings that varied with the avatar's level of trustworthiness and dominance ($b = 0.18 \pm 0.03$, $t(579) = 10.14$, $p < .001$; $b = 0.07 \pm 0.05$, $t(579) = 2.81$, $p = .005$).

Based on Oosterhof and Todorov⁴⁵ and Safra et al.³⁸, we then reconstructed participants' likeability two-dimensional space based on their ratings of dominance and trustworthiness. As can be seen in Figure 1B,

both adolescents with and without ASD combined perceived dominance and trustworthiness to form likeability judgments such that more trustworthy and less dominant faces were rated as more likeable in both groups (TD group: $b_{Trustworthiness} = 0.47 \pm 0.07$, $t(575) = 13.27$, $p < .001$; $b_{Dominance} = -0.32 \pm 0.07$, $t(575) = -9.69$, $p < .001$; ASD group: $b_{Trustworthiness} = 0.52 \pm 0.07$, $t(575) = 14.03$, $p < .001$; $b_{Dominance} = -0.07 \pm 0.07$, $t(575) = -1.87$, $p = .062$).

Impact of social motivation in TD adolescents

We then examine whether, as in adults, social motivation increased the weight granted to trustworthiness for likeability evaluations in TD adolescents. As expected, social motivation increased the relative weight granted to trustworthiness to evaluate faces' likeability ($b_{SocMot*Trust} = 0.08 \pm 0.08$, $t(570) = 2.06$, $p = .039$; Figure 1B). Importantly, this effect was still present after controlling for non-social motivations ($b_{SocMot*Trust} = 0.18 \pm 0.11$, $t(560) = 3.24$, $p = .001$). In addition, higher levels of social motivation were associated with lower sensitivity to high levels of dominance ($b_{SocMot*Dom^2} = 0.14 \pm 0.12$, $t(570) = 2.41$, $p = .016$; after controlling for non-social motivations: $b_{SocMot*Dom^2} = 0.11 \pm 0.10$, $t(560) = 2.29$, $p = .023$). A lower sensitivity to dominance in highly socially motivated participants was also found as a trend ($b_{SocMot*Dom} = 0.06 \pm 0.07$, $t(570) = 1.71$, $p = .088$; after controlling for non-social motivations: $b_{SocMot*Dom} = 0.13 \pm 0.15$, $t(560) = 1.67$, $p = .098$; all other effects, $p > .108$). Given that social motivation was associated with an increase in the perceived intensity of trustworthiness ($b = 0.03 \pm 0.03$, $t(578) = 2.01$, $p = .045$; other effects on cue detection: all $ps >$

.104), we conducted a similar model using avatars' objective dominance and trustworthiness as predictors. This analysis confirmed that social motivation was associated with a larger weight granted to trustworthiness for likeability evaluations ($b_{\text{SocMot*Trust}} = 0.03 \pm 0.01$, $t(570) = 3.42$, $p < .001$; after controlling for non-social motivations: $b_{\text{SocMot*Trust}} = 0.03 \pm 0.02$, $t(560) = 3.345240$, $p < .001$; no other significant effect of social motivation: all $ps > .124$). As a conclusion, social motivation had a similar impact in adolescents as reported in adults³⁸.

Effect of social motivation in ASD

We then tested whether social motivation had a uniform impact across populations by analysing the ASD and TD groups together and including Group (ASD vs TD) as a regressor. The interaction between social motivation and the weight granted to trustworthiness was different in the two groups ($b_{\text{Group*SocMot*Trust}} = -0.17 \pm 0.15$, $t(1140) = -2.18$, $p = .029$; no other difference in the effect of social motivation between the two groups was found: all $ps > .127$), such that social motivation was not associated with an increased weight granted to trustworthiness in the ASD group ($b_{\text{SocMot*Trust}} = 0.05 \pm 0.11$, $t(1140) = 0.97$, $p > .250$). Importantly, controlling for non-social motivations confirmed the difference between the groups ($b_{\text{Group*SocMot*Trust}} = -0.38 \pm 0.21$, $t(1120) = -3.47$, $p < .001$; no other significant interaction between Group and social motivation: all $ps > .250$) and revealed that social motivation was associated with a decrease in the weight granted to trustworthiness for likeability evaluations in the ASD group

($b_{\text{SocMot*Trust}} = -0.10 \pm 0.09$, $t(1120) = -2.17$, $p = .031$). Finally, these findings were replicated using avatars' levels of dominance and trustworthiness instead of participants' subjective evaluations to predict likeability evaluations, suggesting that this difference was not due to differences in the way adolescents evaluate dominance and trustworthiness cues ($b_{\text{Group*SocMot*Trust}} = 0.13 \pm 0.08$, $t(1140) = 2.94$, $p = .003$; after controlling for non-social motivations: $b_{\text{Group*SocMot*Trust}} = 0.19 \pm 0.12$, $t(1120) = 3.05$, $p = .002$; all other interactions between Group and social motivation: all $ps > .130$). To summarize, social motivation does not have a uniform effect across the TD and the ASD groups.

In line with this finding, while the ASD group was less socially motivated than the TD group ($t(38) = -2.43$, $p = .019$), the comparison of these two groups did not match the difference between lowly and highly socially motivated adolescents without ASD. More precisely, compared to TD adolescents, adolescents with ASD perceived dominance cues as less intense ($b = -0.09 \pm 0.06$, $t(1158) = -3.04$, $p = .002$; no other significant difference in cue detection: all $ps > .250$), and granted less weight to dominance for evaluating likeability ($b_{\text{SocMot*Dom}} = 0.26 \pm 0.10$, $t(1150) = 5.03$, $p < .001$; no other significant effect of diagnosis: all $ps > 0.113$). Importantly, this effect was preserved when taking objective cues of dominance and trustworthiness as predictors ($b_{\text{SocMot*Dom}} = 0.05 \pm 0.02$, $t(1150) = 4.62$, $p < .001$), which indicates that the weighting difference between the ASD and TD groups could not be explained by differences in explicit cue evaluation.

Discussion

The aim of this study was to assess whether social motivation had a similar impact in adolescents with and without ASD. Previous work demonstrated that higher levels of social motivation increase the weight granted to trustworthiness cues during likeability evaluations ³⁸. We replicated this effect in TD adolescents. However, social motivation had the opposite effect in the ASD group and was associated with a decrease in the weight granted to trustworthiness. This shows that social motivation can have contradictory effects in clinical and non-clinical populations. Given that ASD participants are overall less socially motivated than typically developing participants ¹⁹, it would be tempting to construe autism as a simple case of extreme diminished social motivation and to use findings describing the effect of low social motivation in the general population as a guide to predict ASD cognition. Our results challenge this view, by revealing that atypicalities found in participants with ASD do not match the behavioural differences associated with diminished social motivation in a non-clinical sample: lowly socially motivated adolescents without ASD displayed a decreased sensitivity to trustworthiness, adolescents with ASD displayed a decreased sensitivity to dominance.

These findings have important implications for the understanding and investigation of ASD. First, our results suggest that the effect of social motivation uncovered in non-clinical populations cannot always be applied to ASD. Social disinterest in ASD may indeed be associated with emergent

properties that cannot be derived by simply extrapolating the effects of mildly diminished social motivation. In addition, long-lasting difficulties in social interactions may have a retroactive action on social behaviour ⁴⁶⁻⁴⁹. Finally, ASD are well-known for being associated with other conditions such as social anxiety and hyperactivity disorders ^{35,36} that may interact with diminished social motivation. In this context, it is important to underline that findings obtained in subclinical or non-clinical populations should be applied to ASD with a great deal of caution.

Regarding the present study, we wish to underline two potential sources of noise. First, self-reports of social motivation might be biased differently in the ASD and the TD group: individuals with ASD indeed have difficulties reporting their own feelings and may lack insight⁵⁰⁻⁵²; conversely TD participants are more likely to be susceptible to social desirability effects ⁵³. Our results should thus be replicated using more objective measures of social motivation ^{10,54}. Second, it is widely recognised that autism should not be construed as a unique neuropsychological condition and that a composite view is needed in order to take into account the existence of distinct partial phenotypes in ASD ^{55,56}. It is thus possible that adolescents in our study belong to different subtypes of ASD that are characterized by different levels of social motivation deficits. Addressing these questions would require a much larger sample size than the one we had access to in the context of this study. Despite these caveats, we believe that our results are relevant for the understanding of social processing in ASD. In particular, we demonstrated that individuals with ASD are able to detect dominance

and trustworthiness cues and to combine these cues to produce likeability evaluation (albeit differently from TD participants). These results are in line with a growing body of evidence showing preserved abilities to detect and to combine social cues in individuals with ASD ^{43,44,57-59}. In addition, our study also extends previous results on social cues combination by showing that individuals with ASD are not only able to modulate their perception of faces by integrating different social signals but also that they are able to create new social judgments by combining social cues. Indeed, the detection of dominance and trustworthiness cues and their combination are processed in distinct brain regions ⁶⁰⁻⁶².

Moreover, the impact of ASD on the way social cues are combined to produce likeability evaluations is particularly robust. Differences in likeability evaluations indeed remained while using both avatars' objective and subjective levels of dominance and trustworthiness, which demonstrates an actual difference in dominance salience when producing likeability evaluations. Overall, our results thus suggest that individuals with ASD are able to detect and use dominance, but they place less weight on this trait compared to TD individuals. Why that might be the case is an empirical question in need of further investigation. Nevertheless, a number of experiments have revealed that more masculine individuals are less sensitive to dominance ^{63,64}. Therefore, our results may be in line with previous research showing an exaggerated male pattern of neural activation in ASD during face evaluation ⁶⁵. More widely, our results are relevant for dimensional approaches in psychiatry beyond the precise case of social

motivation ²⁹⁻³¹. A growing body of research has indeed emphasized the importance of studying the impact of specific traits on behaviour by pooling clinical and non-clinical populations ²⁸⁻³¹. This approach has notably been applied by investigating biologically-relevant traits, such as reward anticipation across populations ⁶⁶; or by applying multi-dimensional diagnostic measures, such as the Autism Quotient, to sub- and non-clinical population ⁶⁷. However, our results suggest that findings obtained in non-clinical samples cannot always be directly mapped onto clinical populations. As suggested for the case of social motivation, it is indeed possible that variations in specific traits affect behaviour non-linearly to produce emergent behavioural peculiarities. Therefore, our findings are in favour of a strict application of the original Research Domain Criteria Framework recommendations emphasizing the importance of investigating biologically relevant traits across conditions and of identifying points of disjunction that may potentially give rise to non-linear effects of a given trait on behaviour ²⁸⁻³¹.

To summarise, our study replicates previous findings obtained in healthy adults showing that social motivation increases the weight granted to trustworthiness to produce likeability judgments. In contrast with our prediction however, social motivation did not have the same impact in ASD and in TD. Despite an overall diminished social motivation in the ASD group, the impact of autism was quite different than the simple effect of social motivation in TD adolescents. This result suggests that it may be misleading to construe social motivation in isolation and that it is vital to

further understand how social motivation interacts with other dimensions of ASD.

Material and Methods

Participants

A minimum target of 20 TD adolescents and 20 adolescents with ASD was fixed *a priori*. The exact number was determined by scheduling constraints. A final number 22 TD adolescents (6 females) and 22 adolescents with ASD (4 females), aged between 12 and 17 years old (TD: $M = 13.70 \pm 0.61$; ASD: $M = 14.45 \pm 0.89$; Table 1), participated in this study. The experiment was approved by the local research ethics committee (ClinicalTrials.gov Identifier: NCT02628808, Protocol Study ID: 2008-A00019-46). The TD adolescents were recruited from a mainstream school and the adolescents with ASD were recruited from the University Hospital Robert Debré (Paris, France). The adolescents with ASD had received an official diagnosis of autism by an independent clinician according to the criteria of the Diagnostic and Statistical Manual for mental disorders-IV TR (DSM IV TR ⁶⁸). The Autism Diagnostic Interview Revised (ADI-R ⁶⁹) and the Autism Diagnostic Observational Schedule (ADOS ⁷⁰) were used to further assess the ASD group. The mean ADOS score for the ASD group was 13.59 ± 1.81 (Table 1). All participants with ASD had normal vision (Freiburg Visual Acuity and Contrast Test version 3.8.2 ⁷¹ adapted to the distance used in the experiment of 30 cm), no participant was on medication during the period of the study. Preliminary interviews confirmed that TD adolescent

participants did not have any special needs or history of psychiatric illness or developmental delay and all of them had normal or corrected to normal vision.

Before testing, all parents and children provided their written informed consent to participate in the study. IQ was measured using the full Wechsler Intelligence Scale for Children version IV (WISC IV ⁷²; Mean = 103.40 ± 10.86; range: 70 – 148; Table 1) in adolescents with ASD and with the Wechsler Abbreviated Scale of Intelligence in TD adolescents due to time constraints (in the four subsets form as it has been shown to give the most representative score of the full IQ, ⁷³; range: 87 – 138; Table 1). Finally, at the end of the experiment, participants completed the Kazdin's Pleasure Scale ⁷⁴, a self-rated questionnaire to assess their levels of anhedonia (see description below). In addition, trait anxiety was assessed using an abbreviated form of the State-Trait Anxiety Inventory (STAI ^{75,76}; Table 1).

Materials and design

The pleasure scale

The Kazdin Pleasure Scale for Children (three subforms: social, physical and other; ⁷⁴; Table 1) was used to assess participants' anhedonia levels. This scale is a validated self-rated instrument to measure anhedonia in both children with and without ASD ¹⁹. It consists of 39 items pertaining to social (e.g., "You accidentally overhear your teacher telling the principal what a terrific student you are"), physical (e.g., "You are cycling down the street very fast while still in good control of yourself") or other sources of pleasure

(e.g., "On a Saturday night, you stay up watching television as long as you want"). Participants were asked to read each item out loud and to rate their feeling in the corresponding situation on a 3-point Likert scale ("Very happy", "Happy" or "Neither happy nor unhappy"). The three scales were reverse-coded in order to reflect participants' levels of motivation.

The Face Evaluation Task

The experiment was programmed on ePrime (Psychology Software Tools, 2002) and lasted approximately 15 minutes. 30 faces varying parametrically on dominance and trustworthiness were generated using FaceGen 3.1 (<http://www.facegen.com>). Previous research has shown that these faces elicit dominance and trustworthiness judgments both at the explicit and the implicit level ^{39,40}. Following Oosterhof and Todorov's methodology ⁴⁵, the questions bearing on the three traits of interest, i.e. trustworthiness, dominance or likeability, were presented in separate blocks. The three-block sequence and the sequence of trials within each block were randomized between blocks and between participants. Participants had to answer: 'How [trait] is this person?' using a cursor on a 9-point scale ranging from 1 'not at all [trait]' to 9 'extremely [trait]' (recoded from -1 to +1 for the analyses). Depending on the block, [trait] was replaced by 'trustworthy', 'dominant' or 'likeable'. The face, the question and the scale appeared simultaneously after a 400 ms blank screen. Participants were instructed to follow their first impression and they were told that there was no right or wrong answer. The mouse was initially set to the middle of the screen in

order to reinforce the salience of the positive and the negative sides of the scale. The name of the dimension was displayed in each trial (Figure 1A).

Procedure

Participants were tested individually in a quiet room. Participants were seated at a 30 cm distance from the laptop. They completed three separate blocks of the face evaluation task, each block consisted of the same 30 faces. Participants could rest between each block. Following completion of the 90-trial experiment, participants filled out the STAI and the Kazdin Pleasure scale with the experimenter.

Data cleaning

2 TD participants and 2 ASD participants were excluded from the analysis for using only one side of the scales.

Group differences

We first checked that the included participants of the ASD and TD groups were matched on age, gender, IQ and anxiety. The ASD and TD groups did not differ on any of these variables (Table 1). We then measured the difference between the ASD and TD groups in the different types of motivation. As in previous studies ¹⁹, the ASD group was significantly less socially motivated than the TD group ($t(38) = -2.43, p = .019$) but did not differ in the two other types of motivation (Table 1). However, it is worth noting that contrary to Chevallier et al. (2012) ¹⁹, social motivation did not

significantly correlate with ADOS severity scores (coded as indicated in Gotham et al., ⁷⁷; $r = -.02 \pm 0.44$, $N = 20$, $t(16) = 0.09$, $p > .250$)

Finally, to test for possible differences between the ASD and the TD groups in the way participants performed the task, we ran several t-tests on scale use variables (i.e., ratings variance, number of different ratings, lower and higher ratings) for each scale. None of these values were significantly different between the two groups either for the dominance scale (all $ps > .250$), the trustworthiness scale (all $ps > .250$) or the likeability scale (all $ps > .118$).

Data analysis

Cues decoding: To measure the influence of social motivation on participants' ability to decode trustworthiness and dominance cues, we ran mixed linear regressions on trustworthiness and dominance ratings, taking avatars' levels of trustworthiness / dominance as well as participants' level of social motivation as regressors and participants' ID as a random factor.

Likeability evaluations: To investigate the impact of social motivation on the composition of likeability judgments, we ran a mixed linear regression on likeability evaluations, taking social motivation, ratings of dominance and ratings of trustworthiness as predictors and participants' ID as a random factor. Following Todorov et al. ⁶¹, this model included linear and quadratic interaction effects of perceived trustworthiness and perceived dominance as well as interaction terms between these two factors. In addition, we also ran a similar model taking avatars' objective levels of trustworthiness and

dominance instead of participants' ratings of trustworthiness and dominance.

References

1. Kaplan, H. S., Hooper, P. L. & Gurven, M. The evolutionary and ecological roots of human social organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 3289–3299 (2009).
2. Baumeister, R. F. & Leary, M. R. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* **117**, 497–529 (1995).
3. Epley, N. & Schroeder, J. Mistakenly seeking solitude. *J. Exp. Psychol. Gen.* **143**, 1980–1999 (2014).
4. Vouloumanos, A., Hauser, M. D., Werker, J. F. & Martin, A. The Tuning of Human Neonates' Preference for Speech. *Child Dev.* **81**, 517–527 (2010).
5. Salva, O. R., Farroni, T., Regolin, L., Vallortigara, G. & Johnson, M. H. The Evolution of Social Orienting: Evidence from Chicks (*Gallus gallus*) and Human Newborns. *PLOS ONE* **6**, e18802 (2011).
6. Rekers, Y., Haun, D. B. M. & Tomasello, M. Children, but Not Chimpanzees, Prefer to Collaborate. *Curr. Biol.* **21**, 1756–1758 (2011).
7. Langton, S. R. H., Law, A. S., Burton, A. M. & Schweinberger, S. R. Attention capture by faces. *Cognition* **107**, 330–342 (2008).
8. Lavie, N., Ro, T. & Russell, C. The Role of Perceptual Load in Processing Distractor Faces. *Psychol. Sci.* **14**, 510–515 (2003).

9. Ro, T., Friggel, A. & Lavie, N. Attentional biases for faces and body parts. *Vis. Cogn.* **15**, 322–348 (2007).
10. Chevallier, C. et al. Measuring Social Motivation Using Signal Detection and Reward Responsiveness. *PLOS ONE* **11**, e0167024 (2016).
11. Jones, R. M. et al. Behavioral and Neural Properties of Social Reinforcement Learning. *J. Neurosci.* **31**, 13039–13045 (2011).
12. Lin, A., Adolphs, R. & Rangel, A. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* **7**, 274–281 (2012).
13. Brown, L. H., Silvia, P. J., Myin-Germeys, I. & Kwapil, T. R. When the need to belong goes wrong: the expression of social anhedonia and social anxiety in daily life. *Psychol. Sci.* **18**, 778–82 (2007).
14. Germine, L. T., Garrido, L., Bruce, L. & Hooker, C. Social anhedonia is associated with neural abnormalities during face emotion processing. *NeuroImage* **58**, 935–945 (2011).
15. Gooding, D. C. & Tallent, K. A. Spatial, object, and affective working memory in social anhedonia: an exploratory study. *Schizophr. Res.* **63**, 247–260 (2003).
16. Kwapil, T. R. et al. The social world of the socially anhedonic: Exploring the daily ecology of asociality. *J. Res. Personal.* **43**, 103–106 (2009).
17. Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S. & Schultz, R. T. The social motivation theory of autism. *Trends Cogn. Sci.* **16**, 231–239 (2012).

18. Harvey, P.-O., Bodnar, M., Sergerie, K., Armony, J. & Lepage, M. Relation between emotional face memory and social anhedonia in schizophrenia. *J. Psychiatry Neurosci. JPN* **34**, 102–110 (2009).
19. Chevallier, C., Grèzes, J., Molesworth, C., Berthoz, S. & Happé, F. Brief Report: Selective Social Anhedonia in High Functioning Autism. *J. Autism Dev. Disord.* **42**, 1504–1509 (2012).
20. Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J. & Brown, E. Children with Autism Fail to Orient to Naturally Occurring Social Stimuli. *J. Autism Dev. Disord.* **28**, 479–485 (1998).
21. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. & Jones, W. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* **459**, 257–261 (2009).
22. Blanchard, J. L., Horan, W. P. & Brown, S. A. Diagnostic differences in social anhedonia: A longitudinal study of schizophrenia and major depressive disorder. *J. Abnorm. Psychol.* **110**, 363–371 (2001).
23. Watson, K. K., Werling, D. M., Zucker, N. & Platt, M. Altered Social Reward and Attention in Anorexia Nervosa. *Front. Psychol.* **1**, (2010).
24. Gliga, T. & Csibra, G. Seeing the face through the eyes: a developmental perspective on face expertise. in *Progress in Brain Research* (ed. Rosander, C. von H. and K.) **164**, 323–339 (Elsevier, 2007).
25. Grelotti, D. J., Gauthier, I. & Schultz, R. T. Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Dev. Psychobiol.* **40**, 213–225 (2002).

26. Klin, A., Jones, W., Schultz, R. & Volkmar, F. The enactive mind, or from actions to cognition: lessons from autism. *Philos. Trans. R. Soc. B Biol. Sci.* **358**, 345–360 (2003).
27. Mundy, P. Annotation: The neural basis of social impairments in autism: the role of the dorsal medial-frontal cortex and anterior cingulate system. *J. Child Psychol. Psychiatry* **44**, 793–809 (2003).
28. Cuthbert, B. N. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry* **13**, 28–35 (2014).
29. Insel, T. et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* **167**, 748–751 (2010).
30. National Institute of Mental Health. Strategic Plan for Research. (2015).
31. Sanislow, C. A. et al. Developing constructs for psychopathology research: research domain criteria. *J. Abnorm. Psychol.* **119**, 631–639 (2010).
32. National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria. Behavioral Assessment Methods for RDoC Constructs. *Bethesda MD Natl. Inst. Ment. Health* (2016).
33. Parish-Morris, J. et al. Visual attention to dynamic faces and objects is linked to face processing skills: a combined study of children with autism and controls. *Front. Psychol.* **4**, (2013).

34. Olsen, E. K., Bjorkquist, O. A., Bodapati, A. S., Shankman, S. A. & Herbener, E. S. Associations between trait anhedonia and emotional memory deficits in females with schizophrenia versus major depression. *Psychiatry Res.* **230**, 323–330 (2015).
35. Matson, J. L. & Nebel-Schwalm, M. S. Comorbid psychopathology with autism spectrum disorder in children: An overview. *Res. Dev. Disabil.* **28**, 341–352 (2007).
36. Simonoff, E. et al. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *J. Am. Acad. Child Adolesc. Psychiatry* **47**, 921–929 (2008).
37. Todorov, A., Olivola, C. Y., Dotsch, R. & Mende-Siedlecki, P. Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annu. Rev. Psychol.* **66**, 519–545 (2015).
38. Safra, L., Wyart, V., Baumard, N. & Chevallier, C. Face evaluations confirm that social motivation serves cooperative goals. *Nat. Hum. Behav.* (submitted).
39. Stewart, L. H. et al. Unconscious evaluation of faces on social dimensions. *J. Exp. Psychol. Gen.* **141**, 715–727 (2012).
40. Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N. & Falvello, V. B. Validation of data-driven computational models of social perception of faces. *Emotion* **13**, 724–738 (2013).

41. Cogsdill, E. J., Todorov, A. T., Spelke, E. S. & Banaji, M. R. Inferring Character From Faces A Developmental Study. *Psychol. Sci.* **25**, 1132–1139 (2014).
42. Ewing, L., Caulfield, F., Read, A. & Rhodes, G. Appearance-based trust behaviour is reduced in children with autism spectrum disorder. *Autism* **19**, 1002–1009 (2015).
43. Philip, R. C. M. et al. Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychol. Med.* **40**, 1919–1929 (2010).
44. White, S., Hill, E., Winston, J. & Frith, U. An islet of social ability in Asperger Syndrome: Judging social attributes from faces. *Brain Cogn.* **61**, 69–77 (2006).
45. Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proc. Natl. Acad. Sci.* **105**, 11087–11092 (2008).
46. Chamberlain, B., Kasari, C. & Rotheram-Fuller, E. Involvement or Isolation? The Social Networks of Children with Autism in Regular Classrooms. *J. Autism Dev. Disord.* **37**, 230–242 (2006).
47. Kasari, C., Locke, J., Gulsrud, A. & Rotheram-Fuller, E. Social Networks and Friendships at School: Comparing Children With and Without ASD. *J. Autism Dev. Disord.* **41**, 533–544 (2010).
48. Locke, J., Ishijima, E. H., Kasari, C. & London, N. Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting. *J. Res. Spec. Educ. Needs* **10**, 74–81 (2010).

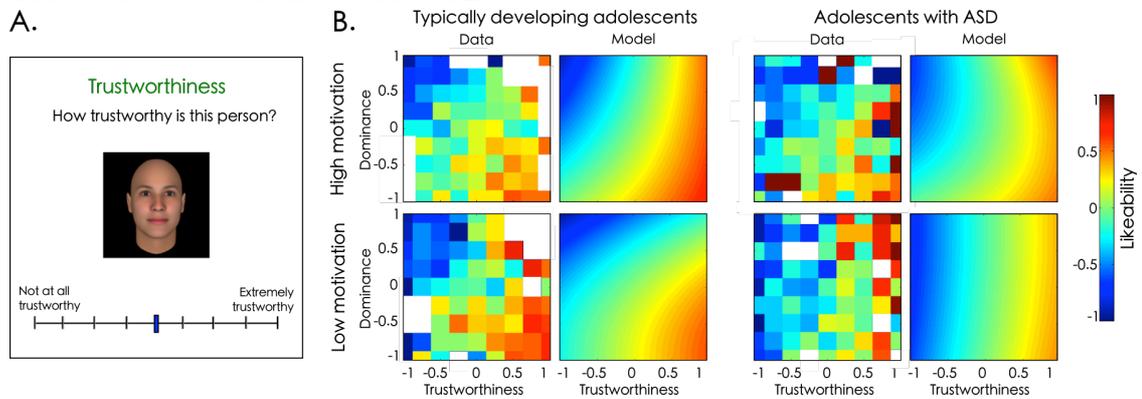
49. Shattuck, P. T., Orsmond, G. I., Wagner, M. & Cooper, B. P. Participation in Social Activities among Adolescents with an Autism Spectrum Disorder. *PLOS ONE* **6**, e27176 (2011).
50. Lombardo, M. V., Barnes, J. L., Wheelwright, S. J. & Baron-Cohen, S. Self-Referential Cognition and Empathy in Autism. *PLOS ONE* **2**, e883 (2007).
51. Lombardo, M. V. & Baron-Cohen, S. Unraveling the paradox of the autistic self. *Wiley Interdiscip. Rev. Cogn. Sci.* **1**, 393–403 (2010).
52. Lombardo, M. V. & Baron-Cohen, S. The role of the self in mindblindness in autism. *Conscious. Cogn.* **20**, 130–140 (2011).
53. Chevallier, C., Molesworth, C. & Happé, F. Diminished Social Motivation Negatively Impacts Reputation Management: Autism Spectrum Disorders as a Case in Point. *PLOS ONE* **7**, e31107 (2012).
54. Dubey, I., Ropar, D. & Hamilton, A. F. de C. Measuring the value of social engagement in adults with and without autism. *Mol. Autism* **6**, 35–35 (2015).
55. Happé, F., Ronald, A. & Plomin, R. Time to give up on a single explanation for autism. *Nat. Neurosci.* **9**, 1218–1220 (2006).
56. Jeste, S. S. & Geschwind, D. H. Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat. Rev. Neurol.* **10**, 74–81 (2014).
57. Caulfield, F., Ewing, L., Bank, S. & Rhodes, G. Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children. *Br. J. Psychol.* **107**, 503–518 (2016).

58. Caulfield, F., Ewing, L., Burton, N., Avard, E. & Rhodes, G. Facial Trustworthiness Judgments in Children with ASD Are Modulated by Happy and Angry Emotional Cues. *PLOS ONE* **9**, e97644 (2014).
59. Ioannou, C. et al. Shared mechanism for emotion processing in adolescents with and without autism. *Sci. Rep.* **7**, srep42696 (2017).
60. Todorov, A. & Engell, A. D. The role of the amygdala in implicit evaluation of emotionally neutral faces. *Soc. Cogn. Affect. Neurosci.* **3**, 303–312 (2008).
61. Todorov, A., Said, C. P., Oosterhof, N. N. & Engell, A. D. Task-invariant Brain Responses to the Social Value of Faces. *J. Cogn. Neurosci.* **23**, 2766–2781 (2011).
62. Vecchiato, G. et al. Neuroelectrical Correlates of Trustworthiness and Dominance Judgments Related to the Observation of Political Candidates. *Comput. Math. Methods Med.* **2014**, e434296 (2014).
63. Watkins, C. D., Jones, B. C. & DeBruine, L. M. Individual differences in dominance perception: Dominant men are less sensitive to facial cues of male dominance. *Personal. Individ. Differ.* **49**, 967–971 (2010).
64. Watkins, C. D., Quist, M. C., Smith, F. G., DeBruine, L. M. & Jones, B. C. Individual differences in women's perceptions of other women's dominance. *Eur. J. Personal.* **26**, 79–86 (2012).
65. Hall, J. et al. Social Cognition, the Male Brain and the Autism Spectrum. *PLOS ONE* **7**, e49033 (2012).

66. Hägele, C. et al. Dimensional psychiatry: reward dysfunction and depressive mood across psychiatric disorders. *Psychopharmacology (Berl.)* **232**, 331–341 (2015).
67. Neufeld, J., Ioannou, C., Korb, S., Schilbach, L. & Chakrabarti, B. Spontaneous Facial Mimicry is Modulated by Joint Attention and Autistic Traits. *Autism Res. Off. J. Int. Soc. Autism Res.* **9**, 781–789 (2016).
68. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR)*. (American Psychiatric Publishing, 2000).
69. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **24**, 659–685 (1994).
70. Lord, C. et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* **30**, 205–223 (2000).
71. Bach, M. The Freiburg Visual Acuity test--automatic measurement of visual acuity. *Optom. Vis. Sci. Off. Publ. Am. Acad. Optom.* **73**, 49–53 (1996).
72. Wechsler, D. Wechsler intelligence scale for children--Fourth Edition (WISC-IV). *San Antonio TX Psychol. Corp.* (2003).
73. Grégoire, J. *L'examen clinique de l'intelligence de l'enfant: fondements et pratique du WISC-IV*. (Editions Mardaga, 2009).

74. KAZDIN, A. E. Evaluation of the Pleasure Scale in the Assessment of Anhedonia in Children. *J. Am. Acad. Child Adolesc. Psychiatry* **28**, 364–372 (1989).
75. Fioravanti-Bastos, A. C. M., Cheniaux, E. & Landeira-Fernandez, J. Development and validation of a short-form version of the Brazilian state-trait anxiety inventory. *Psicol. Reflex. E Crítica* **24**, 485–494 (2011).
76. Spielberger, C. D. Manual for the State-Trait Anxiety Inventory STAI (Form Y) ('Self-Evaluation Questionnaire'). (1983).
77. Gotham, K., Pickles, A. & Lord, C. Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. *J. Autism Dev. Disord.* **39**, 693–705 (2009).

Figure 1 – Social motivation has distinct effects on face evaluations in adolescents with and without ASD



(A) Example of an evaluation trial. Participants had to rate each face by moving a cursor. **(B) Likeability ratings as a function of trustworthiness (x axis) and dominance ratings (y axis) in typically developing adolescents (left) and adolescents with ASD (right).** Rating intensity is represented on a scale ranging from blue for lower ratings to red for higher ratings. Pixelized figures correspond to averaged data in the initial study (data) for the most (upper row) and least (lower row) socially motivated participants (median split). Smoothed figures represent the predictions of the regression models ran separately on the two participant samples. While in typically developing adolescents, higher levels of social motivation are associated with an increase in the weight granted to trustworthiness, this is not the case in adolescents with ASD.

Table 1 - Descriptive statistics for age, gender and IQ and anxiety of the ASD and TD groups

	TD (N = 20)	ASD (N = 20)	Statistics
Age	13.70 ± 0.61	14.00 ± 0.88	$t(38) = 0.58, p > .250$
Gender ratio	25 % female	15 % female	$\chi^2(2, N = 42) = 0.01, p > .250$
IQ	106.60 ± 6.14	100.85 ± 10.13	$t(38) = -1.02, p > .250$
STAI	13.80 ± 1.80	14.35 ± 1.96	$t(38) = 0.43, p > .250$
Social Motivation	30.10 ± 1.74	34.10 ± 2.97	$t(38) = -2.43, p = .019$
Physical Motivation	12.80 ± 0.91	13.10 ± 1.23	$t(38) = -0.41, p = .684$
Other Sources of Motivation	18.65 ± 2.06	21.15 ± 1.73	$t(38) = -1.94, p = .059$

Social motivation as a high-stake cooperative strategy

Across a total of six studies, we have shown that social motivation is, as soon as adolescence, associated with an increase in the weight granted to cooperation-related signals for social decisions. Our results thus confirm a central prediction of the evolutionary model of social motivation that proposes to construe social motivation as an adaptation to highly cooperative environments (Chevallier, Kohls, et al., 2012).

In addition to confirming this prediction, the cooperation-power framework (Oosterhof & Todorov, 2008) also allowed us to formulate a new hypothesis about the social strategy associated with social motivation, which would correspond to a high-stake cooperative strategy. Indeed, the increased importance of trustworthiness in highly socially motivated participants was associated with a higher probability to approach dominant and trustworthy partners. Importantly, while trustworthiness provides information about the probability of reciprocation (Chang et al., 2010), dominance-power also impacts cooperation outcomes by modifying the stakes of the cooperative activity. Indeed, cooperative interactions with powerful individuals can provide higher gains, in terms of social benefits or amount of resources (Kaufmann, 1983; N. Lin, 1999; Stockley & Bro-Jørgensen, 2011; von Rueden, 2014; von Rueden, Gurven, & Kaplan, 2010) but can also cause larger losses, with, for instance, lower possibilities of retaliation (Carré,

McCormick, & Mondloch, 2009b; Carmen E. Lefevre, Etchells, Howell, Clark, & Penton-Voak, 2014). In summary, cooperation with more powerful partners seems to bear higher stakes than cooperation with less powerful partners. Our results thus suggest that highly socially motivated participants are more likely to engage in cooperative activities that bear larger stakes. This cooperative behavior may maximize the number of cooperators while increasing the potential gains of cooperative activities, which may constitute an optimal strategy in highly cooperative environments. Our findings thus open the path to further investigations of the functional consequences of social motivation as well as to a new type of study on cooperation. Indeed, the suggested ability of the cooperation-power framework to identify variations in cooperative strategies invites to go one step further and to apply this framework to the study of individual differences in cooperation more generally.

Towards new investigations of cooperative strategies

By showing individual differences in the combination of trustworthiness and dominance signals for approach behavior, our results suggest that individuals vary in their willingness to cooperate with more or less trustworthy and more or less dominant partners. Our findings thus call for a new type of investigation of cooperation that goes beyond general cooperative tendencies. Indeed, classical investigations of individual

variability in cooperation mostly rely on self-reported questionnaires (Petersen & Aarøe, 2015), anonymous real-life situations (Holland, Silva, & Mace, 2012) or economic games played with unknown partners (Nettle, Colléony, & Cockerill, 2011), all taking cooperation as a trait. However, according to the investment definition of cooperation (Baumard, André, & Sperber, 2013), individuals' cooperative behavior should vary depending on their partners.

In particular, based on the economic literature on investment, individuals should be sensitive to the probability of success and to the size of the potential losses of cooperative activities, as they are for any investment (Ehrlich & Becker, 1972). For each individual a self-protection and a self-insurance profile can thus be defined that refers to her attitudes towards these two risks respectively. Applying the same reasoning to cooperation, the systematic measure of cooperation self-protection and self-insurance can help to identify individuals' risk attitudes in cooperative activities, and thus their cooperative strategy, as well as the mechanisms underlying variations in cooperation.

This framework may notably provide new insights into the link between individuals' current state and cooperative strategies. As is the case in this chapter, individuals' state can be defined by internal variables such as personality traits, gender or age (Güroğlu, van den Bos, & Crone, 2014; Sheskin et al., 2014). In addition, external factors, such as individuals' current or past level of resources or the environment mortality rate, may

also influence individuals' cooperative behavior. Indeed, these external variables may change the cost and benefits associated with cooperation and thus shape the development of different cooperation strategies. In line with this idea, poverty has been shown to be associated with decreased cooperation (Holland et al., 2012; Nettle, Colléony, et al., 2011; Petersen & Aarøe, 2015). The application of the cooperation self-insurance and self-protection framework could help to further characterize this association by revealing if this link is due to specific cooperation-related risk attitudes of individuals exposed to resource scarcity, and thereby could inform about the mechanisms by which economic poverty influences cooperation.

More generally, as different environments may be associated with different success-maximizing strategies, such as more cooperative long-term strategies or more aggressive short-term ones, environmental factors may modulate the importance individuals grant to cooperation- and to power-related cues not only for cooperative activities but also in many other social domains. The way different social choices respond to these environmental variations can provide new insights into the type of social cues that guide specific social decisions, and thereby into the underlying cognitive mechanisms. In a third chapter, we will explore variations in the importance of cooperation- and power-related signals for social decisions in response to environmental factors to better understand the cognitive processes involved in political choices.

Chapter three. Investigation of the cognitive processes underlying social choices: study of leader choices

An ecological approach to human behavior

The idea that animals react differently in different situations is far from new; it is even deeply anchored in folk psychology and has been at the basis of countless studies. For instance, transient manipulations of the environment during psychology experiments produce changes in a large range of behavior, from economic decision-making to social behavior (e.g., Griskevicius, Goldstein, Mortensen, Cialdini, & Kenrick, 2006; Krosch & Amodio, 2014; Little et al., 2010; Liu, Feng, Suo, Lee, & Li, 2012; Miles,

Lumsden, Richardson, & Macrae, 2011). As an example amongst thousands of experiments, Little et al. (2010) have shown that exposure to pathogens cues increased the preference for symmetric and more gender-stereotypical faces.

Humans' behavioral flexibility does not only operate in the lab, real-life events also influence individuals' behavior. For instance, interviews conducted after terrorist attacks revealed that these acute social threats are associated with increased conservatism and authoritarianism (Bonanno & Jost, 2006; Van de Vyver, Houston, Abrams, & Vasiljevic, 2016). Similarly, Li et al. (2012) have shown that inhabitants of an area that experienced important damages following the Wenchuan earthquake (China, 2008) displayed an increased preference for immediate rewards compared to inhabitants from a non-devastated area of the same region. Individuals do not only react to acute events, variations in behavior are also associated with the long-term experience of different environments (Holland et al., 2012; Murray, Schaller, & Suedfeld, 2013; Nettle, Colléony, et al., 2011). Cooperation has notably been shown to vary in this way, with cooperative tendencies being dramatically reduced between wealthy and deprived neighborhoods (Holland et al., 2012; Nettle, Colléony, et al., 2011). This sensitivity to the environment appears as early as childhood. For instance, the association between economic deprivation and lower cooperative tendencies has been evidenced in children as young as 7-year old (Benenson, Pascoe, & Radmore, 2007; Safra et al., 2016- see Appendix).

Adopting an ecological perspective, these behavioral variations can be understood as cases of phenotypic plasticity. More precisely, they can be construed as the output of evolutionary selected responses to specific environmental signals. In other words, under this view, evolutionary pressures have acted on the way cognitive mechanisms respond to different environmental cues. In this perspective, the investigation of variations in behavior with environmental factors can inform us about the underlying cognitive mechanisms and shed light on their evolutionary roots.

This ecological approach on human behavior provides a new framework to investigate the link between environment and behavior, and notably incites to study this link in an ontological perspective. Indeed, individuals' early environment provides cues about the kind of environment they will likely face as adults and the kind of somatic resources they can rely on for their development. Thereby, they are crucial to calibrate current and future behavior (Nettle et al., 2013). Accordingly, studies on non-human animals have shown that signals perceived during the juvenile period influence behavior throughout the individual's lifespan (Roff, 2002; Stearns, 2000). Even when they live in low-stress environments as adults, non-human animals, who have experienced high levels of stress during the juvenile period go on to be more present-orientated, prioritizing immediate survival and reproduction over long-term benefits, compared to individuals who experienced low-stress environments from their younger age (Bloxham, Bateson, Bedford, Brilot, & Nettle, 2014).

In the past twenty years, evolutionary psychologists have applied this framework to study the influence of early environmental signals on human behavior. In very much the same way as animal models, early adversity has been shown to be correlated with differences in time discounting, somatic investment and reproduction timing in humans. Even after controlling for current socio-economic status, people born with low birth-weight or who experienced poverty, psychosocial stress or family disruption during childhood mature earlier and have their first child sooner than control populations (Bateson et al., 2004; Ellis, McFadyen-Ketchum, Dodge, Pettit, & Bates, 1999; Mell, Safra, Algan, Baumard, & Chevallier, 2017; Nettle, Coall, & Dickins, 2011, 2010; Webster, Graber, Gesselman, Crosier, & Schember, 2014- see Appendix). Similarly, adults who suffered from harsh conditions during childhood go on to exhibit greater behavioral disinhibition and steeper rates of time discounting independently of their economic status later in life (Griskevicius et al., 2013; Mittal, Griskevicius, Simpson, Sung, & Young, 2015b).

By revealing the influence of early environmental signals on human behavior, these studies shed light on the development of these behaviors. In addition, they provide evidence that individual differences in behavior can be construed as the outputs of biologically rooted responses to the environment. Studying the influence of early external signals on behavior can thus inform on the way cognitive mechanisms respond to environmental factors and, thereby on their functioning and evolutionary bases.

Importantly, this approach is not restricted to the behavioral domains that have been studied in animal models, but can be extended to any behavior for which early environmental signals provide relevant information about the individuals' future (Nettle et al., 2013).

Childhood environment and social preferences

Harshness in early stages of development is notably likely to induce important changes in social preferences. Indeed, some social strategies may be better suited to highly cooperative environments while others may be more adapted to environments in which social competition is fiercer and social trust lower. In line with this idea, independently from their socioeconomic status later in life, adults who grew up in stressful environments are more sensitive to social threats and negative emotions than adults who grew up in less stressful environments, which may be favorable in environments presenting higher levels of violence (Javanbakht et al., 2015; Kim et al., 2013). Going one step further, the study of the influence of early environmental signals should not only provide insights into the precise cognitive mechanisms involved in social behavior, such as emotion processing, but also into more systemic social attitudes.

This approach appears particularly promising for better understanding the social attitudes that have already been shown to be sensitive to individuals' current environment. This is the case for political attitudes, which have been

theorized to be notably guided by an evolved leader selection mechanism that responds to environmental variations (King, Johnson, & Van Vugt, 2009). In line with this idea, analysis of historical records and sociological data have revealed systematic variations in the preference for strong leaders with the exposure to environmental threats (i.e., economic and social threats; Doty, Peterson, & Winter, 1991; McCann & Stewin, 1990; Onraet, van Hiel, & Cornelis, 2013; Sales, 1973). For instance, Onraet et al. (2013) have shown that economic threats (e.g., inflation and unemployment) are strong predictors of between-countries variations in authoritarianism. At the individual level also, perceived threat to safety and dangerous worldviews correlate reliably with the preference for strong leaders (Mirisola, Roccato, Russo, Spagna, & Vieno, 2014; Onraet, Dhont, & Hiel, 2014; Perry, Sibley, & Duckitt, 2013; Sibley, Wilson, & Duckitt, 2007). This link has also been confirmed experimentally: social threat scenarios systematically induce a preference for more masculine leaders (Little, Burriss, et al., 2007a; Re, DeBruine, Jones, & Perrett, 2013; Spisak, Dekker, et al., 2012).

To further investigate the idea that the preference for strong leaders in times of threat corresponds to a deeply rooted response to external stressors, we designed a study in which we investigated the effect of childhood poverty on authoritarianism. As environmental threats are associated with an increased level of social competition and a decreased level of social trust (Ellis, Figueredo, Brumbach, & Schlomer, 2009; Kaplan &

Gangestad, 2005; Nettle, Colléony, et al., 2011), the influence of external threats on political attitudes may be reflected in the importance of cooperation- and power-related signals for leader choice. Accordingly, studies on face evaluations have consistently shown that more dominant and less trustworthy leaders are preferred in times of social threat (Laustsen & Petersen, 2015; Little, Burriss, et al., 2007a; Little et al., 2012). Based on these results, in our study, we used face evaluations of avatars varying in dominance and trustworthiness to investigate the influence of early environmental harshness on leader preferences

We first assessed whether early signals of environmental harshness had an immediate impact on leader preferences. To do so, we asked 41 children coming from two contrasted neighborhoods in terms of socio-economic status to choose the individual they would prefer as their leader in a mountain trip. Replicating the effect of economic threats shown in adults (Little, Burriss, et al., 2007a; Spisak, Homan, et al., 2012), we found that children from lower socio-economic backgrounds prefer stronger leaders, i.e., more dominant and less trustworthy leaders, than their wealthier peers. In other words, this experiment revealed that environmental harshness is linked with a preference for stronger leaders as early as childhood, providing support to the hypothesis that this association reflects a deeply-rooted response to the environment.

We then extended these results by investigating whether the association between childhood poverty and the preference for strong leaders lasted in

adulthood (representative sample of the French population: $N = 1000$). In line with the hypothesis of a long-term impact of environmental signals on social attitudes, adults reporting having suffered from poverty during their childhood are more likely to select more dominant and less trustworthy avatars as leaders, independently from their socio-economic status late in life. Importantly, this long-term effect of childhood poverty was replicated using participants' self-reports of authoritarianism. Indeed, participants who suffered from poverty during their childhood are also more likely to strongly agree with the idea that 'having a strong leader who does not have to bother with the elections and the parliament is a good thing', both in the representative sample of the French population and in an independent population constituted of representative samples of 46 European countries (European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008), 2015). As a conclusion, our experiments reveal that the early experience of poverty is associated with a preference for stronger leaders in both children and adults, thereby providing new insights into the ontology of political attitudes. In particular, by showing a link between early environmental signals on political behavior, and although our experimental design did not allow us to establish a causal link between childhood poverty and authoritarianism or to exclude the existence of a mediating variable, our study suggests that the cognitive mechanisms involved in political behavior respond to environmental signals perceived during development. Together, our three studies provide new evidence in favor of the hypothesis that

differences in political behavior can, at least partially, be understood as the product of a deeply rooted response to environmental signals.

The manuscript in which these results are reported has been published in Evolution and Human Behavior in May 2017. My contribution to this work was as follows: design of the experiment, analysis of the data and writing of the paper.



Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehonline.org

Childhood harshness predicts long-lasting leader preferences

Safrá Lou^{a,*}, Algan Yann^b, Tecu Teodora^c, Grèzes Julie^a, Baumard Nicolas^d, Chevallier Coralie^{a,*}^a Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, Inserm, U960, Laboratoire de Neurosciences Cognitives (LNC), F-75005 Paris, France^b SciencesPo, Department of Economics, Paris, France^c University of Bucharest, Faculty of Philosophy, Bucharest, Romania^d Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, CNRS, UMR8129, Institut Jean-Nicod, F-75005 Paris, France

ARTICLE INFO

Article history:

Initial receipt 5 January 2017

2 May 2017

Final revision received 2 May 2017

Available online xxxx

Keywords:

Authoritarianism

Poverty

Childhood

Social perception

Dominance

Politics

ABSTRACT

Understanding the origins of political authoritarianism is of key importance for modern democracies. Recent works in evolutionary psychology suggest that human cognitive preferences may be the output of a biological response to early stressful environments. In this paper, we hypothesized that people's leader preferences are partly driven by early signals of harshness. We experimentally elicited children's (Study 1) and adults' (Study 2) political preferences using faces controlled for dominance and trustworthiness and showed that early childhood harshness has an enduring effect on adult political attitudes. Importantly, this effect was further confirmed using self-reported extreme authoritarianism (Study 2) and by the analysis of the large database of the European Value Survey (Study 3). We discuss the potential political implications of this early calibration of leader preferences.

© 2017 Elsevier Inc. All rights reserved.

For decades, empirical works have demonstrated that political preferences vary systematically with environmental harshness. Perceived threat to safety and dangerous worldviews indeed correlates with right-wing authoritarianism (Onraet, Dhont, & Hiel, 2014; Sibley, Wilson, & Duckitt, 2007), and threatening or war scenarios are associated with a preference for taller, more masculine, more dominant and less trustworthy leaders (Laustsen & Petersen, 2015; Little, Roberts, Jones, & DeBruine, 2012; Re, DeBruine, Jones, & Perrett, 2013). Importantly, this authoritarianism shift also appears in response to non-political threats. For instance, pathogen- and disease-avoidance, a major issue in human evolution, correlate with the degree of authoritarianism at the country level (Murray, Schaller, & Suedfeld, 2013; Thornhill, Fincher, & Aran, 2009) and at the individual level (see Terrizzi, Shook, & McDaniel, 2013 for a meta-analysis). Taken together, these studies suggest that the preference for strong leaders is a deeply rooted evolutionary response to external stressors.

However, cues from individuals' current environment are not the only signals affecting behavior. Signals perceived during childhood are indeed crucial to calibrate current and future behaviors (Bateson et al., 2004; Del Giudice, Ellis, & Shirtcliff, 2011; Nettle, Frankenhuis, & Rickard, 2013). Environmental childhood provides cues about the kind of environment individuals will likely face as adults or the kind of somatic resources they can rely on for their development (Nettle et al.,

2013). In line with this idea, empirical research shows that phenotypes are adjusted to early conditions through multiple developmental mechanisms (Frankenhuis, Panchanathan, & Nettle, 2016). For instance, non-human animals who experience a period of high stress in the juvenile period go on to be more present-oriented, and to prioritize immediate survival and reproduction over long-term benefits (Bateson et al., 2004). Similarly, people born with low birth-weight or who experience psychosocial stress and family disruption in childhood mature earlier and have their first child sooner than control populations (Adair, 2001; Nettle, Coall, & Dickins, 2011; Sloboda, Hart, Doherty, Pennell, & Hickey, 2007). Harshness in early stages of development also induces important changes in social cognition. For example, independently of their socioeconomic status later in life, adults who grew up in stressful environments are more sensitive to social threats and negative emotions, which may be adaptive in more competitive and violent environments (Javanbakht et al., 2015; Kim et al., 2013). The goal of the present paper is to study whether leader preferences are also influenced by cues of childhood harshness, independently of individuals' current circumstances. To test this hypothesis, we relied on a robust measure of harshness in childhood: resource deprivation. Childhood resource deprivation indeed reflects both lower levels of resources and increased instability and exposure to adverse events (Evans, 2004). In other words, resource deprivation constitutes an interesting proxy for the level of external stress experienced during childhood. Therefore, we assess the association between deprivation during childhood and children's leader preferences (Study 1). We then evaluate the persistence of this effect in adulthood (Studies 2 & 3).

* Corresponding authors.

E-mail addresses: lou.safrá@gmail.com (L. Safrá), coralie.chevallier@gmail.com (C. Chevallier).

To measure leader preferences consistently in children and adults, we relied on participants' perception of faces. Extensive research in psychology has indeed shown that facial cues are used for leader choice both in adults and children, and that their use reliably predicts election outcomes (Antonakis & Dalgas, 2009; Olivola & Todorov, 2010; Todorov, Mandisodza, Goren, & Hall, 2005). In addition, cross-national studies and experimental manipulations have shown that the importance granted to specific facial cues such as trustworthiness or dominance is sensitive to environmental factors (Laustsen & Petersen, 2015; Little et al., 2012). Similarly, real-life political differences between Democrats and Republicans have been linked with stable differences in facial preferences when choosing a leader (Laustsen & Petersen, 2015; Olivola, Sussman, Tsetsos, Kang, & Todorov, 2012). These results suggest that people's reliance on specific facial cues to choose a leader is a reliable proxy for their actual political preferences.

In the present project, we exploited the differential impact of dominance and trustworthiness in social judgments (Oosterhof & Todorov, 2008) to investigate the relationship between childhood deprivation and the preference for authoritarian leaders. More precisely, we used faces controlled for both dominance and trustworthiness to measure how early adverse experience may shape the use of these two facial cues for choosing a leader. Finally, we confirmed the association between leader preferences and childhood deprivation by analyzing the effect of self-reported extreme authoritarianism in a nationally representative sample of French adults (Study 2) and in a large-scale survey on 46 European countries (Study 3, *European Values Study Longitudinal Data File 1981–2008* (EVS 1981–2008), 2015).

1. Study 1

The aim of this first study was to investigate the immediate effect of childhood poverty on children's preference for strong leaders. To do so, we adapted an existing experimental design that successfully elicits political preferences in children (Antonakis & Dalgas, 2009).

1.1. Materials and methods

1.1.1. Participants

41 children from the city of Slatina, Romania, aged 6 to 8 years ($M = 6.85 \pm 0.13$ years; results are given in the standard form: mean \pm 95% confidence intervals) were recruited in two schools situated in a deprived neighborhood and in a working-class neighborhood, about 1 km away from the city center. A minimum target of 20 participants per group was pre-planned based on the number of 6 to 8 year-olds in the lower-SES school; the exact number was determined by scheduling constraints and by the number of parental consent forms we received. The study was approved by the schools' management team and by the School Inspectorate. Parents signed a written informed consent form and children provided verbal assent at the start of the procedure. All study procedures complied with the Declaration of Helsinki. Children received a small gift to thank them for their participation.

Two *Childhood Deprivation Groups* (*Deprived* and *Not Deprived*) were formed based on children's neighborhood. Because neighborhood status may interact with children's status, we excluded children whose parental income did not match their neighborhood's status. With this goal in mind, we asked parents whether their income was above or below the legal minimum wage in Romania (i.e. 850 lei per month, which corresponds approximately to 216\$). One participant was excluded in the *Deprived Group* and one was excluded in the *Not Deprived Group* resulting in 20 participants in the *Deprived Group* and 19 in the *Not Deprived Group*. The two groups were further characterized by lack of material possessions and lower access to cultural activities, they had younger and less educated parents, more siblings and more crowded houses and more unstable families (all $ps < 0.035$; Table S1). Children in the *Deprived and Not Deprived Childhood Groups* were matched on

age ($t(37) = 0.19, p > 0.250$) and gender ratio ($\chi^2(1, N = 39) = 0.62, p > 0.250$; Table S1).

1.1.2. Procedure and analysis

Following Antonakis and Dalgas (2009), we asked children to choose the individual they would prefer as the captain of their team to go on a mountain trip. They had to repeat this choice five times, on five different pairs of faces. Each pair included two versions of a single avatar identity: a more dominant one and a more trustworthy one (Fig. 1A). The identities were selected by a native Romanian to match the most common types of faces in Romania. Morphs were created with the XxMorphs freeware and using the Facegen 3.1 open database (Oosterhof & Todorov, 2008). The more dominant faces corresponded to a 50% morph between a 3-point dominant and a 1-point trustworthy face. Theoretically, the obtained average faces were 1.5-point dominant and 0.5-point Trustworthy. Symmetrically, the more trustworthy faces corresponded to a morph between a 1-point dominant and a 3-point trustworthy face, resulting in 0.5-point dominant and 1.5-point trustworthy faces.

Validation of the stimuli by 60 Amazon MTurkers (MTurk, <http://www.mturk.com>) confirmed that the more dominant faces were judged as more dominant and less trustworthy than the more trustworthy faces (mean of participants' correct identification rate: most dominant face: $86\% \pm 6\%$; most trustworthy face: $79\% \pm 6\%$; mean percentage of correct identification for each face: most dominant face: $86\% \pm 4\%$; most trustworthy face: $79\% \pm 13\%$). The pairs were presented in a random order and the position of the more dominant face (on the right or on the left of the more trustworthy face) was counterbalanced between pairs. The task lasted approximately 5 min.

The effect of *Childhood Deprivation* on the probability of choosing the more dominant face was investigated using a logistic regression taking *Childhood Deprivation* as a predictor. All the results obtained with this model were confirmed using *Parental income status* instead of *Childhood Deprivation* as the predictor (See Supplementary information).

1.2. Results

In line with our hypotheses, children experiencing deprivation were more likely to choose the more dominant and less trustworthy face as the captain for their team (logistic regression: $b = 0.68 \pm 0.57, z = 2.33, p = 0.020$). This difference was explained by a preference for strong leaders in the, most deprived group only ($M = 0.62 \pm 11, t(19) = 2.30, p = 0.033$; Least deprived group: $M = 0.45 \pm 10, t(18) = -0.99, p = 0.337$).

1.3. Conclusions

This first study demonstrates that children's leader preferences are sensitive to their environment and that deprivation biases these preferences towards strong leaders. Thus, in a second study, we investigate the long-lasting effect of early exposure to stressors on a nationally representative sample of adults with diverse socio-economic backgrounds.

2. Study 2

Adult participants had to choose whom they would vote for in a national election between avatar faces parametrically varying on trustworthiness and dominance. Participants also reported their preference for authoritarian leaders to investigate the relationship between childhood deprivation and explicit authoritarian attitudes.

2.1. Materials and methods

2.1.1. Participants

The number of subjects was fixed *a priori* with IPSOS polling institute to constitute a sample that was representative of the French population

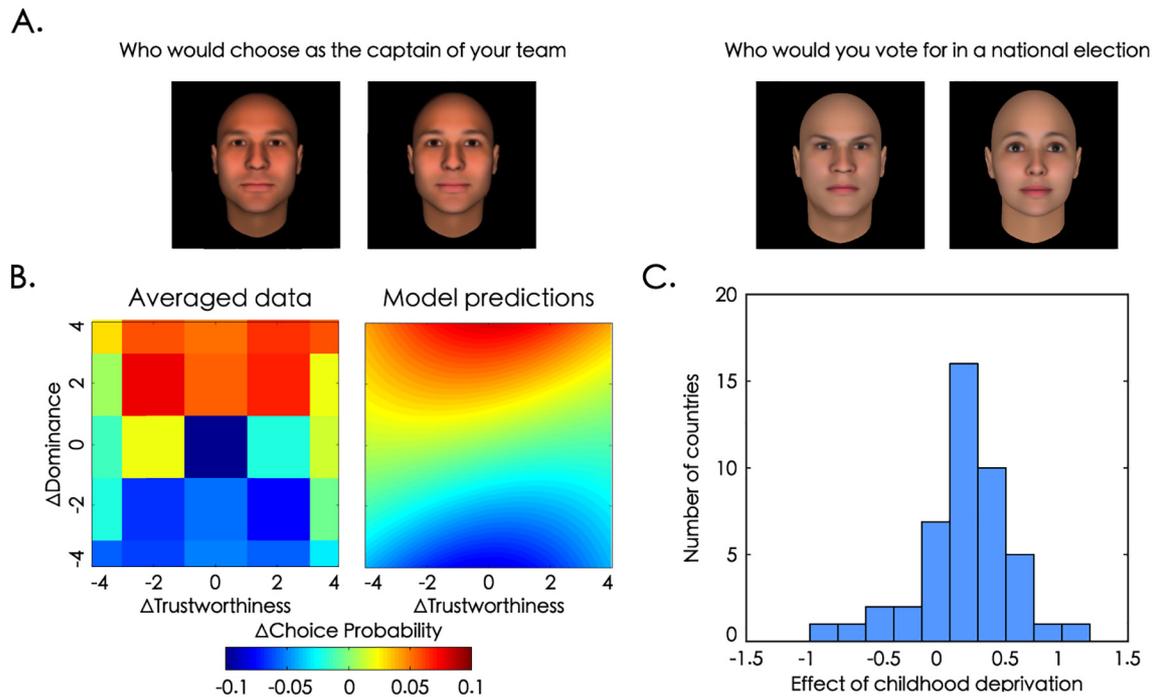


Fig. 1. Influence of childhood deprivation on leader preferences. **A.** Facial preferences experiment design. Children had to choose, for 5 pairs of faces, the individual they would prefer as the captain of their team. The faces of each pair corresponded to the same identity varying in the level of dominance (1.5 or 2.5 points) and trustworthiness (2.5 or 1.5 points; left). Adults had to choose the individual they would vote for in a national election for 36 pairs of faces varying parametrically on dominance and trustworthiness (right). **B.** Differences between choice probabilities for the different face combinations in *Deprived* and *Not Deprived Childhood* participants in Study 2. Difference in choice probability is represented from red for a higher probability for the *Not Deprived* ($N = 649$) than for the *Deprived Childhood Group* ($N = 163$) to blue for a lower probability of choice for the *Deprived Childhood Group*. The pixelized figure corresponds to the averaged data for each combination of Δ Dominance and Δ Trustworthiness and the smoothed figure to the models' predictions. Participants who experienced poverty during their childhood had a higher probability of choosing more dominant and less trustworthy leaders than those who did not experience poverty. **C.** Impact of Childhood Deprivation on extreme authoritarianism for 46 European countries. Two-level analyses revealed that having experienced deprivation during childhood increased the probability of reporting extreme authoritarianism. Importantly, this effect was also significant after controlling for parents' cultural capital as well as current status and confidence in politics.

for age, gender, geographical region, urban vs. rural and occupation (quota method). 1006 French participants (544 women) aged 16 to 83 years old (47.42 ± 1.01 years) participated in this online study. Each participant received a description of the study and provided their informed consent before starting the experiment. Following IPSOS' policy, participants were not paid to take part in the study. The experiment lasted around 15 min.

2.1.2. Procedure

The procedure for this experiment was a standard discrete choice task (McFadden, 1980, see Fig. 1A). This methodology was validated in two pre-test experiments, see Supplementary information Study S1 & S2 for details.

The experiment was programmed in Javascript. Eight avatar faces, generated via FaceGen 3.1 software and controlled for their level of trustworthiness and dominance were used for this experiment. These faces have already been successfully demonstrated to elicit dominance and trustworthiness judgments both at the explicit and implicit level (Stewart et al., 2012). These faces spanned every possible combination of dominance and trustworthiness in a range of -2 to $+2$ points with an increment of 2 points (Oosterhof & Todorov, 2008) and were regularly spaced on both dominance and trustworthiness. In each presented pair, the faces were 2 to 4 points different from each other on at least one dimension (dominance or trustworthiness). This resulted in 36 pairs of faces.

Each trial began with a central fixation cross presented for 300 ms. The two faces were then presented simultaneously. Participants had to select who they would vote for in a national election, and had up to 2 s to answer by pressing "e" for the face on the left and "p" for the face on the right. If they failed to answer within 2 s, the next trial was automatically presented. A blank page was presented for around

100 ms after each trial. The 36 possible pairs of faces were presented in a random order. The presentation position of the faces (the more trustworthy face on the left or on the right) of each pair was randomized.

At the beginning of the experiment, participants filled a questionnaire designed to measure their subjective level of resources in childhood and adulthood on 100-point scales (Griskevicius, Tybur, Delton, & Robertson, 2011). This questionnaire has been used in multiple studies demonstrating a link between childhood poverty and adult behaviors, independently of participants' current level of resources (e.g., Griskevicius et al., 2013). Following the literature, we used participants' current level of resources as a control variable to investigate the independent effect of childhood poverty. In addition, participants also completed two other validated questionnaires: one on negative household affect (Bernstein et al., 1994) and one on parental care (Mikach & Bailey, 1999). At the end of the experiment, participants had to indicate their agreement with the sentence "I think having a strong leader who does not have to bother with parliament and elections is a good thing" on a 4-point scale. This self-reported measure of participants' preference for a strong leader was used to assess the link between our experimental measure and explicit political attitudes. Extreme authoritarianism was defined as strong agreement with this sentence.

Eighty-six participants were excluded for not meeting basic quality checks such as indicating their gender consistently at two different times of the procedure. Reaction times inferior to 200 ms were discarded (mean reaction time of the included participants before filtering: 1.16 ± 0.02 s). Thirty-six participants were excluded for pressing the same key on >90% of the trials, thirty-five participants were excluded for having given responses with a reaction time superior to 200 ms in less than half the trials and one participant was excluded for not having given any response with a reaction time superior to 200 ms. Twenty-

four participants were excluded for not having given any response. The analyses were performed on 818 participants.

2.1.3. Childhood deprivation measure

In economics, poverty is defined as an income below 50% of the median income of the population (e.g., Alesina, Glaeser, & Sacerdote, 2001). We applied this definition to participants' subjective levels of resources. The subjective poverty line was computed on a sample constituted of the 1006 participants from Study 2 and 685 additional participants drawn from the same population ($N = 1691$ in total). Poverty thresholds were set at a score of 65.5 for *Childhood Deprivation (Deprived or Not Deprived Childhood Groups)* and 76.5 for *Current Resources (Low or Middle Current Resources Groups)*. 163 participants were included in the *Deprived Childhood Group* and 649 in the *Not Deprived Childhood Group*. In the older subsample (662 participants), 138 participants were included in the *Deprived Childhood Group* and 524 in the *Not Deprived Childhood Group*.

2.1.4. Subjective Maltreatment and Parental Care

Two categories, similar to those for Childhood and Current Deprivation were created for Subjective Maltreatment and Parental Care. The *Negative Household Affect* threshold was set to a score of 115 resulting in 95 participants included in the *Negative Household Affect Group* and 723 included in the *Not Negative Household Affect Group*. For Parental Care, the threshold was set to a score of 125 and 81 participants were included in the *Low Parental Care Group* and 737 in the *Middle Parent Care Group*.

2.1.5. Pre-analyses

We first checked that participants from the two Childhood Deprivation Groups did not differ on basic task performance measures. As no objective measure of accuracy was available for this preference task, we used mean reaction times and number of valid trials as measures of task performance. We found that the two groups did not differ on these basic performance measures (Reaction times - *Not Deprived Childhood Group*: $M = 1164.65 \pm 18.93$; *Deprived Childhood Group*: $M = 1159.69 \pm 37.95$; $t(816) = 0.23$, $p > 0.250$; number of valid trials - *Not Deprived Childhood Group*: $M = 33.25 \pm 0.25$; *Deprived Childhood Group*: $M = 33.75 \pm 0.47$; $t(816) = -0.12$, $p > 0.250$). Therefore, it appears that the two groups differed neither on their attention to or in their comprehension of the task.

2.1.6. Analyses

As recommended by McFadden in his seminal paper on choice experiments (McFadden, 1980), choices were analysed using mixed logit logistic regression for each subject, taking trial number and face position as random factors and faces' levels of trustworthiness and dominance as regressors. We measured the effect of *Childhood Deprivation* on the probability of choosing a strong leader (a more dominant and less trustworthy leader) by computing this probability for each individual based on her individual coefficients. Importantly, this computed probability corresponds to the probability of choosing the strong leader in any pair of faces, i.e., independently of the avatar's level of dominance and trustworthiness. In other words, this measure reflects participants' general preference for a strong leader. The effect of *Childhood Deprivation* and the controlling factors was assessed using logistic regressions on the probability of choosing a strong leader. Finally, we used similar logistic regressions on extreme authoritarianism to assess the effect of *Childhood Deprivation* on self-reported authoritarianism.

2.2. Results

Strong leaders (defined, as in Study 1, as a more dominant and less trustworthy leader) had a lower chance of being chosen than more trustworthy and less dominant leaders in general (probability of choosing a strong leader: $M = 0.36 \pm 0.01$, $t(817) = -35.39$, $p < 0.001$). As

expected, participants who had experienced childhood poverty were more likely to choose a strong leader ($b_{ChildDep} = 0.10 \pm 0.08$, $z = 2.49$, $p = 0.013$; Fig. 1A-B).

Finally, genetic studies have shown that family environment has less influence on political attitudes after the individual has left the family house (see Hatemi & McDermott, 2012 for a review on this topic). To confirm the existence of a long-term effect of childhood poverty, we ran our analyses after excluding participants who were younger than 30 years old at the time of testing. The impact of childhood poverty on leader preferences was also present in these older participants ($b_{ChildDep} = 0.11 \pm 0.09$, $z = 2.52$, $p = 0.012$), which confirms the long-term impact of childhood adversity. Importantly, this effect was preserved after controlling for participants' current level of resources both in the complete sample ($b_{ChildDep} = 0.09 \pm 0.08$, $z = 2.25$, $p = 0.025$) and in the older subsample ($b_{ChildDep} = 0.11 \pm 0.09$, $z = 2.34$, $p = 0.019$).

To assess the role of potential confounds, we tested the effect of childhood deprivation after controlling for participants' current poverty and level of education as well as variables assessing participants' childhood affective environment (parental care or negative household affect). The effect of childhood deprivation was robust to the inclusion of these variables, both in the complete sample ($b_{ChildDep} = 0.09 \pm 0.08$, $z = 2.17$, $p = 0.030$) and in the older subsample ($b_{ChildDep} = 0.09 \pm 0.09$, $z = 2.08$, $p = 0.038$).

We then further tested the link between childhood deprivation, our experimental measure of preference for strong leaders, and participants' self-reported extreme authoritarianism. As expected, participants with a higher probability of choosing a strong leader during the task were more likely to endorse extreme authoritarianism ($b_{Prob} = 2.06 \pm 1.59$, $z = 2.55$, $p = 0.011$). As for facial preferences, childhood poverty predicted extreme authoritarianism both before (complete sample: $b_{ChildDep} = 0.56 \pm 0.42$, $z = 2.62$, $p = 0.009$; older subsample: $b_{ChildDep} = 0.50 \pm 0.46$, $z = 2.14$, $p = 0.033$) and after including the control variables (complete sample: $b_{ChildDep} = 0.59 \pm 0.43$, $z = 2.65$, $p = 0.008$; older subsample: $b_{ChildDep} = 0.47 \pm 0.46$, $z = 1.97$, $p = 0.048$).

2.3. Conclusions

This study revealed a long-term impact of childhood deprivation on adult leader preferences using both facial preferences and self-reported extreme authoritarianism. Importantly, the effect of childhood deprivation was still present after controlling for potential confounding factors such as parental care and level of education. To assess the robustness of the link between extreme authoritarianism and childhood deprivation, we analysed responses from an independent large-scale sociological survey.

3. Study 3

The impact of childhood deprivation on the preference for strong leaders was analysed using a large scale sociological survey conducted on 66,281 respondents living in 46 European countries (the European value survey, Wave 4, [European Values Study Longitudinal Data File 1981–2008 \(EVS 1981–2008\)](#), 2015).

3.1. Material and methods

3.1.1. Dataset and measures

We analysed Wave 4 (collected between 2008 and 2010) of the European Value Survey ([European Values Study Longitudinal Data File 1981–2008 \(EVS 1981–2008\)](#), 2015). The analyses were performed on 66,281 respondents (mean number of respondents per country: 1440.89 ± 94.21) from 46 European countries. Extreme authoritarianism was measured exactly as in Study 2. Childhood Deprivation was measured using the question: "Parent(s) had problems to make ends meet" in the European Value Survey. A childhood poverty threshold was computed using the same method as in Study 2. Because responses

to the childhood deprivation question may have different values for different countries, the poverty threshold was computed for each country separately. This method was also applied to three other subjective childhood measures (parental political involvement, parental interest for the news and parental interest for books) and to three subjective adulthood measures (current level of religiosity, life satisfaction and confidence in political institutions).

3.1.2. Analyses

The effect of Childhood Deprivation on extreme authoritarianism was computed using two-level analyses: logistic regressions were run for each country separately and coefficient significance was assessed using *t*-tests at the group level. Countries for which R identified fitting errors were excluded ($N = 2$ in the all sample analysis with the control variables and $N = 3$ in the older sample analysis with the control variable).

3.2. Results

Our analyses confirmed the effect of childhood deprivation on extreme authoritarianism (all respondents: $M = 0.21 \pm 0.10$, $t(45) = 4.07$, $p < 0.001$; Fig. 1C; respondents older than 30 years old: $M = 0.18 \pm 0.12$, $t(45) = 2.78$, $p = 0.008$). As in Study 2, we further tested the effect of childhood deprivation after controlling for participants' level of education, income level as well as childhood measures assessing parents' cultural capital (interest in books, news and politics as well as their level of education). In addition, we also included additional variables measuring participants' current level of stress and resources (recent episode of unemployment and subjective life satisfaction) as well as variables measuring participants' attitudes that could potentially mediate the link between childhood deprivation and extreme authoritarianism (religiosity, Altemeyer & Hunsberger, 1992, and confidence in political institutions, Knigge, 1998). These analyses confirmed a specific effect of childhood deprivation on adults' political attitudes (all respondents: $M = 0.11 \pm 0.10$, $t(43) = 2.11$, $p = 0.040$; older respondents: $M = 0.13 \pm 0.11$, $t(42) = 2.25$, $p = 0.029$).

4. Discussion

With the present experiments, we aimed to understand the effect of childhood environment on political preferences. In line with our hypothesis, we found that experiencing poverty during childhood was associated with an increased preference for dominant leaders. These results are consistent with the literature on the effect of external threats on political preferences (Laustsen & Petersen, 2015; Perrin, 2005; Sales, 1973; Van de Vyver, Houston, Abrams, & Vasiljevic, 2016). For example, analyses of historical records have revealed increased authoritarianism during periods of social and economic threat (such as the Great Depression, the late 1960s and the late 1970s in the US, Sales, 1973). Similarly, the acute threats of 09/11/2001 in the US (Perrin, 2005) and of the 2005 London bombings (Van de Vyver et al., 2016) are associated with increased conservatism and authoritarianism.

However, the present studies suggest that the effect of external threats is much more pervasive than previously thought. Our results indeed reveal an immediate effect of early adverse experiences on children's preferences (Study 1) as well as a postponed effect on adults' political preferences (Studies 2–3). This postponed effect was evidenced both in the experimental task and in self-reports and suggests the existence of a direct relationship between childhood environmental harshness and political attitudes. However, it is important to note that no causal link between childhood deprivation and political attitudes can be drawn based on these correlational studies. Indeed, we cannot exclude the existence of unobserved variables that might impact both adult political attitudes and childhood environmental harshness independently. In an effort to control for potential confounding factors, we included various control variables in our models. These additional

analyses show that the effect of childhood environmental harshness on political attitudes was independent of participants' current environment (current status and recent life events), cultural factors (education level and parenting style), and confidence in political institutions. This robustness check suggests a limited impact of omitted observable variables in our context.

In addition, it is worth noting that genetic factors may also be partly responsible for our results. Indeed, authoritarianism, lower socio-economic status, as well as educational attainment and cognitive skills are all partly genetically determined (Hatemi & McDermott, 2012; Marioni et al., 2014; Plomin & Deary, 2015; Rietveld et al., 2013). Genetic factors also mediate the association between lower socio-economic status and many social and cognitive traits (Hill et al., 2016; Krapohl & Plomin, 2016; Trzaskowski et al., 2014). Therefore, it is possible that the evidenced association between childhood deprivation and authoritarianism is due to genetic mechanisms, such as the co-transmission of genetic factors associated with a lower socio-economic status and those linked with higher degrees of authoritarianism. The present findings would thus be strengthened by future research involving external shocks in experimental settings or natural experiments.

In addition, we have to acknowledge the small sample size of the study on children (Study 1, $N = 39$) and that further experiments are needed to replicate the early sensitivity of leader preferences to environmental factors. The present results should also be confirmed by using different measures of childhood deprivation in adults. Indeed, even if actual and retrospective socio-economic status appear to have similar effects on adult outcomes (Cohen, Janicki-Deverts, Chen, & Matthews, 2010; Mittal, Griskevicius, Simpson, Sung, & Young, 2015), contrasting perceived and objective deprivation would allow us to test whether individuals' subjective perception of environmental harshness has a more potent effect on behavior than objective environmental harshness.

Despite these caveats, the present findings are the first to point to possible mechanisms that may lead to an early emergence of authoritarianism and orient individuals' behavior throughout their lifespan. This idea is consistent with longitudinal studies showing that right-wing authoritarianism is stable over months (Onraet et al., 2014; Sibley et al., 2007) and years (Mirisola et al., 2014) and with brain imaging research showing substantial differences between adults who experienced stressful environments as children in brain areas involved in face evaluations (Javanbakht et al., 2015; Kim et al., 2013; Mende-Siedlecki, Said, & Todorov, 2012). Therefore, the effect of childhood poverty evidenced in our study may rely on deep changes in brain functioning. In the present studies, however, the immediate and long-term effects of childhood deprivation on political preferences were evidenced using a cross-sectional design. Longitudinal studies are thus needed to confirm the association between childhood deprivation and leader preferences, and to track its evolution across the lifespan.

In addition, our studies may also shed light on the ecological bases of authoritarianism and on the possible adaptive value of such political attitudes. Indeed, we provide further evidence that leader choice is highly responsive to environmental threats and that the plasticity of leader preferences is also present in children. Our results thus extend existing data showing that more masculine and more dominant leaders are preferred in times of threat (Laustsen & Petersen, 2015; Little, Burriss, Jones, & Roberts, 2007; Re et al., 2013) and that right-wing political attitudes increase with perceived threat (Onraet et al., 2014; Sibley et al., 2007). Interestingly, Van Vugt, Hogan, and Kaiser (2008) have suggested that these contextual effects could be construed as the adaptive responses of an evolved leader choice mechanism. In this perspective, choosing a strong leader would be an adaptive response to an increased need for group coordination in harsh environments (Laustsen & Petersen, 2015). Our results showing an early plasticity of leader preferences in response to external threats are compatible with this idea but alternative explanations cannot be ruled out at this stage.

Indeed, leader preferences could be just as well guided by other cognitive mechanisms. Specifically, it has been demonstrated that leader choice is mainly determined by broad competence judgments that are independent from the more specific assessment of leadership qualities (see Olivola & Todorov, 2010 for a review). In our context, this might suggest that dominant leaders are more appealing in harsh environments simply because they are perceived as more competent in these specific environments. This indeed makes sense from an ecological point of view: stressful environments are more competitive and less cooperative (Evans, 2004), which favors dominant individuals, who are likely to outperform others and to acquire larger amounts of resources. This alternative view would predict that leader preferences are guided by competence evaluations that bear no relevance to people's need for group coordination. Further research testing the optimality of leader preferences under different environmental constraints is thus needed to tease apart these two interpretations.

Finally, these studies may offer a new way of explaining long-term changes in political attitudes. The early calibration of social preferences indeed suggests that political life is not only influenced by current circumstances (recessions, wars, acts of terrorism, etc.) but also by early life conditions. This may have a protective effect against aggravating circumstances when voters experienced favorable situations as children, but it can also hinder the beneficial effects of current improvements in people's environment. For instance, in the early 1970's, after 30 years of rising prosperity, more and more people started to embrace 'post-materialist values' such as emphasized autonomy and self-expression (Inglehart, 2015). Yet, cohorts born and raised before World War II, who had experienced the recession of the 30's and wartime restrictions, continued to adhere to 'materialist values', emphasizing economic security and authoritarianism (Inglehart, 2015). Conversely, since the 1970's, Western countries have experienced >30 years of economic stagnation and rising inequalities (Twenge & Kasser, 2013). The effect of stress on people's political attitudes may have been mitigated for a while by the presence of large cohorts of individuals brought up during times of rising prosperity (1945–1975) and whose political attitudes were still influenced by a favorable childhood. But as the younger generations (generations X and Y) are coming of age, and starting to vote, authoritarian candidates may become increasingly popular.

5. Future directions

The present studies are a first step towards establishing an association between early exposure to adversity and authoritarianism. Two lines of research should be developed to further investigate the ontology and development of social attitudes from an ecological perspective. On the one hand, the same evolutionary reasoning could be applied to test whether other environmental factors known to affect adult political preferences also influence leader preferences when they are experienced during childhood. For instance, it would be particularly interesting to test whether exposure to pathogens or diseases during childhood also increase preference for autocracy, as it does in adults (Thornhill, Fincher, Murray, & Schaller, 2010; Thornhill et al., 2009). Based on this perspective, another open question is to understand how people's current and childhood environment interact to produce different political attitudes. Indeed, recent research in evolutionary psychology has suggested that childhood experiences may activate different behavioral responses to current environmental threats (Del Giudice et al., 2011; Griskevicius et al., 2011; Griskevicius et al., 2013). Based on these results, one could hypothesize that external threats will shape political attitudes differently depending on people's prior childhood environment.

On the other hand, our study focused on the ontology of one particular political attitude but it is worth noting that different political and social attitudes often covary. For instance, preference for authoritarian leaders is associated with a preference for conformity and with higher degrees of prejudice against minorities (Altemeyer, 1996; Cohrs,

Moschner, Maes, & Kielmann, 2005; Duckitt, 2006; Laythe, Finkel, & Kirkpatrick, 2001). However, little is known about whether these traits constitute a behavioral syndrome, i.e. the output of a coordinated strategy (Baumard & Chevallier, 2015; Griskevicius et al., 2013; Pepper & Nettle, 2017; Réale et al., 2010), or if it is rather an aggregation of independent phenotypes that respond to similar environmental signals. This behavioral syndrome approach has already been successfully applied to describe the co-variation of somatic and reproductive behaviors and to shed light on the pressures that may have shaped its evolution (Brumbach, Figueredo, & Ellis, 2009; Ellis, Figueredo, Brumbach, & Schlomer, 2009; Mell, Safra, Algan, Baumard, & Chevallier, 2017; Simpson, Griskevicius, Kuo, Sung, & Andrew, 2012). Addressing the question of the co-variation of this set of political attitudes in an ecological perspective would thus provide insight on the evolutionary roots of political behavior.

Acknowledgments

This study was supported by the Institut d'Études Cognitives (ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*), the Institut national de la santé et de la recherche médicale (Inserm), the Fondation Roger de Spoelberch, an ERC consolidator grant (no. 8657) and an "Action Incitative" funding from the École Normale Supérieure. We are grateful to Anne L'Hôte for programming Study 2 and to Dr. Bahador Bahrami for providing us the stimuli. All the data and the analysis scripts are available at <https://osf.io/yut9p/> except the European Value Survey data which are downloadable from the official European Value Survey website (<http://www.europeanvaluesstudy.eu>).

Appendix A. Supplementary information

Supplementary information to this article can be found online at <http://dx.doi.org/10.1016/j.evolhumbehav.2017.05.001>.

References

- Adair, L. S. (2001). Size at birth predicts age at menarche. *Pediatrics*, 107(4), e59.
- Alesina, A., Glaeser, E., & Sacerdote, B. (2001). Why doesn't the United States have a European-style welfare state? *Brookings papers on economic activity*, 187.
- Altemeyer, B. (1996). *The authoritarian specter*. Vol. vii. Cambridge, MA, US: Harvard University Press.
- Altemeyer, B., & Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *The International Journal for the Psychology of Religion*, 2(2), 113–133. http://dx.doi.org/10.1207/s15327582ijpr0202_5.
- Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, 323(5918), 1183. <http://dx.doi.org/10.1126/science.1167748>.
- Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R. A., ... Sultan, S. E. (2004). Developmental plasticity and human health. *Nature*, 430(6998), 419–421. <http://dx.doi.org/10.1038/nature02725>.
- Baumard, N., & Chevallier, C. (2015). The nature and dynamics of world religions: A life-history approach. *Proceedings of the Royal Society B*, 282(1818), 20151593. <http://dx.doi.org/10.1098/rspb.2015.1593>.
- Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., ... Ruggiero, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *The American Journal of Psychiatry*, 151(8), 1132–1136. <http://dx.doi.org/10.1176/ajp.151.8.1132>.
- Brumbach, B. H., Figueredo, A. J., & Ellis, B. J. (2009). Effects of harsh and unpredictable environments in adolescence on development of life history strategies. *Human Nature*, 20(1), 25–51. <http://dx.doi.org/10.1007/s12110-009-9059-3>.
- Cohen, S., Janicki-Deverts, D., Chen, E., & Matthews, K. A. (2010). Childhood socioeconomic status and adult health. *Annals of the New York Academy of Sciences*, 1186(1), 37–55. <http://dx.doi.org/10.1111/j.1749-6632.2009.05334.x>.
- Cohrs, J. C., Moschner, B., Maes, J., & Kielmann, S. (2005). The motivational bases of right-wing authoritarianism and social dominance orientation: Relations to values and attitudes in the aftermath of September 11, 2001. *Personality and Social Psychology Bulletin*, 31(10), 1425–1434. <http://dx.doi.org/10.1177/0146167205275614>.
- Del Giudice, M., Ellis, B. J., & Shirtcliff, E. A. (2011). The adaptive calibration model of stress reactivity. *Neuroscience & Biobehavioral Reviews*, 35(7), 1562–1592. <http://dx.doi.org/10.1016/j.neubiorev.2010.11.007>.
- Duckitt, J. (2006). Differential effects of right wing authoritarianism and social dominance orientation on outgroup attitudes and their mediation by threat from and competitiveness to outgroups. *Personality and Social Psychology Bulletin*, 32(5), 684–696. <http://dx.doi.org/10.1177/0146167205284282>.

- Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk. *Human Nature*, 20(2), 204–268. <http://dx.doi.org/10.1007/s12110-009-9063-7>.
- European Values Study Longitudinal Data File 1981–2008 (EVS 1981–2008) (2015). EVS. GESIS data archive, Cologne. ZA4804 data file version 3.0.0. <http://dx.doi.org/10.4232/1.12253>.
- Evans, G. W. (2004). The environment of childhood poverty. *American Psychologist*, 59(2), 77–92. <http://dx.doi.org/10.1037/0003-066X.59.2.77>.
- Frankenhuis, W. E., Panchanathan, K., & Nettle, D. (2016). Cognition in harsh and unpredictable environments. *Current Opinion in Psychology*, 7, 76–80. <http://dx.doi.org/10.1016/j.copsyc.2015.08.011>.
- Griskevicius, V., Ackerman, J. M., Cantú, S. M., Delton, A. W., Robertson, T. E., Simpson, J. A., ... Tybur, J. M. (2013). When the economy falters, do people spend or save? Responses to resource scarcity depend on childhood environments. *Psychological Science*, 0956797612451471. <http://dx.doi.org/10.1177/0956797612451471>.
- Griskevicius, V., Tybur, J. M., Delton, A. W., & Robertson, T. E. (2011). The influence of mortality and socioeconomic status on risk and delayed rewards: A life history theory approach. *Journal of Personality and Social Psychology*, 100(6), 1015–1026. <http://dx.doi.org/10.1037/a0022403>.
- Hatemi, P. K., & McDermott, R. (2012). The genetics of politics: Discovery, challenges, and progress. *Trends in Genetics*, 28(10), 525–533. <http://dx.doi.org/10.1016/j.tig.2012.07.004>.
- Hill, W. D., Hagenaaars, S. P., Marioni, R. E., Harris, S. E., Liewald, D. C. M., Davies, G., ... Deary, I. J. (2016). Molecular genetic contributions to social deprivation and household income in UK Biobank. *Current Biology*, 26(22), 3083–3089. <http://dx.doi.org/10.1016/j.cub.2016.09.035>.
- Inglehart, R. (2015). *The silent revolution: Changing values and political styles among western publics*. Princeton University Press.
- Javanbakht, A., King, A. P., Evans, G. W., Swain, J. E., Angstadt, M., Phan, K. L., & Liberzon, I. (2015). Childhood poverty predicts adult amygdala and frontal activity and connectivity in response to emotional faces. *Frontiers in Behavioral Neuroscience*, 9. <http://dx.doi.org/10.3389/fnbeh.2015.00154>.
- Kim, P., Evans, G. W., Angstadt, M., Ho, S. S., Sripada, C. S., Swain, J. E., ... Phan, K. L. (2013). Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. *Proceedings of the National Academy of Sciences*, 110(46), 18442–18447. <http://dx.doi.org/10.1073/pnas.1308240110>.
- Knigge, P. (1998). The ecological correlates of right-wing extremism in Western Europe. *European Journal of Political Research*, 34(2), 249–279. <http://dx.doi.org/10.1023/A:1006953714624>.
- Krapohl, E., & Plomin, R. (2016). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, 21(3), 437–443. <http://dx.doi.org/10.1038/mp.2015.2>.
- Laustsen, L., & Petersen, M. B. (2015). Does a competent leader make a good friend? Conflict, ideology and the psychologies of friendship and followership. *Evolution and Human Behavior*, 36(4), 286–293. <http://dx.doi.org/10.1016/j.evolhumbehav.2015.01.001>.
- Laythe, B., Finkel, D., & Kirkpatrick, L. A. (2001). Predicting prejudice from religious fundamentalism and right-wing authoritarianism: A multiple-regression approach. *Journal for the Scientific Study of Religion*, 40(1), 1–10.
- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1), 18–27. <http://dx.doi.org/10.1016/j.evolhumbehav.2006.09.002>.
- Little, A. C., Roberts, S. C., Jones, B. C., & DeBruine, L. M. (2012). The perception of attractiveness and trustworthiness in male faces affects hypothetical voting decisions differently in wartime and peacetime scenarios. *The Quarterly Journal of Experimental Psychology*, 65(10), 2018–2032. <http://dx.doi.org/10.1080/17470218.2012.677048>.
- Marioni, R. E., Davies, G., Hayward, C., Liewald, D., Kerr, S. M., Campbell, A., ... Deary, I. J. (2014). Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence*, 44, 26–32. <http://dx.doi.org/10.1016/j.intell.2014.02.006>.
- McFadden, D. (1980). Econometric models for probabilistic choice among products. *The Journal of Business*, 53(3), S13–S29.
- Mell, H., Safra, L., Algan, Y., Baumard, N., & Chevallier, C. (2017). Childhood environmental harshness predicts coordinated health and reproductive strategies: A cross-sectional study of a nationally representative (France) sample. Retrieved from <https://osf.io/4rzfy/?action=download> (Retrieved from Osf. Io/urdky).
- Mende-Siedlecki, P., Said, C. P., & Todorov, A. (2012). The social evaluation of faces: A meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience*, nsr090. <http://dx.doi.org/10.1093/scan/nsr090>.
- Mikach, S. M., & Bailey, J. M. (1999). What distinguishes women with unusually high numbers of sex partners? *Evolution and Human Behavior*, 20(3), 141–150. [http://dx.doi.org/10.1016/S1090-5138\(98\)00045-2](http://dx.doi.org/10.1016/S1090-5138(98)00045-2).
- Mirisola, A., Rocco, M., Russo, S., Spagna, G., & Vieno, A. (2014). Societal Threat to Safety, Compensatory Control, and Right-Wing Authoritarianism. *Political Psychology*, 35(6), 795–812. <http://dx.doi.org/10.1111/pops.12048>.
- Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, 109(4), 604–621. <http://dx.doi.org/10.1037/pspi0000028>.
- Murray, D. R., Schaller, M., & Suedfeld, P. (2013). Pathogens and politics: Further evidence that parasite prevalence predicts authoritarianism. *PLoS One*, 8(5), e62275. <http://dx.doi.org/10.1371/journal.pone.0062275>.
- Nettle, D., Coall, D. A., & Dickins, T. E. (2011). Early-life conditions and age at first pregnancy in British women. *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1712), 1721–1727. <http://dx.doi.org/10.1098/rspb.2010.1726>.
- Nettle, D., Frankenhuus, W. E., & Rickard, I. J. (2013). The evolution of predictive adaptive responses in human life history. *Proceedings of the Royal Society B*, 280(1766), 20131343. <http://dx.doi.org/10.1098/rspb.2013.1343>.
- Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans prefer republican-looking leaders political facial stereotypes predict candidate electoral success among right-leaning voters. *Social Psychological and Personality Science*, 3(5), 605–613. <http://dx.doi.org/10.1177/194850611432770>.
- Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, 34(2), 83–110. <http://dx.doi.org/10.1007/s10919-009-0082-1>.
- Onraet, E., Dhont, K., & Hiel, A. V. (2014). The relationships between internal and external threats and right-wing attitudes a three-wave longitudinal study. *Personality and Social Psychology Bulletin*, 40(6), 712–725. <http://dx.doi.org/10.1177/0146167214524256>.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087–11092. <http://dx.doi.org/10.1073/pnas.0805664105>.
- Pepper, G. V., & Nettle, D. (2017). The behavioural constellation of deprivation: Causes and consequences. *The Behavioral and Brain Sciences*, 1–72. <http://dx.doi.org/10.1017/S0140525X1600234X>.
- Perrin, A. J. (2005). National threat and political culture: Authoritarianism, anti-authoritarianism, and the September 11 attacks. *Political Psychology*, 26(2), 167–194. <http://dx.doi.org/10.1111/j.1467-9221.2005.00414.x>.
- Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, 20(1), 98–108. <http://dx.doi.org/10.1038/mp.2014.105>.
- Réale, D., Garant, D., Humphries, M. M., Bergeron, P., Careau, V., & Montiglio, P.-O. (2010). Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 365(1560), 4051–4063. <http://dx.doi.org/10.1098/rstb.2010.0208>.
- Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial cues to perceived height influence leadership choices in simulated war and peace contexts. *Evolutionary Psychology*, 11(1), 89–103. <http://dx.doi.org/10.1177/147470491301100109>.
- Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., ... Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, 340(6139), 1467–1471. <http://dx.doi.org/10.1126/science.1235488>.
- Sales, S. M. (1973). Threat as a factor in authoritarianism: An analysis of archival data. *Journal of Personality and Social Psychology*, 28(1), 44–57. <http://dx.doi.org/10.1037/h0035588>.
- Sibley, C. G., Wilson, M. S., & Duckitt, J. (2007). Effects of dangerous and competitive worldviews on right-wing authoritarianism and social dominance orientation over a five-month period. *Political Psychology*, 28(3), 357–371. <http://dx.doi.org/10.1111/j.1467-9221.2007.00572.x>.
- Simpson, J. A., Griskevicius, V., Kuo, S. I. -C., Sung, S., & Andrew, W. (2012). Evolution, stress, and sensitive periods: The influence of unpredictability in early versus late childhood on sex and risky behavior. *Developmental Psychology*, 48(3), 674–686. <http://dx.doi.org/10.1037/a0027293>.
- Sloboda, D. M., Hart, R., Doherty, D. A., Pennell, C. E., & Hickey, M. (2007). Age at menarche: Influences of prenatal and postnatal growth. *The Journal of Clinical Endocrinology & Metabolism*, 92(1), 46–50. <http://dx.doi.org/10.1210/jc.2006-1378>.
- Stewart, L. H., Ajina, S., Getov, S., Bahrami, B., Todorov, A., & Rees, G. (2012). Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology: General*, 141(4), 715–727. <http://dx.doi.org/10.1037/a0027950>.
- Terrizzi, J. A., Jr., Shook, N. J., & McDaniel, M. A. (2013). The behavioral immune system and social conservatism: A meta-analysis. *Evolution and Human Behavior*, 34(2), 99–108. <http://dx.doi.org/10.1016/j.evolhumbehav.2012.10.003>.
- Thornhill, R., Fincher, C. L., & Aran, D. (2009). Parasites, democratization, and the liberalization of values across contemporary countries. *Biological Reviews*, 84(1), 113–131. <http://dx.doi.org/10.1111/j.1469-185X.2008.00062.x>.
- Thornhill, R., Fincher, C. L., Murray, D. R., & Schaller, M. (2010). Zoonotic and non-zoonotic diseases in relation to human personality and societal values: Support for the parasite-stress model. *Evolutionary Psychology*, 8(2), 147470491000800201. <http://dx.doi.org/10.1177/147470491000800201>.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623–1626. <http://dx.doi.org/10.1126/science.1110589>.
- Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., ... Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, 42, 83–88. <http://dx.doi.org/10.1016/j.intell.2013.11.002>.
- Twenge, J. M., & Kasser, T. (2013). Generational changes in materialism and work centrality, 1976–2007 associations with temporal changes in societal insecurity and materialistic role modeling. *Personality and Social Psychology Bulletin*, 39(7), 883–897. <http://dx.doi.org/10.1177/0146167213484586>.
- Van de Vyver, J., Houston, D. M., Abrams, D., & Vasiljevic, M. (2016). Boosting belligerence how the July 7, 2005, London bombings affected liberals' moral foundations and prejudice. *Psychological Science*, 27(2), 169–177. <http://dx.doi.org/10.1177/0956797615615584>.
- Van Vugt, M., Hogan, R., & Kaiser, R. B. (2008). Leadership, followership, and evolution: Some lessons from the past. *American Psychologist*, 63(3), 182–196. <http://dx.doi.org/10.1037/0003-066X.63.3.182>.

1 **Supplementary Information**

2 Impact of gender on facial features preferences

3 As it has been shown that individuals are more likely to vote for candidates that look similar to
4 them (Bailenson, Iyengar, Yee, & Collins, 2009) and because dominant and less trustworthy
5 faces have been demonstrated to be more masculine (Oosterhof & Todorov, 2008), we checked
6 the robustness of our results by adding gender as a covariable. Confirming the impact of
7 childhood deprivation on leader preferences, adding gender as a regressor did not significantly
8 impact the effect of childhood deprivation on leader preferences either in children ($z = -0.24, p$
9 $> .250; b_{ChildDep} = 0.58 \pm 0.29, z = 1.95, p = .052$), adults ($z = -0.10, p > .250; b_{ChildDep} = 0.11 \pm 0.08,$
10 $z = 2.66, p = .008$) or older adults ($z = -0.11, p > .250; b_{ChildDep} = 0.16 \pm 0.09, z = 2.69, p = .007$).

11

12 Study 1 – Parental income analysis

13 We ran a logistic regression analysis on children's leader choices taking parental income as a
14 regressor. This analysis confirmed the impact of childhood deprivation on individuals'
15 preference for strong leader: experiencing childhood deprivation increased by 25.54% the
16 probability of choosing a strong leader ($b = 0.69 \pm 0.56, z = 2.41, p = .016$).

17

18 Study S1 - Experiment design validation

19 In order to assess whether a discrete choice experiment would replicate the well-documented
20 impact of exposure to external threats on leader preferences (Laustsen & Petersen, 2015; Little,

21 Burriss, Jones, & Roberts, 2007; Little, Roberts, Jones, & DeBruine, 2012; Re, DeBruine, Jones, &
22 Perrett, 2013), we first ran a replication of Little et al.'s studies (Little et al., 2007, 2012)
23 showing that war scenarios induce a preference for more dominant and less trustworthy
24 leaders.

25

26 *Participants*

27 60 participants (25 women), aged 20 to 60 years old (mean age: 32.58 ± 2.57 years) were
28 recruited via MTurk to participate in this online study. Each participant received a description of
29 the study, provided their informed consent before starting the experiment, and received \$2 for
30 this 30-minute experiment (which corresponds to the \$6.00 hourly wage of MTurk).

31

32 *Procedure*

33 The experiment was programmed on Qualtrics and was similar to Study 2 except that the set of
34 pairs was presented twice, in two separate blocks, one in a war-time scenario and one in a
35 peace-time scenario ("Who would you vote for in a time of war/peace?"). These scenarios were
36 similar to those used in (Little et al., 2007, 2012), and the order of the two blocks was
37 randomized between participants.

38

39 Reaction times inferior to 200 ms were discarded (mean reaction time of the included
40 participants before filtering: 0.50 ± 0.01 s). Four participants were excluded from the analysis for

41 pressing the same key in more than 90% of the trials, two participants were excluded for having
42 no response with a reaction time superior to 200ms and one participant was excluded for not
43 having given any response. The analyses were performed on 53 participants.

44

45 *Analyses & Results*

46 As in Study 2, choices were analysed using two-level analyses. First, mixed logit logistic
47 regressions, taking trial number and face position as random factors, were ran for each
48 participant. Levels of trustworthiness (Trustworthiness) and dominance (Dominance) as well as
49 the Scenario (war-time or peace-time) were used as regressors. The coefficients from these
50 regressions were then used to compute the probability of choosing a strong leader (i.e., a 1-
51 point more dominant and 1-point less trustworthy leader). A beta regression was then run on
52 these probabilities to assess the impact of the experimental manipulation.

53

54 The analysis of the coefficients revealed that in times of peace, strong leaders had a lower
55 probability of being chosen than more trustworthy and less dominant faces (Probability of
56 choosing a strong leader: $M = 0.34 \pm 0.04$, $t(52) = -8.79$, $p < .001$). As expected, the scenario
57 impacted these choices: in times of war, strong leaders were more likely to be chosen ($b_{War} =$
58 2.72 ± 0.99 , $z = 5.39$, $p < .001$). Indeed, strong leaders were chosen significantly above chance in
59 the war scenario ($M = 0.59 \pm 0.06$, $t(52) = 3.62$, $p < .001$). Interestingly, in wartime, both
60 dominance and untrustworthiness were valued in leaders ($b_T = -0.30 \pm 0.18$, $t(52) = -3.35$, $p =$
61 $.002$; $b_D = 0.15 \pm 0.11$, $t(52) = 2.76$, $p = .008$) while it was the opposite in peace time ($b_T = 0.57 \pm$

62 0.11, $t(52) = 10.12$, $p < .001$; $b_D = -0.13 \pm 0.09$, $t(52) = -2.81$, $p = .007$). In other words, in
63 wartime, stronger leaders were preferred, which replicates Little et al.'s 2007 and 2012 findings
64 (Little et al., 2007, 2012).

65

66 *Conclusions*

67 Our new methodology replicates a robust result of the leader choice literature:
68 participants are biased towards strong leaders after war primes (Little et al., 2007, 2012). This
69 finding demonstrates that our online experiment is sensitive to changes in leader preferences
70 after exposure to an external threat.

71

72 Study S2 – The impact of childhood environment: pilot study

73 *Participants*

74 The number of subjects was fixed *a priori* to 200 based on our past experience with between-
75 subjects social choice experiments on MTurk. 200 American participants (66 women) aged 19 to
76 74 years old (33.31 ± 1.63 years), were recruited via Amazon Mechanical Turk to participate in
77 this online study. Each participant received a description of the study, provided their informed
78 consent before starting the experiment, and received \$1.5 for this 15-minute experiment.

79

80 *Procedure*

81 The same experiment as in Study 2 was programmed on Qualtrics. However, in this experiment,
82 participants filled the subjective level of resources questionnaires on 9-point scales
83 (Griskevicius, Tybur, Delton, & Robertson, 2011) and at the end of the experiment. As in Study
84 2, reaction times inferior to 200 ms were discarded (mean reaction time of the included
85 participants before filtering: 0.49 ± 0.02 s). Four participants were excluded for pressing the
86 same key in more than 90% of the trials, four participants were excluded for having responses
87 with a reaction time superior to 200 ms on less than half the trials, six participants were
88 excluded for not having given any responses with a reaction time superior to 200 ms and one
89 participant was excluded for not having given any response. The analyses were performed on
90 183 participants.

91

92 *Childhood Deprivation Measure*

93 The subjective poverty line was defined as in Study 2. We used previously collected responses
94 on 522 MTurkers in addition to the 200 participants to the study to determine the threshold
95 value for the MTurk population. Poverty thresholds were set at a score of 7 for both *Childhood*
96 (*Deprived or Not Deprived Childhood Groups*) and *Current Resources (Low or Middle Current*
97 *Resources Groups)*. 35 participants were included in the *Deprived Childhood Group* and 148 in
98 the *Not Deprived Childhood Group*. In the older subsample (99 participants), 24 were included
99 in the *Deprived Childhood Group* and 75 in the *Not Deprived Childhood Group*.

100

101 *Pre-analyses and Analyses*

102 As in Study 2, the two *Childhood Deprivation Groups* did not differ on the number of the
103 analysable trials ($t(181) = 0.35, p > .250$) or on their mean reaction times ($t(181) = 1.34, p =$
104 $.183$), indicating that the two groups performed the task similarly. The same models and
105 analyses as in Study 2 were run in this study. As in Study 2, the probability of choosing a more
106 dominant and less trustworthy leader was computed using the model's logit coefficients for
107 each participant and beta regressions were used to assess the impact of childhood deprivation.

108

109 *Results*

110 As in Study 2, strong leaders were generally less chosen than more trustworthy and less
111 dominant leaders ($M = 0.42 \pm 0.02, t(182) = -7.39, p < .001$). After controlling for Current
112 Resources, the effect of Childhood Deprivation was found as a trend in the total sample ($N =$
113 $183; b_{ChildDep} = 0.19 \pm 0.23, z = 1.66, p = .098$) and as a significant effect in the older subsample
114 ($N = 94; b_{ChildDep} = 0.38 \pm 0.28, z = 2.66, p = .008$) such that participants who suffered from
115 poverty during their childhood were more likely to choose a strong leader.

116

117 *Conclusions*

118 This pilot study revealed that experimentally measured leader preferences using facial features
119 are not only sensitive to experimental manipulations (Study S1, Laustsen & Petersen, 2015;
120 Little et al., 2007, 2012) and explicit political affiliations (Laustsen & Petersen, 2015; Olivola,
121 Sussman, Tsetsos, Kang, & Todorov, 2012) but also to the real environment experienced by

122 participants. Moreover, in line with our hypotheses, childhood deprivation induced a long-term

123 preference for stronger leaders.

124

125 **Table S1 Demographic characteristics of the *Deprived* and *Not Deprived Childhood Groups***

	Deprived Childhood Group	Not Deprived Childhood Group	Difference statistical significance
Number of Children	20	19	
Age	6.88 ± 0.23	6.85 ± 0.18	<i>p</i> > .250
Gender ratio (% of Female)	35 %	53 %	<i>p</i> > .250
Access to running water	25 %	100 %	<i>p</i> < .001
Possession of a washing machine	25 %	100 %	<i>p</i> < .001
Possession of a car	0 %	74 %	<i>p</i> < .001
Meat consumption	Monthly	Daily	<i>p</i> < .001
Cinema outings	Never	Monthly	<i>p</i> < .001
Museum outings	Never	Weakly	<i>p</i> = .009
Mean age of the parents	29.65 ± 1.93	34.74 ± 2.17	<i>p</i> < .001
Maximum education level of the parents	High School	Post High School	<i>p</i> = .002
Number of siblings	1.80 ± 0.47	0.47 ± 0.29	<i>p</i> < .001
Number of people/number of rooms ratio	2.33 ± 0.21	1.84 ± 0.32	<i>p</i> = .010
Separation of the parents	30 %	0 %	<i>p</i> = .031

127 **References:**

128 Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2009). Facial Similarity Between Voters and
129 Candidates Causes Influence. *Public Opinion Quarterly*, nfn064.

130 <https://doi.org/10.1093/poq/nfn064>

131 Griskevicius, V., Tybur, J. M., Delton, A. W., & Robertson, T. E. (2011). The Influence of Mortality
132 and Socioeconomic Status on Risk and Delayed Rewards: A Life History Theory
133 Approach. *Journal of Personality and Social Psychology*, 100(6), 1015–1026.

134 <https://doi.org/10.1037/a0022403>

135 Laustsen, L., & Petersen, M. B. (2015). Does a competent leader make a good friend? Conflict,
136 ideology and the psychologies of friendship and followership. *Evolution and Human
137 Behavior*, 36(4), 286–293. <https://doi.org/10.1016/j.evolhumbehav.2015.01.001>

138 Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting
139 decisions. *Evolution and Human Behavior*, 28(1), 18–27.

140 <https://doi.org/10.1016/j.evolhumbehav.2006.09.002>

141 Little, A. C., Roberts, S. C., Jones, B. C., & DeBruine, L. M. (2012). The perception of
142 attractiveness and trustworthiness in male faces affects hypothetical voting decisions
143 differently in wartime and peacetime scenarios. *The Quarterly Journal of Experimental
144 Psychology*, 65(10), 2018–2032. <https://doi.org/10.1080/17470218.2012.677048>

145 Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans Prefer
146 Republican-Looking Leaders Political Facial Stereotypes Predict Candidate Electoral
147 Success Among Right-Leaning Voters. *Social Psychological and Personality Science*, 3(5),
148 605–613. <https://doi.org/10.1177/1948550611432770>

149 Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of*
150 *the National Academy of Sciences*, *105*(32), 11087–11092.

151 <https://doi.org/10.1073/pnas.0805664105>

152 Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial Cues to Perceived Height
153 Influence Leadership Choices in Simulated War and Peace Contexts. *Evolutionary*
154 *Psychology*, *11*(1), 147470491301100109.

155 <https://doi.org/10.1177/147470491301100109>

156

An evolutionary perspective on leader preferences

The preference for untrustworthy leaders in wartime conditions (Little et al., 2012; Safra et al., 2017) challenges the mainstream hypothesis that leader preferences are the output of an evolved mechanism dedicated to leader choice (King et al., 2009; Van Vugt, 2006). Indeed, theoretical and experimental work has shown that trustworthiness is a necessary feature in effective leaders, an idea that is also found in participants' belief (Nichols & Cottrell, 2014; Van Vugt, 2006; von Rueden, Gurven, Kaplan, & Stieglitz, 2014).

Importantly, it is worth pointing out that the preference for untrustworthy leaders is not the only issue faced by this theory. More precisely, the idea that leader-follower dynamics have constituted a major element of social organization throughout the human species evolutionary history, which is at the very center of this theory (Van Vugt & Kurzban, 2007), is also challenged by theoretical, observational and experimental work. Indeed, it has been shown that efficient cooperation can be sustained, at least for small to medium groups and societies, without leaders (Boehm, 1993; Kaplan & Gangestad, 2005; Kaplan et al., 2009; Woodburn, 1982). In addition, the importance of leadership-irrelevant traits, such as attractiveness, for leader choice further questions the optimality of the leader selection heuristics (Little et al., 2012; White et al., 2013). These issues question the existence of a cognitive mechanism dedicated to leader choice (King et al., 2009; Van

Vugt, 2006) and suggest that individuals rather rely on another cognitive mechanism while choosing a leader.

To examine this idea, we reviewed the literature on leader choice and notably cognitive science experiments investigating variation in leader choices using face evaluations. This review revealed that candidates' competence is the most reliable predictor of leader preferences (Little, 2014; Re et al., 2013; Todorov, Mandisodza, Goren, & Hall, 2005b; White et al., 2013). Based on this observation, we developed a theoretical model on leader choice stating that while choosing a leader, individuals actually select the individual they perceived as the most adapted to the current context, i.e., the task the group has to complete or, for the case of political leaders, the general environment of the voter. This theory hypothesizes that leader preferences are guided by candidates' ability to face the challenges of the current context, but not by the needs of the group or by the individuals' benefits of having a particular individual as a leader. This new theoretical framework makes clear predictions about the contexts in which leader preferences are incongruent with the individuals' and the groups' interest that accurately match both experimental and observational data (H. S. Lewis, 1974; Little et al., 2012; Nevicka, De Hoogh, van Vianen, & Ten Velden, 2013; Safra et al., 2017; White et al., 2013; Zebrowitz, Franklin, & Palumbo, 2015).

This theoretical work has been accepted for submission in Trends in Cognitive Sciences. I contributed to every stage of the development of this new theory.

Why would anyone elect a narcissistic untrustworthy leader?

Safra, L.¹, Baumard, N.², Chevallier, C.¹

¹ *Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, Inserm, U960, Laboratoire de Neurosciences Cognitives (LNC), F-75005 Paris, France*

² *Ecole Normale Supérieure, PSL Research University, Département d'études cognitives, CNRS, UMR8129, Institut Jean-Nicod, F-75005 Paris, France*

Under review in *Trends in Cognitive Sciences*

Abstract

Leader choice is a cornerstone of modern democracies and a central topic in cognitive sciences. The standard view posits that leader choice is guided by an adaptive mechanism that favours individuals who maximize group coordination. However, this theory faces important issues: 1) citizens often rely on traits, such as attractiveness, that are unrelated to coordination skills; 2) untrustworthy leaders who can -and do- impair group functioning by placing their interests first are sometimes chosen; 3) many societies are leader-free and group coordination can be sustained without leaders. Based on this research, we argue that individuals do not choose leaders based on who is most likely to maximise group coordination but rather on who displays the best phenotype given the environment.

Main text

1. The standard account: leaders as group coordinators

Among all the choices individuals have to make in modern democracies, electing political leaders has far reaching consequences. Investigating these choices is therefore of primary interest to understand and predict the evolution of democratic societies. In the past decades, studies in social sciences have highlighted the importance of cultural, familial and societal factors in electoral behaviours [1–5]. More recently, cognitive scientists replicated these findings in experiments investigating the influence of personality traits and worldviews, such as the sensitivity to diseases and perceived external threats, on voting behaviour [6–9]. In addition, basic social cues, such as the candidate’s voice or facial features have been identified as important predictors of voting behaviours [10–13].

Based on this research, evolutionary social psychologists have put forward the idea that humans are equipped with an evolved cognitive mechanism dedicated to leader choice, the function of which would be to choose leaders who are good at coordinating the group [14,15]. Put simply, the evolutionary rationale is that having a good leader provides important survival benefits to the individual by increasing coordination efficiency. As a result, a leader choice mechanism is likely to have been selected as part of a set of psychological adaptations for leadership and followership. The idea of a specific adaptation for leader choice is compelling and fits well with experimental findings demonstrating that choosing a leader differs from choosing other social partners, both with respect to the qualities that are

valued and to the kind of environmental signals that modulate these preferences. For instance, in violent situations, dominant leaders, who can coerce others and thus ensure effective group coordination, are indeed preferred [16–18]. There is however, no such dominance premium for other social decisions, and dominance can even be avoided in cooperative contexts [16,19]. These results thus appear consistent with the idea that leader decisions are guided by a specific cognitive mechanism geared to select leaders with good coordination skills. However, this theory also faces a number of challenges.

2. Challenges to the standard account

The standard evolutionary theory of leader choice posits the existence of an adaptive leader selection mechanism geared to select the individual most likely to maximize group coordination [14,20]. As we have just seen, this prediction appears empirically grounded [16,21,22] but it also faces three important challenges: i) leader selection is influenced by traits that bear no relevance to leadership skills (e.g., perceived health and attractiveness) [23,24], ii) people sometimes prefer leaders who impede group success [18,25,26], and iii) successful group coordination can emerge in the absence of leaders [27,28].

With respect to the first point, cues of perceived health play a larger role in predicting people's votes than perceived leadership abilities [29]. In line with this result, attractiveness, which is a reliable proxy of perceived health [30], is one of the important cues guiding leader selection [23,31,32,32].

Praino et al. even demonstrated that candidates who are more attractive than their opponents can be advantaged by up to 10 points in the election outcome [31]. Yet, attractiveness is not associated with behaviours that are relevant to enhance group coordination. Specifically, attractiveness does not correlate with trustworthiness or reasoning skills [33–35]. If the true function of the leader choice mechanism really is to choose the best leader to coordinate the group, such decisions based on coordination-irrelevant traits are difficult to account for.

Even more surprisingly, people sometimes choose leaders who clearly display traits, such as untrustworthiness, that are likely to impair group coordination [18,26]. Leaders are indeed in a position where they can exert considerable power to favour their own interests over those of the group [36] so one would expect trustworthiness to be universally valued among leaders [20,36,37]. Yet, both experimental and anthropological data report that untrustworthy leaders can be preferred in a range of contexts. For instance, less trustworthy and even untrustworthy leaders are preferred in times of war [18], even after controlling for other relevant traits like dominance [26]. Highlanders of New Guinea, a society characterized by a high degree of violence [38], also value aggressive, pushy and obviously selfish leaders [39]. Similarly, in situations of crisis and uncertainty, narcissistic leaders are sometimes chosen, although they hinder group effectiveness and group coordination [25,40].

Based on these results, one can wonder why the output of an evolved mechanism adapted to choose the best coordinator would so strongly

deviate from its intended goal and lead individuals to elect leaders who impair group success. Importantly, these deviations do not happen at random; rather, they seem to constitute a systematic response to threatening environments (e.g., war scenarios, high levels of violence and times of crisis). In other words, it appears that the cognitive process that guides leader choice systematically responds to environmental threats by producing leader preferences that do not favour the interests of the group, i.e., a preference for untrustworthy leaders. This pattern of deviations raises further doubts about the existence of an evolved cognitive mechanism dedicated to leader choice.

Finally, the standard theory predicts that having a leader is a universal and necessary feature of human social organization. However, leaders are totally absent in many hunter gatherer societies [28,41–43] (Figure 1). The !Kung Bushmen in Namibia or the Hadza in Tanzania, for instance, have no formal authority and decisions are generally taken by consensus or individual bargaining [28,43,44]. Anthropological research confirms that choosing a leader is a relatively recent phenomenon that appeared with post-Neolithic societies, when humans started to build large-scale institutions [28]. Computational models and experimental work also reveal that coordination of small to medium sized groups is resilient and just as effective without leaders [27,37]. For instance, a recent study conducted on the Tsimane in Bolivia reports that leaders do not necessarily increase overall group efficiency to complete a collective obstacle course [37]. More generally, the literature on non-human animals also provides ample evidence that group

coordination is possible without group leaders and can rely solely on self-organization across taxa (see Box 1).

Taken together, these observations challenge the existence of an evolved ability to select leaders based on their coordinating skills. This implies that people rely on other social cognitive abilities that were primarily selected to serve other functions. In the next sections, we provide evidence suggesting that individuals do not choose their leaders based on who they think is most likely to maximise the group's interest (or indeed their own) but rather on who they identify as having the best phenotype given the environment.

3. Leader choices based on assessment of individual competence

Since humans do not have dedicated machinery for the recent task of choosing leaders, they co-opt other cognitive mechanisms. Our theory posits that leader choice actually relies on the evaluation of individuals' ability to thrive in a given context, independently of their leadership or coordination skills. This process leads to choices that are based on an overall assessment of the candidate's competence as an individual, rather than her competence as a group coordinator (Figure 2, key figure). In other words, the candidate who appears as the most likely to succeed in the current context is preferred as a leader, whether or not she can successfully lead and coordinate the group. If this theory is correct, people's leader preferences should incorporate each and every phenotypic feature that plays a role in increasing the candidate's probability of success in a given

environment, including physiological or behavioural attitudes that bear no relevance to leadership (e.g., health, attractiveness, dominance, etc.).

In this framework, choosing a leader does not rely on a dedicated mechanism but rather on the more general ability to evaluate the phenotype of potential social partners. Such evaluations are routinely produced in humans and social judgments about others are automatically formed even when scant information is available [see 45 for a review on this topic]. These evaluations guide a wide range of social choices, such as partner choice or mate choice [46–49]. Of course, evaluations are influenced by the kind of role these prospective partners will play: prosocial and cooperative qualities are prioritized for partner choice [46,50,51]; cues of fertility play a massive role in mate choice [30,52]. Here, we posit that leader choice relies heavily on cues of competence.

When selecting leaders for a given task, individuals indeed place a high premium on leaders who appear competent to complete the task at hand. For example, during rabbit drives, the Washo of California follow hunt leaders that are known for their hunting skills rather than for their leadership skills. Similarly, in the Mae Enga horticulturalists of New Guinea, exceptional warriors are called during wartime to be war leaders [53, but see 27 for more examples]. In very much the same way, studies on football team managers reveal that former players are often chosen as coaches [54,55], despite the fact that former professional players do not necessarily increase the team's success [55]. Empirical studies also demonstrate that people's competence to complete a particular task increases their odds of being chosen to lead a

group on that same task [56,57]. In Little's study [57], for instance, participants preferred cooperatively skilled individuals to lead a cooperative task and physically skilled individuals to lead a physical task. These findings indicate that participants' assessment of who will best perform the task has a stronger influence on leader choice than leadership-relevant traits such as cooperativeness.

Competence is also central for the selection of political leaders. Indeed, among all the traits that predict election outcomes, such as height and attractiveness [10,13,58–64], participants' answer to the question '*how competent is this person?*' is the most reliable predictor of voting decisions [13,61,63]. Answers to this question even predict election outcomes as accurately as voting intentions, forecasting up to 70% of election outcomes [13,65]. This effect is robust and persists after controlling for multiple possible confounds such as the candidate's party's name as well as other traits such as attractiveness and even hypothetical vote [13,61].

While 'competence' has a clear definition in a task-specific context (e.g., competence *for* hunting rabbits), the definition of 'competence' that people rely on in the context of political elections is much more vague and it is produced in the absence of any context [see for instance 13 seminal paper]. To select leaders, people thus seem to rely on a more global evaluation of competence, one that is not circumscribed to a specific task. In line with this idea, asking participants who is the most competent in the absence of any political context predicts actual election outcomes to the same extent as asking participants who they would choose to vote for. This result suggests

that individuals do not use the information about the leadership role candidates will endorse and rather rely on general competence evaluations. Competence evaluations even mediate the link between mock and actual election outcomes, which confirms that individuals' votes are guided by competence evaluations that are much more general than an assessment of leadership skills [13]. In the present framework, we propose that choosing political leaders is a specific case of leader choice. When selecting a leader for a circumscribed task, people choose the most competent individual to succeed in that task. When choosing political leaders, who are responsible for broad decisions pertaining to the group, individuals thus evaluate the candidate's overall competence to succeed in every challenge they face themselves in their environment, including confronting aggressive strangers or fighting pathogens.

To sum up, the hypothesis that people rely on competence to choose their leaders fits the general pattern reported in the literature. We now turn to the specific predictions of our theory. In particular, given that individuals produce a general assessment of the candidate adequacy with the environment, our theory predicts that leader preferences should shift in response to different environmental signals. In the following section, we confront the prediction to existing data, with special attention to cases in which leader preferences are misaligned with the group's interests.

4. The influence of context on leader choice

A crucial prediction of our theory is that a given phenotypic feature should have more or less influence on leader preferences depending on its relevance to produce an individual phenotype that is adapted to the environment. Trustworthiness is a case in point: this phenotypic trait is especially appropriate in highly cooperative environments where trustworthy individuals benefit from being chosen as cooperative partners and exploitation risk is low (ref). In line with this idea, more trustworthy and less masculine leaders are preferred in cooperative contexts such as peacetime experimental scenarios [17,18,22,66]. Incidentally, our account and the standard view predict the same leader preferences in cooperative environments, but for different reasons: the former predicts that people prefer trustworthy leaders because trustworthiness is an advantageous phenotypic trait in cooperative ecologies, the latter predicts that people prefer trustworthy leaders because trustworthiness is indispensable to group coordination [14,36,37]. Therefore, the best test-case for our theory is to turn to contexts, such as harsh environments, in which the qualities that boost individual success differ from the qualities that favour group coordination.

Violent and non-cooperative environments are contexts in which the qualities required for individual success are at odds with those necessary to ensure group functioning. Indeed, dominant and untrustworthy individuals are more likely to exploit group members. However, they are also more likely to succeed in violent interactions and less likely to be exploited by others [67–69]. These individuals can thus be described as having the most

appropriate phenotype for harsh environments, and our framework predicts that should be preferentially chosen as leaders. In line with this prediction, both observational and experimental data indicate that individuals prefer more dominant and untrustworthy leaders in times of social threat. Analysis of historical records and sociological data have indeed revealed systematic variations in the preference for strong leaders with the exposure to environmental threats (i.e., economic and social threats) [70–73]. For instance, economic threats such as inflation and unemployment are important predictors of between-countries variations in the preference for strong leaders [72]. At the individual level, perceived threat to safety and dangerous worldviews is also associated with an increased preference for strong leaders [8,74,75]. Similarly, Safra et al. [26] have shown that childhood scarcity is systematically associated with a preference for more dominant and less trustworthy leaders. This link has also been confirmed experimentally: social threat scenarios systematically induce a preference for more dominant and untrustworthy leaders [16,17,26,57].

Finally, according to the present theory, not only social pressures, such as violence and exploitation, but also physical and biological threats, should impact leader choice. For instance, in high morbidity environments, being healthy is all the more important. Our theory thus predicts that health cues will be favoured in environments of high disease prevalence. This idea fits with evidence showing that both subjective and experimentally induced higher perceived levels of disease prevalence are associated with a preference for more healthy-looking leaders [23,24,30].

To summarise, the hypothesis that leader choice is driven by an assessment of competence provides a good account of actual leader choices and of the flexibility of political decisions across contexts.

Conclusion

In the present paper, we propose an alternative to the standard evolutionary theory on leader choice. Our survey of the literature indeed suggests that, when choosing a leader, people do not rely on an evolved mechanism dedicated to leader choice, but rather select the individual who is most likely to thrive and succeed in the particular ecology they live in. This implies that cognitive mechanisms that initially evolved to serve other functions are recruited to produce leader decisions (see Outstanding questions box). Importantly, this framework sheds new light on political behaviours, notably by making testable predictions regarding the kind of information people take into account to elect their political leaders. This theory grants particular importance to the way citizens perceive their environment, which may point towards new directions to improve the dynamics of political choices (see Box 2). One prediction of this theory, for instance, is that the perception of voters' current environment is more important to predict the outcome of an election than their perception of the needs of their country. This theory is therefore especially informative to understand the link between worldviews and voting behaviour [9,75], and may constitute part of the explanation behind the rise of far right parties across the globe (see Outstanding questions box). Economic crises and

terrorism threats may have biased voters' to perceive their environment as increasingly dangerous and competitive, thereby conferring a political advantage to more dominant and less trustworthy leaders [76,77].

Trends Box

- Individuals sometimes choose leaders that can – and do – impair group functioning
- People select leaders based on their ability to succeed in the present context and not on their ability to coordinate the group
- Leader preferences do not rely on an evolved cognitive mechanism dedicated to leader choice but rather exploit other cognitive mechanisms

Outstanding Questions Box

- What is the exact cognitive mechanism underlying leader choice? One possibility is that leader choice co-opts social information gathering mechanism that have nothing to do with collective action. For instance, 'copy-successful-individuals' has been put forward as an adapted strategy to maximize the accuracy of social information gathering [92–94]. Another possibility is that leader choice co-opts the cognitive mechanisms underlying the social influence dynamics of collective decision-making. Indeed, granting more weight to knowledgeable individuals increases the accuracy of group decision-making [95,96]. Future research thus needs to clarify the interplay between social information gathering mechanisms and collective decision mechanisms in leader choice.
- What is the link between leader evaluations and more systemic political attitudes? How do these respond to environmental variations? To what extent do some environments favour specific leaders as well as more systemic political attitudes such as authoritarianism or liberalism? Can the covariation of leader preferences and other social attitudes (e.g., conformity, xenophobia) be traced back to a unique mechanism that simultaneously shapes both phenomena?
- Is the influence of biological (e.g., genetic) and ecological factors on the development of political attitudes mediated by differences in how people perceive their environment? In particular, can differences

in environment perception explain long-lasting effect of childhood environment on leader preferences? Can it also account of the link between negativity bias and conservatism?

Box 1. Who needs a leader anyway? Insights from primate behaviour

The literature on non-human primate behaviour provides ample evidence that group coordination can be achieved without group leaders and can rely solely on self-organization [78]. For instance, in Tonkean macaques (*Maccaca tonkeana*) all individuals take part in the decision process to choose their foraging sites and their night camps [78,79]. When facing two options, the number of macaques preferring each option is used to decide on the direction of the collective movement: once a given number of notifying individuals have chosen one of the two directions, the whole troop moves towards the direction chosen by the majority (Figure 1). These threshold-based dynamics ensure efficient collective decision-making and has been described in multiple taxa [80] from other non-human primate species, such as white-faced capuchins (*Cebus capucinus*) [81], to honey bees (*Apis mellifera*) [82] and sticklebacks (*Gasterosteus aculeatus*) [83]. Although it may be argued that human beings coordinate their decisions for far more complex activities, these data nonetheless highlight the fact that group leaders are not necessary to achieving coordination.

Box 2. Improving political dynamics

Our theory posits that when asked to choose the best leader for their country, individuals select the candidate who displays the global phenotype that is most closely aligned with the constraints imposed by their local ecology. The way citizens perceive their environment is therefore of paramount importance and it is problematic for the functioning of democracies that citizens' perception is so often at odds with reality [84–87]. The annual Gallup survey, for example, shows that the American public greatly overestimates the level of crime: while the U.S. Department of Justice reports that violent crimes have fallen drastically since the 1990s, most American citizens report, year after year, that there is more crime in the US than a year ago [85,88,89]. Similarly, while the number of people living in extreme poverty has fallen from about 2 billion in 1990 to 0.7 billion in 2015, in a survey conducted in South Africa, the Swedish foundation Gapminder has shown that the majority of people think that this number has actually increased [86]. Overall, despite the increase in life expectancy and the massive decrease in child mortality, violence or disease prevalence, people around the world disagree with the idea that the world is getting better and perceive their environment as more dangerous than it is actually is [90]. According to our theory, these misperceptions may induce a stronger preference for untrustworthy and dominant leaders than what an objective analysis of their country's economic and sociological data would have predicted. Closing the gap between perception and reality by

informing citizens may thus have the potential prevent the rise of dangerous political attitudes.

In addition, voting decisions might also be improved by tuning citizens' competence evaluations to the specific roles political leaders play in society. Indeed, citizens are often poorly informed about the different tasks leaders accomplish, which explains why they may fail to evaluate candidates on the appropriate traits. Our theory predicts that if individuals were aware of the decisions political leaders have to make and of the qualities they have to possess to succeed as decision-makers, they would vote, as expected by the democratic ideal, for candidates they perceive as the most able to guide their country. In line with this idea, more politically knowledgeable voters are less influenced by competence evaluations [91]. Therefore, acting on the two major components involved in leader choices (the context and the competence evaluations) offers a promising way to improve political dynamics.

Acknowledgments

This study was supported by the Institut d'Études Cognitives (ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*). We are also grateful to Dr. Jean-Baptiste André, Dr. Guillaume Dezecache and Dr. Thom Scott-Phillips for their comments that greatly helped to improve the manuscript.

References

- 1 Coffé, H. et al. (2007) Fertile grounds for extreme right-wing parties: Explaining the Vlaams Blok's electoral success. *Electoral Studies* 26, 142–155
- 2 Golder, M. (2003) Explaining Variation In The Success Of Extreme Right Parties In Western Europe. *Comparative Political Studies* 36, 432–466
- 3 Hatemi, P.K. and McDermott, R. (2012) The genetics of politics: discovery, challenges, and progress. *Trends in Genetics* 28, 525–533
- 4 Knigge, P. (1998) The ecological correlates of right-wing extremism in Western Europe. *European Journal of Political Research* 34, 249–279
- 5 Rydgren, J. (2007) The Sociology of the Radical Right. *Annual Review of Sociology* 33, 241–262
- 6 Dallago, F. and Roccato, M. (2009) Right-wing authoritarianism, Big Five and perceived threat to safety. *European Journal of Personality* DOI: 10.1002/per.745
- 7 Hodson, G. and Costello, K. (2007) Interpersonal disgust, ideological orientations, and dehumanization as predictors of intergroup attitudes. *Psychol Sci* 18, 691–698

- 8 Onraet, E. et al. (2014) The Relationships Between Internal and External Threats and Right-Wing Attitudes: A Three-Wave Longitudinal Study. *Personality and Social Psychology Bulletin* 40, 712–725
- 9 Perry, R. et al. (2013) Dangerous and competitive worldviews: A meta-analysis of their associations with Social Dominance Orientation and Right-Wing Authoritarianism. *Journal of Research in Personality* 47, 116–127
- 10 Chen, F.F. et al. (2016) Culture Matters The Looks of a Leader Are Not All the Same. *Social Psychological and Personality Science* 7, 570–578
- 11 Olivola, C.Y. and Todorov, A. (2010) Elected in 100 milliseconds: Appearance-Based Trait Inferences and Voting. *J Nonverbal Behav* 34, 83–110
- 12 Tigue, C.C. et al. (2012) Voice pitch influences voting behavior. *Evolution and Human Behavior* 33, 210–216
- 13 Todorov, A. et al. (2005) Inferences of Competence from Faces Predict Election Outcomes. *Science* 308, 1623–1626
- 14 Van Vugt, M. and Grabo, A.E. (2015) The Many Faces of Leadership. *Current Directions in Psychological Science* 24, 484–489
- 15 Van Vugt, M. and Kurzban, R.K. (2007) Cognitive and social adaptations for leadership and followership: Evolutionary game theory and group dynamics. In *Evolution and the Social Mind: Evolutionary Psychology and Social Cognition* (Forgas, J. P. et al., eds), pp. 229–244, Psychology Press Ltd

- 16 Laustsen, L. and Petersen, M.B. (2015) Does a competent leader make a good friend? Conflict, ideology and the psychologies of friendship and followership. *Evolution and Human Behavior* 36, 286–293
- 17 Little, A.C. et al. (2007) Facial appearance affects voting decisions. *Evolution and Human Behavior* 28, 18–27
- 18 Little, A.C. et al. (2012) The perception of attractiveness and trustworthiness in male faces affects hypothetical voting decisions differently in wartime and peacetime scenarios. *The Quarterly Journal of Experimental Psychology* 65, 2018–2032
- 19 Safra, L. et al. (submitted) Face evaluations confirm that social motivation serves cooperative goals. *Nature Human Behaviour*
- 20 Price, M.E. and Van Vugt, M. (2014) The evolution of leader–follower reciprocity: the theory of service-for-prestige. *Front. Hum. Neurosci.* 8, 363
- 21 Laustsen, L. and Petersen, M.B. (2017) Perceived Conflict and Leader Dominance: Individual and Contextual Factors Behind Preferences for Dominant Leaders. *Political Psychology* DOI: 10.1111/pops.12403
- 22 Re, D.E. et al. (2013) Facial Cues to Perceived Height Influence Leadership Choices in Simulated War and Peace Contexts. *Evol Psychol* 11, 89–103
- 23 White, A.E. et al. (2013) Beauty at the Ballot Box Disease Threats Predict Preferences for Physically Attractive Leaders. *Psychological Science* 24, 2429–2436.
- 24 Zebrowitz, L.A. et al. (2015) Ailing Voters Advance Attractive Congressional Candidates. *Evolutionary Psychology* 13, 16–28

- 25 Nevicka, B. et al. (2013) Uncertainty enhances the preference for narcissistic leaders. *Eur. J. Soc. Psychol.* 43, 370–380
- 26 Safra, L. et al. (2017) Childhood harshness predicts long-lasting leader preferences. *Evolution and Human Behavior* DOI: <https://osf.io/v7k73/>
- 27 Hooper, P.L. et al. (2010) A theory of leadership in human cooperative groups. *Journal of Theoretical Biology* 265, 633–646
- 28 Kaplan, H.S. et al. (2009) The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364, 3289–3299
- 29 Kramer, R.S.S. et al. (2010) Perceived health from biological motion predicts voting behaviour. *The Quarterly Journal of Experimental Psychology* 63, 625–632
- 30 Thornhill, R. and Gangestad, S.W. (1999) Facial attractiveness. *Trends in Cognitive Sciences* 3, 452–460
- 31 Praino, R. et al. (2014) Looking Good or Looking Competent? Physical Appearance and Electoral Success in the 2008 Congressional Elections. *American Politics Research* 42, 1096–1117
- 32 Stockemer, D. and Praino, R. (2017) Physical attractiveness, voter heuristics and electoral systems: The role of candidate attractiveness under different institutional designs. *The British Journal of Politics and International Relations* 19, 336–352

- 33 Mitchem, D.G. et al. (2015) No relationship between intelligence and facial attractiveness in a large, genetically informative sample. *Evol Hum Behav* 36, 240–247
- 34 Talamas, S.N. et al. (2016) Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance. *PLOS ONE* 11, e0148284
- 35 Wilson, R.K. and Eckel, C.C. (2006) Judging a Book by its Cover: Beauty and Expectations in the Trust Game. *Political Research Quarterly* 59, 189–202
- 36 Van Vugt, M. (2006) Evolutionary Origins of Leadership and Followership. *Pers Soc Psychol Rev* 10, 354–371
- 37 von Rueden, C. et al. (2014) Leadership in an Egalitarian Society. *Hum Nat* 25, 538–566
- 38 Berndt, R.M. (1964) Warfare in the New Guinea Highlands. *American Anthropologist* 66, 183–203
- 39 Lewis, H.S. (1974) *Leaders and followers: Some anthropological perspectives*, Addison-Wesley Reading, MA:
- 40 Nevicka, B. et al. (2011) Reality at odds with perceptions: narcissistic leaders and group performance. *Psychological Science* 22, 1259–1264
- 41 Boehm, C. (1993) Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology* 34, 227–254
- 42 Fried, M.H. (1967) *The evolution of political society: An essay in political anthropology*, Random House.
- 43 Woodburn, J. (1982) Egalitarian Societies. *Man* 17, 431–451

- 44 Widerquist, K. and McCall, G. (2017) Prehistoric Myths in Modern Political Philosophy.
- 45 Todorov, A. et al. (2015) Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annu. Rev. Psychol.* 66, 519–545
- 46 Bonnefon, J.-F. et al. (2013) The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General* 142, 143–150
- 47 Rezlescu, C. et al. (2012) Unfakeable Facial Configurations Affect Strategic Choices in Trust Games with or without Information about Past Behavior. *PLOS ONE* 7, e34293
- 48 Rhodes, G. (2006) The evolutionary psychology of facial beauty. *Annual review of psychology* 57, 199–226
- 49 Yu, M. et al. (2014) Developing trust: First impressions and experience. *Journal of Economic Psychology* 43, 16–29
- 50 Chang, L.J. et al. (2010) Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology* 61, 87–105
- 51 van 't Wout, M. and Sanfey, A.G. (2008) Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition* 108, 796–803
- 52 Millera, G.F. and Todd, P.M. (1998) Mate choice turns cognitive. *Trends in Cognitive Sciences* 2, 190–198
- 53 Meggitt, M.J. (1977) *Blood is their argument: Warfare among the Mae Enga tribesmen of the New Guinea highlands*, McGraw-Hill Humanities Social.

- 54 Crust, L. and Lawrence, I. (2006) A review of leadership in sport: implications for football management. *Athletic Insight Journal* 8,
- 55 Muehlheusser, G. et al. (2016) The Contribution of Managers to Organizational Success. *Journal of Sports Economics* DOI: 10.1177/1527002516674760
- 56 Bor, A. (2017) Spontaneous categorization along competence in partner and leader evaluations. *Evolution and Human Behavior* in press,
- 57 Little, A.C. (2014) Facial appearance and leader choice in different contexts: Evidence for task contingent selection based on implicit and learned face-behaviour/face-ability associations. *The Leadership Quarterly* 25, 865–874
- 58 Blaker, N.M. et al. (2013) The height leadership advantage in men and women: Testing evolutionary psychology predictions about the perceptions of tall leaders. *Group Processes Intergroup Relations* 16, 17–27
- 59 Castelli, L. et al. (2009) The effects of perceived competence and sociability on electoral outcomes. *Journal of Experimental Social Psychology* 45, 1152–1155
- 60 Chen, F.F. et al. (2014) The looks of a leader: Competent and trustworthy, but not dominant. *Journal of Experimental Social Psychology* 51, 27–33
- 61 Laustsen, L. (2013) Decomposing the Relationship Between Candidates' Facial Appearance and Electoral Success. *Polit Behav* 36, 777–791

- 62 Poutvaara, P. et al. (2009) Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology* 45, 1132–1135
- 63 Rule, N.O. et al. (2010) Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology* 98, 1–15
- 64 Sussman, A.B. et al. (2013) Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology* 49, 771–775
- 65 Ballew, C.C. and Todorov, A. (2007) Predicting political elections from rapid and unreflective face judgments. *PNAS* 104, 17948–17953
- 66 Spisak, B.R. et al. (2012) Warriors and Peacekeepers: Testing a Biosocial Implicit Leadership Hypothesis of Intergroup Relations Using Masculine and Feminine Faces. *PLOS ONE* 7, e30399
- 67 Bos, P.A. et al. (2010) Testosterone decreases trust in socially naïve humans. *PNAS* 107, 9991–9995
- 68 Carré, J.M. et al. (2009) Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science* 20, 1194–1198
- 69 Haselhuhn, M.P. and Wong, E.M. (2011) Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society of London B: Biological Sciences* DOI: 10.1098/rspb.2011.1193
- 70 Doty, R.M. et al. (1991) Threat and authoritarianism in the United States, 1978–1987. *Journal of Personality and Social Psychology* 61, 629–640

- 71 McCann, S.J.H. and Stewin, L.L. (1990) Good and Bad Years: An Index of American Social, Economic, and Political Threat (1920–1986). *The Journal of Psychology* 124, 601–617
- 72 Onraet, E. et al. (2013) Threat and Right-Wing Attitudes: A Cross-National Approach. *Political Psychology* 34, 791–803
- 73 Sales, S.M. (1973) Threat as a factor in authoritarianism: An analysis of archival data. *Journal of Personality and Social Psychology* 28, 44–57
- 74 Mirisola, A. et al. (2014) Societal Threat to Safety, Compensatory Control, and Right-Wing Authoritarianism. *Political Psychology* 35, 795–812
- 75 Sibley, C.G. et al. (2007) Effects of Dangerous and Competitive Worldviews on Right-Wing Authoritarianism and Social Dominance Orientation over a Five-Month Period. *Political Psychology* 28, 357–371
- 76 Kakkar, H. and Sivanathan, N. (2017) When the appeal of a dominant leader is greater than a prestige leader. *PNAS* DOI: 10.1073/pnas.1617711114
- 77 Landau, M.J. (2004) Deliver us from Evil: The Effects of Mortality Salience and Reminders of 9/11 on Support for President George W. Bush. *Personality and Social Psychology Bulletin* 30, 1136–1150
- 78 King, A.J. and Sueur, C. (2011) Where Next? Group Coordination and Collective Decision Making by Primates. *Int J Primatol* 32, 1245–1267
- 79 Sueur, C. et al. (2010) Sequence of quorums during collective decision making in macaques. *Behavioral Ecology and Sociobiology* 64, 1875–1885

- 80 Sumpter, D.J.T. and Pratt, S.C. (2009) Quorum responses and consensus decision making. *Philos Trans R Soc Lond B Biol Sci* 364, 743–753
- 81 Petit, O. et al. (2009) Collective decision-making in white-faced capuchin monkeys. *Proc Biol Sci* 276, 3495–3503
- 82 Seeley, T.D. and Visscher, P.K. (2004) Quorum sensing during nest-site selection by honeybee swarms. *Behav Ecol Sociobiol* 56, 594–601
- 83 Ward, A.J.W. et al. (2008) Quorum decision-making facilitates information transfer in fish shoals. *PNAS* 105, 6948–6953
- 84 Chambers, J.R. et al. (2014) Better Off Than We Know. *Psychol Sci* 25, 613–618
- 85 Gramlich, J. Voters' perceptions of crime continue to conflict with reality. , *Pew Research Center*. (2016)
- 86 Gapminder Ignorance Survey in South Africa 2015. . (2015)
- 87 IPSOS Mori (2015) , Perils of Perception. , *Ipsos MORI*. [Online]. Available: https://www.ipsos.com/ipsos-mori/en-uk/perils-perception-2015?language_content_entity=en-uk. [Accessed: 03-Jul-2017]
- 88 Gallup (2016) , Crime. , *Gallup.com*. [Online]. Available: <http://www.gallup.com/poll/1603/Crime.aspx>. [Accessed: 19-Jul-2017]
- 89 U.S. Department of Justice (2014) , Uniform Crime Reporting. , *FBI*. [Online]. Available: <https://ucr.fbi.gov/ucr>. [Accessed: 19-Jul-2017]
- 90 Roser, M. and Nadgy, M. Optimism & Pessimism. , *Our World In Data*. (2016)

- 91 Lenz, G.S. and Lawson, C. (2011) Looking the Part: Television Leads Less Informed Citizens to Vote Based on Candidates' Appearance. *American Journal of Political Science* 55, 574–589
- 92 Boyd, R. and Richerson, P.J. (1988) *Culture and the Evolutionary Process*, University of Chicago Press.
- 93 Kendal, R.L. et al. (2005) Trade-Offs in the Adaptive Use of Social and Asocial Learning. 35 (Behavior, B.-A. in the S. of, ed), pp. 333–379, Academic Press
- 94 Laland, K.N. (2004) Social learning strategies. *Learn Behav* 32, 4–14
- 95 Hertz, U. et al. (2016) Social influence protects collective decision making from equality bias. *Journal of Experimental Psychology: Human Perception and Performance* 42, 164–172
- 96 Moussaïd, M. et al. (2013) Social Influence and the Collective Dynamics of Opinion Formation. *PLOS ONE* 8, e78433
- 97 Murdock, G.P. and White, D.R. (1969) Standard Cross-Cultural Sample. *Ethnology* 8, 329–369

Figure captions

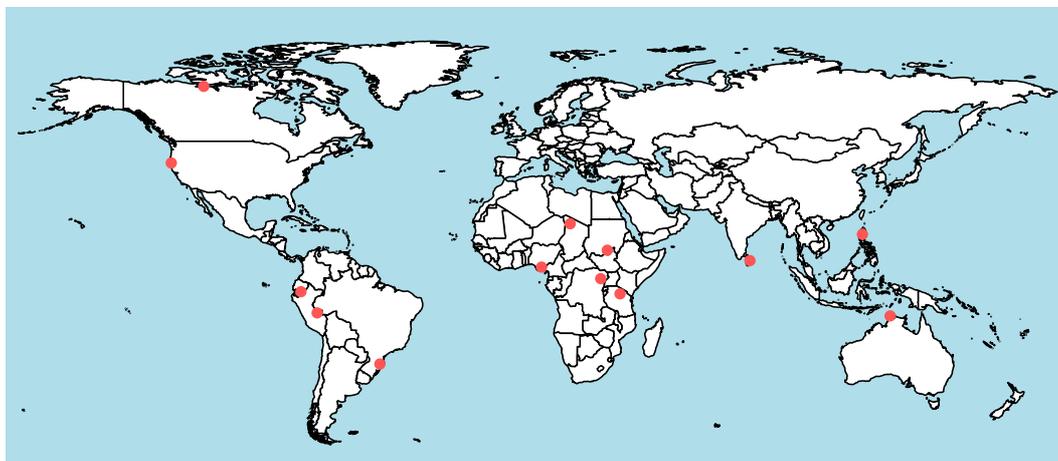


Figure 1. Geographical localisation of 13 societies around the world with no centralized leadership (data from the Standard Cross-Cultural Sample [97])

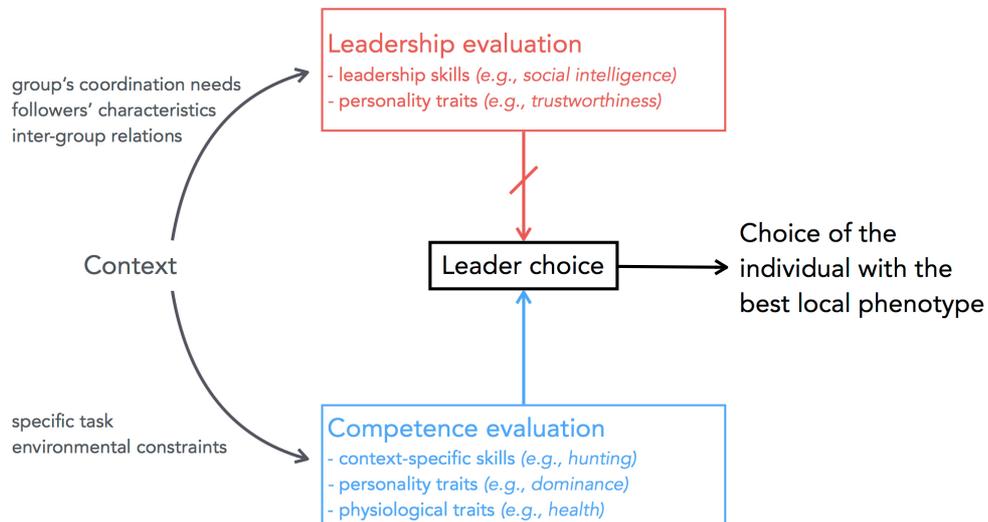


Figure 2 (Key figure). Leader choice mechanism based on individual competence evaluation.

Competence evaluations guide leader preferences and take into account all phenotypic features, from context-specific skills to physiological traits. These evaluations are influenced by the specific context the individual is in (either the specific task her group has to complete or the general environment she experiences). On the contrary, leader preferences are not influenced by leadership evaluations or by the contextual features that impact these evaluations, such as the group's coordination needs. As a result, individuals select the most individually competent candidates as leaders.

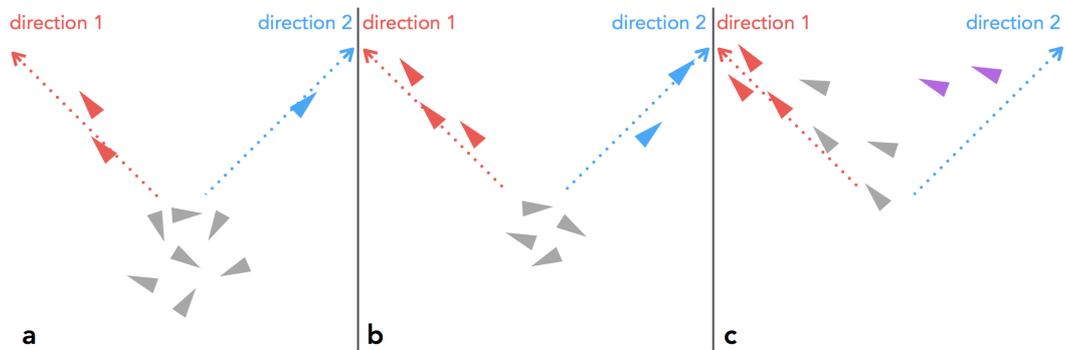


Figure 1. Illustration of the decision-making process in Tonkean macaques (*Macaca tonkeana*) between two directions.

a) one sub-group of macaques proposes a direction (in red) while another individual proposes an alternative direction (direction 2; in blue); b) additional notifying individuals (in color) join the movement in each of the two directions, with a majority moving towards direction 1. c) once the threshold number of notifying individuals is reached, the movement starts in the direction chosen by the majority of notifying individuals (direction 1); the individuals previously notifying direction 2 join the movement in direction 1 (in purple) and then maintain the cohesion of the group (adapted from Sueur, Deneubourg & Petit [79]).

Political attitudes as behavioral responses to the environment

In this chapter, we demonstrated that analyzing political behavior as responses to environmental signals, i.e., adopting an ecological perspective on political behavior, can provide new insights into the cognitive mechanisms underlying these social attitudes in three ways. First, by revealing that early environmental harshness is associated with authoritarian attitudes in adulthood, our experimental studies reveal that the underlying cognitive process integrates information about the environment throughout the individual's lifespan and thus shed a new light on the ontology of political attitudes. In particular, this finding opens the path to the investigation of long-term effects of environmental variables on political attitudes at the individual level and at the country level. Indeed, it suggests the existence of delayed effects of environmental factors, such as economic crises and social threats and taking past events into account, in addition to the current situation, may help to better understand, and predict, the evolution of modern democracies. Second, by showing that environmental harshness is also associated with a preference for strong leaders in children, our study suggests that the cognitive mechanisms underlying political preferences respond to environmental harshness in the same way in politically knowledgeable and not-knowledgeable individuals. Thereby, we provide new evidence in favor of the idea that political preferences may, at least partially, be construed as deeply rooted responses to the environment.

Finally, the ecological perspective on human behavior led us to analyze variations in political choices as differences in the weight granted to various phenotypic features (e.g., strength, prosociality or health) for choosing a leader. Although this approach may appear as highly reductionist, it allowed investigating leader preferences in a uniform way across populations (e.g., in children and in adults) and provided new insights into the way in which leader preferences vary across contexts. In particular, by revealing that in specific cases individuals prefer more untrustworthy leaders, it strongly challenged the mainstream evolutionary theory on leader preferences (King et al., 2009; Van Vugt, 2006) and provided the ground for a new analysis of the mechanisms guiding leader choice and their evolutionary roots. In summary, in this chapter, we exemplified how the study of individual differences in social preferences as responses to environmental factors and as variations in the importance granted to cooperation- and power-related signals, can help to further understand the cognitive mechanisms involved in wide-ranging social attitudes such as political ones.

Chapter four. General discussion

An original approach to study social behavior

Throughout this thesis, we used individual differences in social decisions made from faces to investigate the ecological value and evolutionary roots of two major elements of human beings' social life: social motivation and leader choices. More specifically, our approach relied on two major elements.

First, based on the functional model developed by Oosterhof and Todorov (2008), we conceptualized individual variations in face evaluations as differences in the weighting of cooperation-related and power-related signals. More precisely, we construed individual differences in specific social

decisions (i.e., approachability evaluations and leader choice) as differences in the importance of partner's cooperative vs. power-related motives. In addition to providing a controlled measure of individual differences in social decisions, the developed methodology creates a functional framework to the understanding of these differences. For instance, applying this methodology showed that the preference for more masculine leaders in times of war (Little, Burriss, Jones, & Roberts, 2007b; Spisak, Homan, et al., 2012) is not only driven by a preference for more dominant leaders but also for more untrustworthy ones (Safra et al., 2017). Therefore, it revealed a sub-optimal leader preference in wartime, which contrasts with the acknowledged existence of an evolved cognitive mechanism dedicated to leader choice (King et al., 2009). As demonstrated by this example, and throughout this thesis, the application of the power-cooperation approach to face evaluations can provide new insights into the evolutionary bases of social decisions.

Second, adopting an ecological perspective on human behavior, we construed individual differences in social behavior as responses to environmental constraints. This second element of our approach was applied in two ways. First of all, our investigation of individual differences aimed at understanding individual characteristics, such as personality traits, that influence social behavior in an evolutionary perspective. By adopting an evolutionary perspective on human behavior, these internal factors can be construed as evolutionary selected responses to specific environmental

pressures (Nettle, 2005). Analysis of the variability in social decisions associated with specific individual characteristics, and in particular in the importance granted to different social signals, can thus help us to better understand their evolutionary roots by shedding light on the environmental pressures that may have shaped their development. In the second chapter of this thesis, we showed that social motivation is associated, throughout the lifespan, with an increased weight granted to cooperation-related signals for approach-avoidance decisions rather than with a general tendency to approach more partners indiscriminately. Thereby, using the cooperation-power framework, we confirmed one of the central predictions of the evolutionary model of social motivation stating that social motivation can be construed as a strategy suited to highly cooperative environments (Chevallier, Kohls, et al., 2012).

The second way we analyzed individual differences was directly inspired by the ecological literature on animal behavior. In other words, we investigated behavioral response to signals coming from individuals' environment. Based on a growing literature showing the importance of signals perceived during development on juveniles and adults' behavior both in non-human animals and in humans (Adair, 2001; Bateson et al., 2004; Nettle, Coall, et al., 2011; Sloboda, Hart, Doherty, Pennell, & Hickey, 2007), we reveal an association between childhood deprivation and leader preferences. Our results shed light on the effect of this yet under-investigated factor on the well-studied political attitudes. Thereby, they provide new insights into the ontology of

these social preferences and new evidence in favor of the possibility to, at least partially, construe variations in political attitudes as biological responses to environmental signals.

To summarize, this thesis provides further evidence of the effect of both individual characteristics and external variables on social decisions made from faces. In addition, and more interestingly, it demonstrates how investigating individual differences in the weight granted to cooperation- and power-related signals for making social decisions can provide new insights into the cognitive processes underlying social behavior and on their evolutionary roots. The work presented in this thesis can thus be seen as an example of the potential applications of this more general approach to study social behavior in an ecological and evolutionary perspective.

Beyond faces and social interactions

The use of face evaluations in the present studies was a methodologically useful tool to investigate the importance of partners' cooperation- and power-related features during social decisions. Indeed, face evaluations of neutral faces provided us a controlled setting to investigate individual differences in social decisions. However, the cooperation-power dichotomy has been theorized to be more widely applicable to social behavior (Fiske et al., 2007) and the present methodology should thus extend to other types of social signals.

A first direct candidate for such an extension is voice. Indeed, as they do for faces, individuals form first impressions from others' voices to make social decisions, such as leader choice and trustworthiness evaluations (e.g., Klofstad, Anderson, & Peters, 2012; Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). Importantly, it has been recently shown that the same dimensions of dominance and trustworthiness accounted for evaluations made from voices (McAleer et al., 2014). Individual differences in voice evaluations, and more generally in social decisions made from voices, could therefore be used to highlight differences in the weighting of cooperation- and power-related cues in social decisions and, thus, the cognitive processes that underlie social behavior.

The cooperation-power framework is not restricted to social cues like voices or faces that are used for inferring others' traits, but also extends to signals of others' states. Indeed, cooperation- and power-related behavior can vary within an individual in different contexts. For instance, Güroğlu et al. (2014) have shown that from 9 years old, individuals' cooperation tendencies are sensitive to their partner's status: children, adolescents and young adults trust more a friend than an anonymous peer. Specific social cues have been shown to signal these different states. More precisely, emotions such as joy and fear have been tagged as affiliative signals, indicating a more cooperative state (Hammer & Marsh, 2015; Niedenthal, Mermillod, Maringer, & Hess, 2010). On the opposite dimension, other emotions like

anger have been linked with power and dominance. Indeed, facial displays of anger have been shown to mimic facial features of dominance (Sell, Cosmides, & Tooby, 2014). Similarly, Hareli et al. (2009) have revealed that faces expressing anger were perceived as more dominant than those showing displays of fear or of sadness. The links between joy and cooperation, as well as between anger and dominance have been further demonstrated using computerized faces. Indeed, neutral avatars representing high levels of trustworthiness are perceived as happier and neutral avatars representing high levels of dominance are perceived as angrier (Oosterhof & Todorov, 2008; Said, Haxby, & Todorov, 2011).

These transient emotional displays affect cooperative interactions (Centorrino, Djemai, Hopfensitz, Milinski, & Seabright, 2015b; Hareli et al., 2009; Johnston, Miles, & Macrae, 2010; Krumhuber et al., 2007; Campellone & Kring, 2013). For instance, Krumhuber et al. (2007) have shown that a partner's dynamic smile expressions of genuine smiles significantly increased participants' tendency to engage in a trust game compared to expressions of fake smiles. In summary, social decisions also rely on transient emotional displays carrying cooperation- and power-related cues. The approach developed in this thesis could thus be applied to conceptualize individual differences in the sensitivity to different types of emotions as differences in the sensitivity to cooperation- and power-related social signals.

Similarly, the cooperation-power framework should also extend beyond perceived cues and apply to more objective cues of others' cooperation- and power-related motives, such as partner's behavior. Indeed, individuals take other's behavior into account to update their beliefs about her personality traits (e.g., Chang et al., 2010; Rezsescu et al., 2012). For instance, it has been shown that participants use their partner's history to adjust their cooperative strategy in a trust game, with cooperation decreasing with exposure to the partner's exploitative behavior (Chang et al., 2010). Similarly, third-party provided information about others' behavior also shapes social decisions. As an example, in an experiment on leader choice, Little (2014) has revealed that statements on others' physical or cooperative behavior such as 'helps children in training for various sports, including boxing' or 'volunteers his time at a care home for the elderly' significantly impact social decisions. More generally, first- or second-hand information about others' actions plays a major role in human social evolution (Mohtashemi & Mui, 2003; Sperber & Baumard, 2012). The present framework could thus be used to investigate individual and contextual differences in social decision by measuring variations in the weight granted to any type of cooperation- and power-related information for social decisions.

Finally, the framework presented in this thesis could also be used to investigate the consequences of cooperation- and power-related information on social behavior more generally. Social behavior does not

only include behavior towards others (social interactions) but also behavior in the presence of others (individual behavior in a social context). Indeed, the presence of others significantly impacts the cognitive processing of non-social information (e.g., Conty, Gimmig, Belletier, George, & Huguet, 2010; Safra, Chevallier, Blakemore, & Palminteri, in prep.; Wang & Apperly, 2017-see Appendix). While the studies presented in this thesis exclusively focused on the impact of cooperation- and power-related cues on minimal cases of social interactions, the developed approach could be extended to the study of the influence of the social context on individual behavior. For instance, it has been shown that social information gathering, an important cognitive process in the human species, is highly sensitive to the information source (Kendal, Coolen, van Bergen, & Laland, 2005; Laland, 2004). Our approach could thus be applied to evaluate the influence of cooperation- and power-related information to orient the information source selection as well as the weighting of socially- and self-provided information. In summary, the approach developed in this thesis, based on the cooperation/power dichotomy and on the investigation of individual differences can provide new insights into central components of social behavior, such as cooperation partner choice and social learning.

In addition, this approach may not only be applied to study how individuals use social information but also how individuals provide social information. Indeed, another fruitful way this framework could be used is to study the type of information individuals advertise to others. More precisely, as the

way individuals choose their social partners is guided by perceived social information, on the other hand, potential partners may choose to advertise specific features in order to meet other's expectations and to be preferentially chosen.

From the observer to the target

Across the animal kingdom, individuals tend to act in order to shape the way they are perceived by advertising specific traits. For instance, during reproduction periods, birds deploy sexual parades that advertise their mate quality (Smith & Harper, 2003). In a wide range of taxa, males engage in repetitive energetically expensive behaviors during mating season, such as leg-waving displays of jumping spiders or vocalization of red deers, that signal their physical vigor and thus genetic value to females (see Byers, Hebets, & Podos, 2010 for more examples). Similarly, human beings act more prosocially in presence of others, thereby advertising their quality as cooperation partners (Sperber & Baumard, 2012; Tennie, Frith, & Frith, 2010). In the same line of idea, Tingley (2014) has shown that participants preferentially select highly trustworthy avatars for representing themselves as potential partners in a trust game. As for the use of social cues by the observer for making decisions, the cooperation-power framework can provide insights into the importance of cooperation- and power-related features from the perspective of the target.

At this point, it is important to note that the efficiency of one's signal is highly dependent on the receiver's motives. For instance, signaling trustworthiness is efficient only to the extent that others grant importance to this feature. Therefore, by revealing the traits individuals want to advertise, the study of self-representation would also inform on the importance others grant to cooperation- and power-related cues to choose their social partners. This methodology can notably provide data on the societies for which no direct access to social values is possible, such as past societies. For instance, it has been theorized that from the late Middle Ages, European societies have known important cultural and societal changes, notably described as the civilizing process (Elias, 1978). However, no direct quantitative measure of such a change in mentalities is accessible, making the investigation of the factors that may have influenced this evolution difficult. The study of self-representation in the cooperation-power framework can fill this gap by providing quantitative, yet indirect, data of this historical change of mentalities.

Such an investigation could notably rely on portraits. Indeed, among the different types of self-representation, portraits have been a major medium of representation from the Middle Ages to the rise of photography in the 20th century. Portraits were notably used during marriage negotiation to present potential spouses as well as by kings to disseminate the image of their power to large audiences (West, 2004). In addition, portraits are particularly interesting as they aim to represent not only what the face really

looks like but also the internal characteristics, such as personality traits, of the sitters (Brilliant, 2013; Freeland, 2007; West, 2004). Features of portraits may thus inform us about the types of traits people wanted to advertise. For instance, individuals who would like to be seen as more cooperative would have an interest in selecting a portrait of themselves emphasizing cues of trustworthiness, such as curved mouth corners (Dotsch & Todorov, 2012). On the contrary those who would like to represent themselves as powerful should emphasize features of dominance such as a large width-to-height facial ratio (Alrajih & Ward, 2014).

The study of portraits may thus give new insights into the mentalities at different time periods and thus new data for the exploration of the variation of social values, notably of cooperative motives, through historical ages. Indeed, the factors driving the changes in the level of trust of a society remain today an open question. While some researchers have hypothesized that institutional changes drive variations in prosociality (Boyd & Richerson, 2009; Henrich, 2004; Richerson & Henrich, 2012), others have put forward that environmental factors, such as resources scarcity, may account for differences in cooperation (Baumard & Chevallier, 2015; Baumard, Hyafil, Morris, & Boyer, 2015; Pepper & Nettle, 2017). By providing quantitative, although indirect, data on the importance of cooperative features at different time points, the study of the evolution of portraits could offer the opportunity to test these hypotheses by measuring the relative influence of economical, social and institutional factors on the development of

cooperation, and thus on the evolution of human societies. However, it is worth noting that as any model for understanding human behavior, the developed framework based on the cooperation-power dichotomy is limited to certain types of social behavior, i.e., those for which these features are relevant, such as cooperation partner choice. On the contrary, it may not account for other types of social choices, such as sexual partner choice, for which other indices may also be crucial.

Limitations of the cooperation-power model

Although the developed approach based on the dichotomy between cooperation and power aims to offer a unified framework to study social decisions, it is worth noting that it may not apply to every type of social decisions. For instance, when the present approach offers new insights on leader and cooperation partner choice, one can wonder to what extent it would apply to another crucial type of social decisions: mate choice. Indeed, mate choice has been shown to be sensitive to social cues, such as signals of warmth and aggression (Fletcher, Tither, O'Loughlin, Friesen, & Overall, 2004; Snyder et al., 2011; Valentine et al., 2014) that can easily be decomposed in terms of cooperation- and power-related traits. However, other types of facial features, such as fluctuating asymmetry (Thornhill et al., 1999; Thornhill & Gangestad, 1994; Waynforth, 1998), influence mate choice and can hardly be placed on the cooperation-power two

dimensional space.

In line with this idea, Sutherland et al. (2013) and more recently Vernon et al. (2014) have shown that a three-dimensional space better accounted for face evaluations. More precisely, in addition to the trustworthiness and dominance dimensions evidenced by Oosterhof and Todorov (2008), a third dimension that corresponds to ratings of youth and attractiveness has also been evidenced by these authors (Sutherland et al., 2013; Vernon et al., 2014). As attractiveness has been shown to be a reliable signal of health (Thornhill et al., 1999), this third dimension may be construed as a health dimension. Adding health as a new dimension for understanding face evaluations may better account for mate choice by allowing to code for mate genetic value, an important determinant of mate across the animal kingdom (Andersson & Simmons, 2006; Holveck & Riebel, 2010; Neff & Pitcher, 2005). This third dimension may also be relevant for analyzing behavior from the perspective of the target in self-representation. Potential mates indeed advertise signals of their genetic value in order to increase their chances of mating (Byers et al., 2010; Smith & Harper, 2003).

In addition, including genetic value, or health, may also help disentangle different effects of multidimensional facial features such as masculinity. Indeed, it has been shown that more masculine individuals are preferred as leaders in times of war but also as mates in times of parasite threat (DeBruine et al., 2010; Little, Burriss, et al., 2007a; Spisak, Dekker, et al., 2012). However, masculinity may not hold the same value for these two

types of choices. Experimental data suggest that the preference for more masculine leaders corresponds to a preference for both more dominant and more untrustworthy leaders (Safra et al., 2017). However, in the case of mate choice, it has been theorized that the preference for masculine individuals is driven by the correlation between the male's fitness value and the presentation of masculine facial features (Folstad & Karter, 1992; Rhodes, Chan, Zebrowitz, & Simmons, 2003; Thornhill & Gangestad, 2006). The central importance of fitness value for mate choice makes this hypothesis more parsimonious for explaining the choice of masculine mates than a preference for less power-oriented or more cooperation-oriented individuals (e.g., Andersson & Simmons, 2006; Holveck & Riebel, 2010; Neff & Pitcher, 2005).

In addition, it is worth noting that social decisions are far more complex and that different factors interact to shape social choices and social perception. For instance, perceived ethnicity modulates the evaluation of faces (e.g., Birkás, Dzhelyova, Lábadi, Bereczkei, & Perrett, 2014; Hareli et al., 2009). In an experiment on 266 participants from four different ethnicities (USA Caucasian, Hungarian Caucasian, East Asian and South Asian), Birkás et al. (2014) have shown that faces' ethnicity biased perception of trustworthiness such that the two Caucasian samples rated Caucasian faces as more trustworthy than the two Asian samples did. Projecting these face evaluations, or the subsequent social decisions, on a two- or three-dimensional space may fail to reveal the underlying cognitive processes. As

a consequence, in order to preserve its explanatory power on the computational processes underlying social behavior, the approach developed in this thesis should only be applied to the study of the social choices for which dominance and trustworthiness hold an important value and in contexts that are not disrupted by other social factors like ethnicity.

In summary, in the present thesis, we developed an original approach based on the cooperation-power dichotomy that we applied to two types of social behavior presenting a high individual variability: approach-avoidance behavior and political attitudes. Going one step further, our investigation of individual weighting differences of cooperation- and power-related cues enlightened the functional consequences of internal and external factors on social decisions and thereby provided new insights into the evolutionary roots of human social behavior by both confirming evolutionary hypotheses on social motivation and improving the understanding of the mechanisms driving leader preferences. Importantly, as we only focused on a few specific social decisions as well as on a few specific individual characteristics, the presented studies should be seen as a demonstration of the explanatory power of the developed approach. The approach developed in the present thesis can indeed be extended to other types of social information signaling cooperation- and power-related tendencies, both from the point of view of the receiver and the emitter, as well as to any type of behavior in social contexts that can be influenced by such signals.

Epilogue. Face evaluations, an evolutionary puzzle

An evolutionary perspective on face evaluations

In this thesis, we adopted an ecological and evolutionary perspective on social decisions from faces, evaluating how individuals perceive others as a response to environmental constraints. However, adopting such an evolutionary view on these social choices naturally leads to the question of the evolutionary roots of the use of face evaluations. From an evolutionary perspective, reliance on a specific signal is adaptive only if the provided information is reliable. Indeed, individuals will incur a fitness cost if they rely on misleading signals (Dawkins, 1984; Smith & Harper, 2003). Applying this reasoning to face evaluations, the reliance of human beings on face evaluations for social decisions and the persistent impact of these

evaluations on social interactions (Chang et al., 2010; Rezlescu et al., 2012; J. P. Wilson & Rule, 2015b) should be linked with a high accuracy of these evaluations. Indeed, as face evaluations strongly influence social behavior, and notably cooperation, which is of major importance for the human species (Kaplan et al., 2009), inaccurate evaluations of others from their faces should be associated with important fitness costs.

In line with this idea, it has been shown that perceived dominance and facial features used to assess dominance correlates with actual physical and social power (Carmen E. Lefevre et al., 2014; G. J. Lewis et al., 2012; Mileva, Cowan, Cobey, Knowles, & Little, 2014). Similarly, some experiments have shown that participants can distinguish cooperators from defectors from their faces (Bonnenfon et al., 2013; De Neys et al., 2013; Tognetti et al., 2013). However, it is worth noting that the accuracy of these trustworthiness judgments is limited. Indeed, some experiments have also failed to find such an association between perceived and actual trustworthiness (Efferson & Vogt, 2013; Rule et al., 2013). In addition, as often pointed out in the literature, when found above chance, the accuracy of perceived trustworthiness is still relatively low (Bonnenfon et al., 2017a; Bonnenfon, Hopfensitz, & De Neys, 2017b; Little, 2017). This limited accuracy of trustworthiness evaluations raises a first evolutionary puzzle. Indeed, while partner choice is construed as a major evolutionary drive in shaping human cooperation (McNamara, Barta, Fromhage, & Houston, 2008; Sperber &

Baumard, 2012; Tennie et al., 2010), one could wonder why these evaluations are so loosely accurate.

Puzzle 1. Face evaluations accuracy

First of all to examine this evolutionary puzzle, it is worth noting that reliance on fixed facial features for making social decisions may not be a selected cognitive process but only the reuse of an adapted mechanism dedicated to another cognitive process. Thereby, evaluations made from fixed facial cues would not have been evolutionary selected and no selection pressure would have acted on their accuracy. For instance, one can hypothesize that the use of fixed facial cues to assess cooperative tendencies is only an extension of the use of emotional displays to evaluate others' cooperative states. Indeed, it has been shown that emotional displays, such as genuine smiles, are used to assess prosocial tendencies throughout the lifespan (Krumhuber et al., 2007; Reed, Zeglen, & Schmidt, 2012; Song, Over, & Carpenter, 2016).

In addition, based on the fact that Duchenne smiles are difficult to fake and actually predict higher rates of cooperation, it has been argued that genuine smiles could be construed as a reliable signal of cooperative intentions (Brown, Palameta, & Moore, 2003; Centorrino, Djemai, Hopfensitz, Milinski, & Seabright, 2015a; Centorrino et al., 2015b). In particular, Centorrino et al. (2015b), found that individuals who were rated as presenting a genuine smile were perceived as more trustworthy but are

also more likely to reciprocate in a trust game. This bulk of evidence suggests that Duchenne smiles may be used as a reliable signal of cooperation. Therefore, one can speculate that perceived trustworthiness from fixed facial features is an extension of the cooperativeness detection mechanism based on emotions, and that faces presenting characteristics similar to Duchenne smiles exploit this cooperativeness detection mechanism and are thus perceived as trustworthy (Oosterhof & Todorov, 2009; Said, Sebe, & Todorov, 2009).

On the other hand, it is also possible that evaluations based on fixed facial features correspond to an evolutionary selected mechanism. In this perspective, Little (2017) has proposed that the limited accuracy of trustworthiness evaluations is linked with the fact that untrustworthiness is a cue emitters have an interest in concealing. Indeed, exploitation is an efficient strategy only if the partner decides to cooperate, i.e., only when it is hidden to the others. Under this view, the evolution of cheater detection mechanism from faces would thus have been constrained by the opposite pressure for cheaters to hide their intentions.

A third way of solving this puzzle is to adopt an error-management perspective (Haselton & Buss, 2000). More precisely, it is possible that the limited accuracy to detect cheaters from their faces is not due to random errors. Individuals may not misclassify potential cooperative partners randomly but they may rather over-classify unknown individuals as potential cheaters thereby minimizing the risk of exploitation. However, to our

knowledge, the existing literature on face evaluations does not allow to disentangle between these different possibilities, leaving this question to future research. Nevertheless, all of these hypotheses raise a second evolutionary puzzle: the existence of untrustworthy faces.

Puzzle 2. Signaling untrustworthiness

Given the fact that being chosen as a cooperative partner has acted as a major evolutionary pressure on human cognition (Sperber & Baumard, 2012; Tennie et al., 2010), being perceived as untrustworthy carries important evolutionary costs. Strikingly, the proportion of faces rated as untrustworthy is not marginal. For instance, in our study on threats evaluations, about half of the presented faces were rated as untrustworthy by most of the participants (Safra, Wyart, Baumard, & Chevallier, submitted). The robustness of trustworthiness ratings across people indicates that specific facial cues are universally used to assess untrustworthiness (Dotsch & Todorov, 2012; Engell et al., 2007; Rule et al., 2013), which implies that some faces are universally perceived as untrustworthy, thus raising an evolutionary puzzle. Indeed, having a face perceived as untrustworthy strongly undermines individuals' success in environments where cooperation is important for resource acquisition, i.e., in the humans' ecological niche (Kaplan et al., 2009). Importantly, faces perceived as untrustworthy are not only perceived as less trustworthy than others, untrustworthy-looking individuals also receive less investment in

cooperative games and are less likely to be chosen as cooperative partners (Bonnenfon et al., 2013; Chang et al., 2010; Rezlescu et al., 2012; van 't Wout & Sanfey, 2008). One can thus wonder how facial features universally perceived as signaling untrustworthiness could have persisted throughout human evolutionary history.

A first possibility to explain the existence of untrustworthy faces is congruent with the hypothesis of an evolved cheater detection mechanism from faces and relies on the idea that untrustworthiness actually signals lower cooperative tendencies. Under this view, honestly signaling ones' untrustworthiness would have some fitness advantage over being perceived as more trustworthy than one really is. In line with this idea, it has been shown that individuals that are perceived as trustworthy are more severely sentenced after a crime than those perceived as untrustworthy (Korva, Porter, O'Connor, Shaw, & Brinke, 2013). A similar result was found by Vullioud et al. (2017) on the impact of commitment on trust. Indeed, these authors have shown that while confident advisors are more likely to be initially trusted, if their advice turns out to be unreliable, they are then less trusted than less confident advisors. Going one step further, advisors initially trusted due to their objective competence and not to their advertised confidence experienced lower reputation drops. In other words, deceiving others entailed higher costs for individuals who initially advertised their advices' trustworthiness. Signaling, at least partially, ones' own

untrustworthiness may thus limit the social cost associated with having exploitative tendencies.

An alternative explanation of the existence of faces rated as universally untrustworthy is that facial features used to assess are biologically constrained. More precisely, these facial cues would be an emergent feature of individuals' developmental trajectory. As an analogy, under this view, facial features used to assess untrustworthiness would develop in a similar way as fluctuating asymmetry, a cue testifying developmental issues due to either genetic or environmental factors, that reduces the chances of being selected as a mate across species (Beasley, Bonisoli-Alquati, & Mousseau, 2013; Little, Apicella, & Marlowe, 2007; Livshits & Kobylansky, 1991; Møller & Thornhill, 1998; Perrett et al., 1999; Thornhill & Gangestad, 2006). In other words, as fluctuating asymmetry may not be explained in terms of fitness benefits, features of untrustworthiness may not be associated with any fitness benefit. According to this view, facial features used to assess untrustworthiness would only correspond to developmentally constrained facial elements that are used by others as an indicator of lower cooperative tendencies, with no direct biological link between these facial features and this behavioral trait.

Importantly, because individuals actually integrate multiple signals for making social decisions such as emotional displays or reputation, it is possible that the disadvantage associated with these facial features may only be marginal. Indeed, although face evaluation still influence

cooperative behavior after exposure to one's behavioral history and after multiple interactions, this effect is relatively small compared to more objective sources of information (Chang et al., 2010; Rezsescu et al., 2012). For instance, Rezsescu et al. (2012) found that, when informed of their partners' cooperation history, participants invested on average 50 more virtual pounds in partners with a reciprocation history than in those with an exploitation history, while they invested 3 additional virtual pounds in partners represented by a trustworthy avatar compared to those represented by an untrustworthy avatar. These results question the actual fitness cost of presenting untrustworthy facial features in ecological settings, and in particular in small-scales societies that are thought to constitute the environment of evolutionary adaption of human beings (Foley, 1995). In addition, Gill et al. (2014) have shown that facial dynamics can modulate perceived trustworthiness from faces. In particular, these authors revealed that a highly untrustworthy face can be perceived as trustworthy when associated with good facial dynamics. In summary, presenting facial features of untrustworthiness may not significantly impair individuals' success in the cooperative environments that constitute human ecological niche, and one can thus hypothesize that no strong evolutionary pressure may have acted on these facial cues. However, at this point, one could thus wonder why would trustworthiness ratings be accurate if they only rely on a disconnected emergent feature.

Two main answers can be given to this question. The first answer is based on social dynamics. Indeed, it has been suggested that individuals complete a self-fulfilling prophecy, being perceived as untrustworthy inducing untrustworthy behavior. A recent experimental study by Slepian & Ames (2016) have provided support to this view, showing that individuals perceived as untrustworthy believe others to expect them to act untrustworthily. Furthermore, in this experiment, the link between perceived and actual trustworthiness was mediated by these beliefs and the target's intentions to meet these expectations.

The second answer is ecological. Indeed, it is possible that the same factors influence both the development of facial features of perceived untrustworthiness and those lowering cooperation. For instance, it has been shown that environmental variables, such as resource scarcity, significantly reduce cooperative tendencies (Grueter et al., 2016; Holland et al., 2012; Nettle, Colléony, et al., 2011). Importantly, not only current but also childhood environmental insecurity influences individuals' cooperation behavior (Hörl, Kesternich, Smith, & Winter, 2016; McCullough, Pedersen, Schroder, Tabak, & Carver, 2013; Viitanen, 2014). Applying the analogy of fluctuating asymmetry, it is possible that developmental insecurity would both influence the emergence of facial features used to assess untrustworthiness and lower cooperation tendencies. Under this hypothesis, accuracy of trustworthiness ratings would thus be explained by common factors inducing both lower cooperation tendencies and the emergence of

facial features of untrustworthiness. However, the current state of the literature does not allow to choose between these different hypotheses and further work, involving biological investigation of the emergence of facial features interpreted as signaling untrustworthiness, is needed in order to better understand the evolutionary dynamics that have led to the existence of untrustworthy faces.

As a conclusion, if, as demonstrated in this thesis, the use of face evaluations can help to enlighten the evolutionary roots of some elements of human social behavior, it is important to note that human's reliance on facial cues also raises several evolutionary puzzles that call for further investigation. In particular, the important reliance of human beings on fixed facial features to make social decisions naturally raises the questions of the accuracy of these evaluations and of the development of facial features of untrustworthiness.

General references

- Adair, L. S. (2001). Size at Birth Predicts Age at Menarche. *Pediatrics*, *107*(4), e59–e59.
- Alrajih, S., & Ward, J. (2014). Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *British Journal of Psychology*, *105*(2), 153–161.
<https://doi.org/10.1111/bjop.12035>
- Andersson, M., & Simmons, L. W. (2006). Sexual selection and mate choice. *Trends in Ecology & Evolution*, *21*(6), 296–302.
<https://doi.org/10.1016/j.tree.2006.03.015>
- Antonakis, J., & Dalgas, O. (2009). Predicting Elections: Child's Play! *Science*, *323*(5918), 1183–1183.
<https://doi.org/10.1126/science.1167748>
- Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948–17953.
<https://doi.org/10.1073/pnas.0705435104>

- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6*(2), 269–278. <https://doi.org/10.1037/1528-3542.6.2.269>
- Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R. A., ... Sultan, S. E. (2004). Developmental plasticity and human health. *Nature, 430*(6998), 419–421. <https://doi.org/10.1038/nature02725>
- Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: the evolution of fairness by partner choice. *The Behavioral and Brain Sciences, 36*(1), 59–78. <https://doi.org/10.1017/S0140525X11002202>
- Baumard, N., & Chevallier, C. (2015). The nature and dynamics of world religions: a life-history approach. *Proc. R. Soc. B, 282*(1818), 20151593. <https://doi.org/10.1098/rspb.2015.1593>
- Baumard, N., Hyafil, A., Morris, I., & Boyer, P. (2015). Increased Affluence Explains the Emergence of Ascetic Wisdoms and Moralizing Religions. *Current Biology, 25*(1), 10–15. <https://doi.org/10.1016/j.cub.2014.10.063>
- Beasley, D. A. E., Bonisoli-Alquati, A., & Mousseau, T. A. (2013). The use of fluctuating asymmetry as a measure of environmentally induced developmental instability: A meta-analysis. *Ecological Indicators, 30*, 218–226. <https://doi.org/10.1016/j.ecolind.2013.02.024>
- Begeer, S., Banerjee, R., Lunenburg, P., Terwogt, M. M., Stegge, H., & Rieffe, C. (2008). Brief Report: Self-Presentation of Children with Autism Spectrum Disorders. *Journal of Autism and Developmental*

- Disorders*, 38(6), 1187–1191. <https://doi.org/10.1007/s10803-007-0503-0>
- Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior*, 28(3), 168–175. <https://doi.org/10.1016/j.evolhumbehav.2006.10.003>
- Birkás, B., Dzhelyova, M., Lábadi, B., Bereczkei, T., & Perrett, D. I. (2014). Cross-cultural perception of trustworthiness: The effect of ethnicity features on evaluation of faces' observed trustworthiness across four samples. *Personality and Individual Differences*, 69, 56–61. <https://doi.org/10.1016/j.paid.2014.05.012>
- Bloxham, L., Bateson, M., Bedford, T., Brilot, B., & Nettle, D. (2014). The memory of hunger: developmental plasticity of dietary selectivity in the European starling, *Sturnus vulgaris*. *Animal Behaviour*, 91, 33–40. <https://doi.org/10.1016/j.anbehav.2014.02.025>
- Boehm, C. (1993). Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology*, 34(3), 227–254. <https://doi.org/10.1086/204166>
- Boksem, M. A. S., Mehta, P. H., Bergh, B. V. den, Son, V. van, Trautmann, S. T., Roelofs, K., ... Sanfey, A. G. (2013). Testosterone Inhibits Trust but Promotes Reciprocity. *Psychological Science*, 0956797613495063. <https://doi.org/10.1177/0956797613495063>
- Bonanno, G. A., & Jost, J. T. (2006). Conservative Shift Among High-Exposure Survivors of the September 11th Terrorist Attacks. *Basic*

- and *Applied Social Psychology*, 28(4), 311–323.
https://doi.org/10.1207/s15324834basp2804_4
- Bonnefon, J.-F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, 142(1), 143–150. <https://doi.org/10.1037/a0028930>
- Bonnefon, J.-F., Hopfensitz, A., & De Neys, W. (2017a). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*. Retrieved from http://www.wdeney.org/data/Bonnefon_CDPS-1.pdf
- Bonnefon, J.-F., Hopfensitz, A., & De Neys, W. (2017b). Trustworthiness perception at zero acquaintance: Consensus, accuracy, and prejudice. *Behavioral and Brain Sciences*, 40. Retrieved from <https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/trustworthiness-perception-at-zero-acquaintance-consensus-accuracy-and-prejudice/A2C8BBD5C33C7B20591964F29037DD17>
- Bonnefon, J.-F., Hopfensitz, A., & Neys, W. D. (2015). Face-ism and kernels of truth in facial inferences. *Trends in Cognitive Sciences*, 19(8), 421–422. <https://doi.org/10.1016/j.tics.2015.05.002>
- Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, 43(4), 703–706. <https://doi.org/10.1016/j.jrp.2009.03.007>

- Bos, P. A., Terburg, D., & Honk, J. van. (2010). Testosterone decreases trust in socially naïve humans. *Proceedings of the National Academy of Sciences*, 107(22), 9991–9995. <https://doi.org/10.1073/pnas.0911700107>
- Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1533), 3281–3288. <https://doi.org/10.1098/rstb.2009.0134>
- Brilliant, R. (2013). *Portraiture*. Reaktion Books.
- Brown, W. M., Palameta, B., & Moore, C. (2003). Are there Nonverbal Cues to Commitment? An Exploratory Study Using the Zero-Acquaintance Video Presentation Paradigm. *Evolutionary Psychology*, 1(1), 147470490300100104. <https://doi.org/10.1177/147470490300100104>
- Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour*, 79(4), 771–778. <https://doi.org/10.1016/j.anbehav.2010.01.009>
- Cacioppo, J. T., & Hawkley, L. C. (2009). Perceived social isolation and cognition. *Trends in Cognitive Sciences*, 13(10), 447–454. <https://doi.org/10.1016/j.tics.2009.06.005>
- Cage, E., Pellicano, E., Shah, P., & Bird, G. (2013). Reputation Management: Evidence for Ability But Reduced Propensity in Autism. *Autism Research*, 6(5), 433–442. <https://doi.org/10.1002/aur.1313>

- Campellone, T. R., & Kring, A. M. (2013). Who do you trust? The impact of facial emotion and behaviour on decision making. *Cognition and Emotion*, 27(4), 603–620. <https://doi.org/10.1080/02699931.2012.726608>
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009a). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198. <https://doi.org/10.1111/j.1467-9280.2009.02423.x>
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009b). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198. <https://doi.org/10.1111/j.1467-9280.2009.02423.x>
- Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015a). A Model of Smiling as a Costly Signal of Cooperation Opportunities. *Adaptive Human Behavior and Physiology*, 1(3), 325–340. <https://doi.org/10.1007/s40750-015-0026-4>
- Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015b). Honest signaling in trust interactions: smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior*, 36(1), 8–16. <https://doi.org/10.1016/j.evolhumbehav.2014.08.001>
- Chamberlain, B., Kasari, C., & Rotheram-Fuller, E. (2006). Involvement or Isolation? The Social Networks of Children with Autism in Regular

- Classrooms. *Journal of Autism and Developmental Disorders*, 37(2), 230–242. <https://doi.org/10.1007/s10803-006-0164-4>
- Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, 61(2), 87–105. <https://doi.org/10.1016/j.cogpsych.2010.03.001>
- Chen, F. F., Jing, Y., & Lee, J. M. (2014). The looks of a leader: Competent and trustworthy, but not dominant. *Journal of Experimental Social Psychology*, 51, 27–33. <https://doi.org/10.1016/j.jesp.2013.10.008>
- Chen, F. F., Jing, Y., Lee, J. M., & Bai, L. (2016). Culture Matters The Looks of a Leader Are Not All the Same. *Social Psychological and Personality Science*, 7(6), 570–578. <https://doi.org/10.1177/1948550616644962>
- Chevallier, C., Grèzes, J., Molesworth, C., Berthoz, S., & Happé, F. (2012). Brief Report: Selective Social Anhedonia in High Functioning Autism. *Journal of Autism and Developmental Disorders*, 42(7), 1504–1509. <https://doi.org/10.1007/s10803-011-1364-0>
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4), 231–239. <https://doi.org/10.1016/j.tics.2012.02.007>
- Chevallier, C., Parish-Morris, J., Tonge, N., Le, L., Miller, J., & Schultz, R. T. (2014). Susceptibility to the audience effect explains performance gap between children with and without autism in a theory of mind

- task. *Journal of Experimental Psychology: General*, 143(3), 972–979.
<https://doi.org/10.1037/a0035483>
- Chevallier, C., Tonge, N., Safra, L., Kahn, D., Kohls, G., Miller, J., & Schultz, R. T. (2016). Measuring Social Motivation Using Signal Detection and Reward Responsiveness. *PLOS ONE*, 11(12), e0167024.
<https://doi.org/10.1371/journal.pone.0167024>
- Conty, L., Gimmig, D., Belletier, C., George, N., & Huguet, P. (2010). The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition*, 115(1), 133–139.
<https://doi.org/10.1016/j.cognition.2009.12.005>
- Cuthbert, B. N. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, 13(1), 28–35.
<https://doi.org/10.1002/wps.20087>
- DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L. M., & Little, A. C. (2010). The health of a nation predicts their mate preferences: cross-cultural variation in women’s preferences for masculinized male faces. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1692), 2405–2410.
<https://doi.org/10.1098/rspb.2009.2184>
- De Neys, W., Hopfensitz, A., & Bonnefon, J.-F. (2013). Low second-to-fourth digit ratio predicts indiscriminate social suspicion, not improved trustworthiness detection. *Biology Letters*, 9(2), 20130037.
<https://doi.org/10.1098/rsbl.2013.0037>

- DeWall, C. N., & Bushman, B. J. (2011). Social Acceptance and Rejection: The Sweet and the Bitter. *Current Directions in Psychological Science*, 20(4), 256–260. <https://doi.org/10.1177/0963721411417545>
- Dezecache, G., Mercier, H., & Scott-Phillips, T. C. (2013). An evolutionary approach to emotional communication. *Journal of Pragmatics*, 59, Part B, 221–233. <https://doi.org/10.1016/j.pragma.2013.06.007>
- Dotsch, R., & Todorov, A. (2012). Reverse Correlating Social Face Perception. *Social Psychological and Personality Science*, 3(5), 562–571. <https://doi.org/10.1177/1948550611430272>
- Doty, R. M., Peterson, B. E., & Winter, D. G. (1991). Threat and authoritarianism in the United States, 1978–1987. *Journal of Personality and Social Psychology*, 61(4), 629–640. <https://doi.org/10.1037/0022-3514.61.4.629>
- Efferson, C., & Vogt, S. (2013). Viewing men's faces does not lead to accurate predictions of trustworthiness. *Scientific Reports*, 3, 1047. <https://doi.org/10.1038/srep01047>
- Ehrlich, I., & Becker, G. S. (1972). Market Insurance, Self-Insurance, and Self-Protection. *Journal of Political Economy*, 80(4), 623–648. <https://doi.org/10.1086/259916>
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does Rejection Hurt? An fMRI Study of Social Exclusion. *Science*, 302(5643), 290–292. <https://doi.org/10.1126/science.1089134>
- Elias, N. (1978). *The Civilizing Process*. Urizen Books Pantheon Books.

- Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental Dimensions of Environmental Risk. *Human Nature, 20*(2), 204–268. <https://doi.org/10.1007/s12110-009-9063-7>
- Ellis, B. J., McFadyen-Ketchum, S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (1999). Quality of early family relationships and individual differences in the timing of pubertal maturation in girls: A longitudinal test of an evolutionary model. *Journal of Personality and Social Psychology, 77*(2), 387–401. <https://doi.org/10.1037/0022-3514.77.2.387>
- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit Trustworthiness Decisions: Automatic Coding of Face Properties in the Human Amygdala. *Journal of Cognitive Neuroscience, 19*(9), 1508–1519. <https://doi.org/10.1162/jocn.2007.19.9.1508>
- Epley, N., & Schroeder, J. (2014). Mistakenly seeking solitude. *Journal of Experimental Psychology: General, 143*(5), 1980–1999. <https://doi.org/10.1037/a0037323>
- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008). (2015). EVS. *GESIS Data Archive, Cologne. ZA4804 Data File Version 3.0.0*. <https://doi.org/doi:10.4232/1.12253>
- Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science, 18*(2), 327–334. <https://doi.org/10.1111/desc.12218>

- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Fletcher, G. J. O., Tither, J. M., O'Loughlin, C., Friesen, M., & Overall, N. (2004). Warm and Homely or Cold and Beautiful? Sex Differences in Trading Off Traits in Mate Selection. *Personality and Social Psychology Bulletin*, 30(6), 659–672. <https://doi.org/10.1177/0146167203262847>
- Foley, R. (1995). The adaptive legacy of human evolution: A search for the environment of evolutionary adaptedness. *Evolutionary Anthropology: Issues, News, and Reviews*, 4(6), 194–203. <https://doi.org/10.1002/evan.1360040603>
- Folstad, I., & Karter, A. J. (1992). Parasites, Bright Males, and the Immunocompetence Handicap. *The American Naturalist*, 139(3), 603–622. <https://doi.org/10.1086/285346>
- Freeland, C. (2007). Portraits in Painting and Photography. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 135(1), 95–109.
- Getov, S., Kanai, R., Bahrami, B., & Rees, G. (2015). Human brain structure predicts individual differences in preconscious evaluation of facial dominance and trustworthiness. *Social Cognitive and Affective Neuroscience*, 10(5), 690–699. <https://doi.org/10.1093/scan/nsu103>
- Gill, D., Garrod, O. G. B., Jack, R. E., & Schyns, P. G. (2014). Facial Movements Strategically Camouflage Involuntary Social Signals of

- Face Morphology. *Psychological Science*, 25(5), 1079–1086.
<https://doi.org/10.1177/0956797614522274>
- Griskevicius, V., Ackerman, J. M., Cantú, S. M., Delton, A. W., Robertson, T. E., Simpson, J. A., ... Tybur, J. M. (2013). When the Economy Falters, Do People Spend or Save? Responses to Resource Scarcity Depend on Childhood Environments. *Psychological Science*, 0956797612451471. <https://doi.org/10.1177/0956797612451471>
- Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Cialdini, R. B., & Kenrick, D. T. (2006). Going along versus going alone: When fundamental motives facilitate strategic (non)conformity. *Journal of Personality and Social Psychology*, 91(2), 281–294. <https://doi.org/10.1037/0022-3514.91.2.281>
- Grueter, C. C., Ingram, J. A., Lewisson, J. W., Bradford, O. R., Taba, M., Coetzee, R. E., & Sherwood, M. A. (2016). Human altruistic tendencies vary with both the costliness of selfless acts and socioeconomic status. *PeerJ*, 4, e2610. <https://doi.org/10.7717/peerj.2610>
- Güroğlu, B., van den Bos, W., & Crone, E. A. (2014). Sharing and giving across adolescence: an experimental study examining the development of prosocial behavior. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00291>
- Hammer, J. L., & Marsh, A. A. (2015). Why do fearful facial expressions elicit behavioral approach? Evidence from a combined approach-

- avoidance implicit association test. *Emotion*, 15(2), 223–231.
<https://doi.org/10.1037/emo0000054>
- Hareli, S., Shomrat, N., & Hess, U. (2009). Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion*, 9(3), 378–384.
<https://doi.org/10.1037/a0015958>
- Harvey, P.-O., Bodnar, M., Sergerie, K., Armony, J., & Lepage, M. (2009). Relation between emotional face memory and social anhedonia in schizophrenia. *Journal of Psychiatry & Neuroscience: JPN*, 34(2), 102–110.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91.
<https://doi.org/10.1037/0022-3514.78.1.81>
- Hayden, B. Y., Parikh, P. C., Deaner, R. O., & Platt, M. L. (2007). Economic principles motivating social attention in humans. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1619), 1751–1756. <https://doi.org/10.1098/rspb.2007.0368>
- Helman, E., M, A., Flake, J. K., & Slepian, M. L. (2017). The Unique Contributions of Perceiver and Target Characteristics in Person Perception. *Journal of Personality and Social Psychology*, No Pagination Specified. <https://doi.org/10.1037/pspa0000090>
- Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior &*

Organization, 53(1), 3–35. [https://doi.org/10.1016/S0167-2681\(03\)00094-5](https://doi.org/10.1016/S0167-2681(03)00094-5)

Holland, J., Silva, A. S., & Mace, R. (2012). Lost Letter Measure of Variation in Altruistic Behaviour in 20 Neighbourhoods. *PLOS ONE*, 7(8), e43294. <https://doi.org/10.1371/journal.pone.0043294>

Holveck, M.-J., & Riebel, K. (2010). Low-quality females prefer low-quality males when choosing a mate. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1678), 153–160. <https://doi.org/10.1098/rspb.2009.1222>

Hörl, M., Kesternich, I., Smith, J. P., & Winter, J. K. (2016). *Early-Life Circumstances Predict Measures of Trust Among Adults: Evidence from Hunger Episodes in Post-War Germany* (SSRN Scholarly Paper No. ID 2866849). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=2866849>

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, 167(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>

Izuma, K., Matsumoto, K., Camerer, C. F., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proceedings of the National Academy of Sciences*, 108(42), 17302–17307. <https://doi.org/10.1073/pnas.1107038108>

- Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of Social and Monetary Rewards in the Human Striatum. *Neuron*, 58(2), 284–294. <https://doi.org/10.1016/j.neuron.2008.03.020>
- Javanbakht, A., King, A. P., Evans, G. W., Swain, J. E., Angstadt, M., Phan, K. L., & Liberzon, I. (2015). Childhood Poverty Predicts Adult Amygdala and Frontal Activity and Connectivity in Response to Emotional Faces. *Frontiers in Behavioral Neuroscience*, 9. <https://doi.org/10.3389/fnbeh.2015.00154>
- Jessen, S., & Grossmann, T. (2016). Neural and Behavioral Evidence for Infants' Sensitivity to the Trustworthiness of Faces. *Journal of Cognitive Neuroscience*, 28(11), 1728–1736. https://doi.org/10.1162/jocn_a_00999
- Johnston, L., Miles, L., & Macrae, C. N. (2010). Why are you smiling at me? Social functions of enjoyment and non-enjoyment smiles. *British Journal of Social Psychology*, 49(1), 107–127. <https://doi.org/10.1348/014466609X412476>
- Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., ... Casey, B. J. (2011). Behavioral and Neural Properties of Social Reinforcement Learning. *The Journal of Neuroscience*, 31(37), 13039–13045. <https://doi.org/10.1523/JNEUROSCI.2972-11.2011>
- Kaplan, H. S., & Gangestad, S. W. (2005). Life History Theory and Evolutionary Psychology. In *The handbook of evolutionary psychology* (pp. 68–95). Hoboken, NJ, US: John Wiley & Sons Inc.

- Kaplan, H. S., Hooper, P. L., & Gurven, M. (2009). The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1533), 3289–3299. <https://doi.org/10.1098/rstb.2009.0115>
- Kasari, C., Locke, J., Gulsrud, A., & Rotheram-Fuller, E. (2010). Social Networks and Friendships at School: Comparing Children With and Without ASD. *Journal of Autism and Developmental Disorders*, 41(5), 533–544. <https://doi.org/10.1007/s10803-010-1076-x>
- Kaufmann, J. H. (1983). On the Definitions and Functions of Dominance and Territoriality. *Biological Reviews*, 58(1), 1–20. <https://doi.org/10.1111/j.1469-185X.1983.tb00379.x>
- Kendal, R. L., Coolen, I., van Bergen, Y., & Laland, K. N. (2005). Trade-Offs in the Adaptive Use of Social and Asocial Learning. In B.-A. in the S. of Behavior (Ed.) (Vol. 35, pp. 333–379). Academic Press. Retrieved from <http://www.sciencedirect.com/science/article/pii/S006534540535008X>
- Kim, P., Evans, G. W., Angstadt, M., Ho, S. S., Sripada, C. S., Swain, J. E., ... Phan, K. L. (2013). Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. *Proceedings of the National Academy of Sciences*, 110(46), 18442–18447. <https://doi.org/10.1073/pnas.1308240110>

- King, A. J., Johnson, D. D. P., & Van Vugt, M. (2009). The Origins and Evolution of Leadership. *Current Biology*, 19(19), R911–R916. <https://doi.org/10.1016/j.cub.2009.07.027>
- Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1430), 345–360. <https://doi.org/10.1098/rstb.2002.1202>
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society of London B: Biological Sciences*, 279(1738), 2698–2704. <https://doi.org/10.1098/rspb.2012.0311>
- Korva, N., Porter, S., O'Connor, B. P., Shaw, J., & Brinke, L. ten. (2013). Dangerous Decisions: Influence of Juror Attitudes and Defendant Appearance on Legal Decision-Making. *Psychiatry, Psychology and Law*, 20(3), 384–398. <https://doi.org/10.1080/13218719.2012.692931>
- Kramer, R. S. S., King, J. E., & Ward, R. (2011). Identifying personality from the static, nonexpressive face in humans and chimpanzees: evidence of a shared system for signaling personality. *Evolution and Human Behavior*, 32(3), 179–185. <https://doi.org/10.1016/j.evolhumbehav.2010.10.005>
- Krosch, A. R., & Amodio, D. M. (2014). Economic scarcity alters the perception of race. *Proceedings of the National Academy of*

- Sciences*, 111(25), 9079–9084.
<https://doi.org/10.1073/pnas.1404448111>
- Kross, E., Berman, M. G., Mischel, W., Smith, E. E., & Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proceedings of the National Academy of Sciences*, 108(15), 6270–6275. <https://doi.org/10.1073/pnas.1102693108>
- Krumhuber, E., R, S., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, 7(4), 730–735. <https://doi.org/10.1037/1528-3542.7.4.730>
- Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, 32(1), 4–14.
- Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, 107(1), 330–342. <https://doi.org/10.1016/j.cognition.2007.07.012>
- Laustsen, L. (2013). Decomposing the Relationship Between Candidates' Facial Appearance and Electoral Success. *Political Behavior*, 36(4), 777–791. <https://doi.org/10.1007/s11109-013-9253-1>
- Laustsen, L., & Petersen, M. B. (2015). Does a competent leader make a good friend? Conflict, ideology and the psychologies of friendship and followership. *Evolution and Human Behavior*, 36(4), 286–293. <https://doi.org/10.1016/j.evolhumbehav.2015.01.001>

- Lavie, N., Ro, T., & Russell, C. (2003). The Role of Perceptual Load in Processing Distractor Faces. *Psychological Science*, *14*(5), 510–515. <https://doi.org/10.1111/1467-9280.03453>
- Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P., & Penton-Voak, I. S. (2014). Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biology Letters*, *10*(10), 20140729. <https://doi.org/10.1098/rsbl.2014.0729>
- Lefevre, C. E., Wilson, V. A. D., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Facial Width-To-Height Ratio Relates to Alpha Status and Assertive Personality in Capuchin Monkeys. *PLOS ONE*, *9*(4), e93369. <https://doi.org/10.1371/journal.pone.0093369>
- Lewis, G. J., Lefevre, C. E., & Bates, T. C. (2012). Facial width-to-height ratio predicts achievement drive in US presidents. *Personality and Individual Differences*, *52*(7), 855–857. <https://doi.org/10.1016/j.paid.2011.12.030>
- Lewis, H. S. (1974). *Leaders and followers: Some anthropological perspectives*. Addison-Wesley Reading, MA: Retrieved from http://www.academia.edu/download/5076567/HSL_-_Leadership.pdf
- Li, J.-Z., Gui, D.-Y., Feng, C.-L., Wang, W.-Z., Du, B.-Q., Gan, T., & Luo, Y.-J. (2012). Victims' Time Discounting 2.5 Years after the Wenchuan Earthquake: An ERP Study. *PLOS ONE*, *7*(7), e40316. <https://doi.org/10.1371/journal.pone.0040316>

- Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, 7(3), 274–281. <https://doi.org/10.1093/scan/nsr006>
- Lin, N. (1999). Social Networks and Status Attainment. *Annual Review of Sociology*, 25, 467–487.
- Little, A. C. (2014). Facial appearance and leader choice in different contexts: Evidence for task contingent selection based on implicit and learned face-behaviour/face-ability associations. *The Leadership Quarterly*, 25(5), 865–874. <https://doi.org/10.1016/j.leaqua.2014.04.002>
- Little, A. C. (2017). Evolutionary explanations for financial and prosocial biases: Beyond mating motivation. *Behavioral and Brain Sciences*, 40. <https://doi.org/10.1017/S0140525X16000558>
- Little, A. C., Apicella, C. L., & Marlowe, F. W. (2007). Preferences for symmetry in human faces in two cultures: data from the UK and the Hadza, an isolated group of hunter-gatherers. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1629), 3113–3117. <https://doi.org/10.1098/rspb.2007.0895>
- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007a). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1), 18–27. <https://doi.org/10.1016/j.evolhumbehav.2006.09.002>

- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007b). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1), 18–27. <https://doi.org/10.1016/j.evolhumbehav.2006.09.002>
- Little, A. C., DeBruine, L. M., & Jones, B. C. (2010). Exposure to visual cues of pathogen contagion changes preferences for masculinity and symmetry in opposite-sex faces. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20101925. <https://doi.org/10.1098/rspb.2010.1925>
- Little, A. C., Roberts, S. C., Jones, B. C., & DeBruine, L. M. (2012). The perception of attractiveness and trustworthiness in male faces affects hypothetical voting decisions differently in wartime and peacetime scenarios. *The Quarterly Journal of Experimental Psychology*, 65(10), 2018–2032. <https://doi.org/10.1080/17470218.2012.677048>
- Liu, L., Feng, T., Suo, T., Lee, K., & Li, H. (2012). Adapting to the Destitute Situations: Poverty Cues Lead to Short-Term Choice. *PLOS ONE*, 7(4), e33950. <https://doi.org/10.1371/journal.pone.0033950>
- Livshits, G., & Kobylansky, E. (1991). Fluctuating Asymmetry As a Possible Measure of Developmental Homeostasis in Humans: A Review. *Human Biology*, 63(4), 441–466.
- Locke, J., Ishijima, E. H., Kasari, C., & London, N. (2010). Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting. *Journal of Research in Special Educational Needs*, 10(2), 74–81. <https://doi.org/10.1111/j.1471-3802.2010.01148.x>

- Matson, J. L., & Nebel-Schwalm, M. S. (2007). Comorbid psychopathology with autism spectrum disorder in children: An overview. *Research in Developmental Disabilities, 28*(4), 341–352. <https://doi.org/10.1016/j.ridd.2005.12.004>
- McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE, 9*(3), e90779. <https://doi.org/10.1371/journal.pone.0090779>
- McCann, S. J. H., & Stewin, L. L. (1990). Good and Bad Years: An Index of American Social, Economic, and Political Threat (1920–1986). *The Journal of Psychology, 124*(6), 601–617. <https://doi.org/10.1080/00223980.1990.10543253>
- McCullough, M. E., Pedersen, E. J., Schroder, J. M., Tabak, B. A., & Carver, C. S. (2013). Harsh childhood environmental characteristics predict exploitation and retaliation in humans. *Proceedings of the Royal Society B: Biological Sciences, 280*(1750), 20122104. <https://doi.org/10.1098/rspb.2012.2104>
- McNamara, J. M., Barta, Z., Fromhage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature, 451*(7175), 189–192. <https://doi.org/10.1038/nature06455>
- Mell, H., Safra, L., Algan, Y., Baumard, N., & Chevallier, C. (2017). Childhood Environmental Harshness Predicts Coordinated Health and Reproductive Strategies: A Cross-sectional Study of A Nationally Representative (france) Sample. Retrieved from *Osf. Io/urdky*. Retrieved from <https://osf.io/4rzfy/?action=download>

- Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, C. N. (2011). Do birds of a feather move together? Group membership and behavioral synchrony. *Experimental Brain Research*, 211(3-4), 495–503. <https://doi.org/10.1007/s00221-011-2641-z>
- Mileva, V. R., Cowan, M. L., Cobey, K. D., Knowles, K. K., & Little, A. C. (2014). In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences*, 69, 115–118. <https://doi.org/10.1016/j.paid.2014.05.019>
- Mirisola, A., Roccato, M., Russo, S., Spagna, G., & Vieno, A. (2014). Societal Threat to Safety, Compensatory Control, and Right-Wing Authoritarianism. *Political Psychology*, 35(6), 795–812. <https://doi.org/10.1111/pops.12048>
- Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015a). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, 109(4), 604–621. <https://doi.org/10.1037/pspi0000028>
- Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015b). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, 109(4), 604–621. <https://doi.org/10.1037/pspi0000028>

- Mohtashemi, M., & Mui, L. (2003). Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism. *Journal of Theoretical Biology*, 223(4), 523–531. [https://doi.org/10.1016/S0022-5193\(03\)00143-7](https://doi.org/10.1016/S0022-5193(03)00143-7)
- Møller, A. P., & Thornhill, R. (1998). Bilateral Symmetry and Sexual Selection: A Meta-Analysis. *The American Naturalist*, 151(2), 174–192. <https://doi.org/10.1086/286110>
- Montano, K. J., Tigue, C. C., Isenstein, S. G. E., Barclay, P., & Feinberg, D. R. (2017). Men's voice pitch influences women's trusting behavior. *Evolution and Human Behavior*, 38(3), 293–297. <https://doi.org/10.1016/j.evolhumbehav.2016.10.010>
- Mundy, P. (2003). Annotation: The neural basis of social impairments in autism: the role of the dorsal medial-frontal cortex and anterior cingulate system. *Journal of Child Psychology and Psychiatry*, 44(6), 793–809. <https://doi.org/10.1111/1469-7610.00165>
- Murray, D. R., Schaller, M., & Suedfeld, P. (2013). Pathogens and Politics: Further Evidence That Parasite Prevalence Predicts Authoritarianism. *PLOS ONE*, 8(5), e62275. <https://doi.org/10.1371/journal.pone.0062275>
- National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria. (2016). Behavioral Assessment Methods for RDoC Constructs. *Bethesda, MD: National Institute of Mental Health*.

- National Institute of Mental Health. (2015). Strategic Plan for Research. Retrieved from https://www.nimh.nih.gov/about/strategic-planning-reports/nimh_strategicplanforresearch_508compliant_corrected_final_149979.pdf
- Neff, B. D., & Pitcher, T. E. (2005). Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology*, *14*(1), 19–38. <https://doi.org/10.1111/j.1365-294X.2004.02395.x>
- Nettle, D. (2005). An evolutionary approach to the extraversion continuum. *Evolution and Human Behavior*, *26*(4), 363–373. <https://doi.org/10.1016/j.evolhumbehav.2004.12.004>
- Nettle, D., Coall, D. A., & Dickins, T. E. (2010). Birthweight and paternal involvement predict early reproduction in British women: Evidence from the National Child Development Study. *American Journal of Human Biology*, *22*(2), 172–179. <https://doi.org/10.1002/ajhb.20970>
- Nettle, D., Coall, D. A., & Dickins, T. E. (2011). Early-life conditions and age at first pregnancy in British women. *Proceedings of the Royal Society of London B: Biological Sciences*, *278*(1712), 1721–1727. <https://doi.org/10.1098/rspb.2010.1726>
- Nettle, D., Colléony, A., & Cockerill, M. (2011). Variation in Cooperative Behaviour within a Single City. *PLOS ONE*, *6*(10), e26922. <https://doi.org/10.1371/journal.pone.0026922>

- Nettle, D., Frankenhuys, W. E., & Rickard, I. J. (2013). The evolution of predictive adaptive responses in human life history. *Proc. R. Soc. B*, 280(1766), 20131343. <https://doi.org/10.1098/rspb.2013.1343>
- Nevicka, B., De Hoogh, A. H. B., van Vianen, A. E. M., & Ten Velden, F. S. (2013). Uncertainty enhances the preference for narcissistic leaders. *European Journal of Social Psychology*, 43(5), 370–380. <https://doi.org/10.1002/ejsp.1943>
- Nichols, A. L., & Cottrell, C. A. (2014). What do people desire in their leaders? The role of leadership level on trait desirability. *The Leadership Quarterly*, 25(4), 711–729. <https://doi.org/10.1016/j.leaqua.2014.04.001>
- Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences*, 33(6), 417–433. <https://doi.org/10.1017/S0140525X10000865>
- O’Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41(2), 147–155. [https://doi.org/10.1016/S0028-3932\(02\)00145-8](https://doi.org/10.1016/S0028-3932(02)00145-8)
- Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-Based Trait Inferences and Voting. *Journal of Nonverbal Behavior*, 34(2), 83–110. <https://doi.org/10.1007/s10919-009-0082-1>
- Olsen, E. K., Bjorkquist, O. A., Bodapati, A. S., Shankman, S. A., & Herbener, E. S. (2015). Associations between trait anhedonia and

- emotional memory deficits in females with schizophrenia versus major depression. *Psychiatry Research*, 230(2), 323–330. <https://doi.org/10.1016/j.psychres.2015.09.012>
- Olson, I. R., & Marshuetz, C. (2005). Facial Attractiveness Is Appraised in a Glance. *Emotion*, 5(4), 498–502. <https://doi.org/10.1037/1528-3542.5.4.498>
- Onraet, E., Dhont, K., & Hiel, A. V. (2014). The Relationships Between Internal and External Threats and Right-Wing Attitudes A Three-Wave Longitudinal Study. *Personality and Social Psychology Bulletin*, 40(6), 712–725. <https://doi.org/10.1177/0146167214524256>
- Onraet, E., van Hiel, A., & Cornelis, I. (2013). Threat and Right-Wing Attitudes: A Cross-National Approach. *Political Psychology*, 34(5), 791–803. <https://doi.org/10.1111/pops.12014>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087–11092. <https://doi.org/10.1073/pnas.0805664105>
- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, 9(1), 128–133. <https://doi.org/10.1037/a0014520>
- Pepper, G. V., & Nettle, D. (2017). The behavioural constellation of deprivation: causes and consequences. *The Behavioral and Brain Sciences*, 1.
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and Human Facial Attractiveness.

- Evolution and Human Behavior*, 20(5), 295–307.
[https://doi.org/10.1016/S1090-5138\(99\)00014-8](https://doi.org/10.1016/S1090-5138(99)00014-8)
- Perry, R., Sibley, C. G., & Duckitt, J. (2013). Dangerous and competitive worldviews: A meta-analysis of their associations with Social Dominance Orientation and Right-Wing Authoritarianism. *Journal of Research in Personality*, 47(1), 116–127.
<https://doi.org/10.1016/j.jrp.2012.10.004>
- Petersen, M. B., & Aarøe, L. (2015). Birth Weight and Social Trust in Adulthood: Evidence for Early Calibration of Social Cognition. *Psychological Science*, 26(11), 1681–1692.
<https://doi.org/10.1177/0956797615595622>
- Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage*, 101, 124–137.
<https://doi.org/10.1016/j.neuroimage.2014.06.061>
- Porter, S., Brinke, L. ten, & Gustaw, C. (2010). Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, 16(6), 477–491. <https://doi.org/10.1080/10683160902926141>
- Rademacher, L., Krach, S., Kohls, G., Irmak, A., Gründer, G., & Spreckelmeyer, K. N. (2010). Dissociation of neural networks for anticipation and consumption of monetary and social rewards.

- NeuroImage*, 49(4), 3276–3285.
<https://doi.org/10.1016/j.neuroimage.2009.10.089>
- Ramsey, J. L., Langlois, J. H., Hoss, R. A., Rubenstein, A. J., & Griffin, A. M. (2004). Origins of a stereotype: categorization of facial attractiveness by 6-month-old infants. *Developmental Science*, 7(2), 201–211.
<https://doi.org/10.1111/j.1467-7687.2004.00339.x>
- Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial Cues to Perceived Height Influence Leadership Choices in Simulated War and Peace Contexts. *Evolutionary Psychology*, 11(1), 89–103.
<https://doi.org/10.1177/147470491301100109>
- Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior*, 33(3), 200–209. <https://doi.org/10.1016/j.evolhumbehav.2011.09.003>
- Rekers, Y., Haun, D. B. M., & Tomasello, M. (2011). Children, but Not Chimpanzees, Prefer to Collaborate. *Current Biology*, 21(20), 1756–1758. <https://doi.org/10.1016/j.cub.2011.08.066>
- Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable Facial Configurations Affect Strategic Choices in Trust Games with or without Information about Past Behavior. *PLOS ONE*, 7(3), e34293.
<https://doi.org/10.1371/journal.pone.0034293>
- Rhodes, G., Chan, J., Zebrowitz, L. A., & Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proceedings of the*

- Royal Society of London B: Biological Sciences*, 270(Suppl 1), S93–S95. <https://doi.org/10.1098/rsbl.2003.0023>
- Richerson, P., & Henrich, J. (2012). Tribal Social Instincts and the Cultural Evolution of Institutions to Solve Collective Action Problems. *Cliodynamics: The Journal of Theoretical and Mathematical History*, 3(1). Retrieved from <http://escholarship.org/uc/item/981121t8>
- Roff, D. A. (2002). *Life history evolution*. Sunderland: Sinauer Associates Inc.
- Ro, T., Friggel, A., & Lavie, N. (2007). Attentional biases for faces and body parts. *Visual Cognition*, 15(3), 322–348. <https://doi.org/10.1080/13506280600590434>
- Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, 15(8), 549–562. <https://doi.org/10.1038/nrn3776>
- Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, 104(3), 409–426. <https://doi.org/10.1037/a0031050>
- Safra, L., Algan, Y., Grèzes, J., Teodora, T., Baumard, N., & Chevallier, C. (2017). Childhood harshness predicts long-lasting leader preferences. *Evolution and Human Behavior*. <https://doi.org/https://osf.io/v7k73/>
- Safra, L., Chevallier, C., Blakemore, S.-J., & Palminteri, S. (in prep.). Social information impairs reward learning in depressive subjects: behavioral and computational evidence.

- Safra, L., Tecu, T., Lambert, S., Sheskin, M., Baumard, N., & Chevallier, C. (2016). Neighborhood Deprivation Negatively Impacts Children's Prosocial Behavior. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01760>
- Safra, L., Wyart, V., Baumard, N., & Chevallier, C. (submitted). Face evaluations confirm that social motivation serves cooperative goals. *Nature Human Behaviour*.
- Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain systems for assessing the affective value of faces. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366(1571), 1660–1670. <https://doi.org/10.1098/rstb.2010.0351>
- Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, 9(2), 260–264. <https://doi.org/10.1037/a0014681>
- Sales, S. M. (1973). Threat as a factor in authoritarianism: An analysis of archival data. *Journal of Personality and Social Psychology*, 28(1), 44–57. <https://doi.org/10.1037/h0035588>
- Salva, O. R., Farroni, T., Regolin, L., Vallortigara, G., & Johnson, M. H. (2011). The Evolution of Social Orienting: Evidence from Chicks (*Gallus gallus*) and Human Newborns. *PLOS ONE*, 6(4), e18802. <https://doi.org/10.1371/journal.pone.0018802>
- Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European*

- Journal of Cognitive Psychology*, 19(3), 470–480.
<https://doi.org/10.1080/09541440600757426>
- Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinsen, R. K., ... Cuthbert, B. N. (2010). Developing constructs for psychopathology research: research domain criteria. *Journal of Abnormal Psychology*, 119(4), 631–639.
<https://doi.org/10.1037/a0020909>
- Satow, K. L. (1975). Social approval and helping. *Journal of Experimental Social Psychology*, 11(6), 501–509. [https://doi.org/10.1016/0022-1031\(75\)90001-3](https://doi.org/10.1016/0022-1031(75)90001-3)
- Scheeren, A. M., Begeer, S., Banerjee, R., Terwogt, M. M., & Koot, H. M. (2010). Can you tell me something about yourself? Self-presentation in children and adolescents with high functioning autism spectrum disorder in hypothetical and real life situations. *Autism*.
<https://doi.org/10.1177/1362361310366568>
- Sell, A., Cosmides, L., & Tooby, J. (2014). The human anger face evolved to enhance cues of strength. *Evolution and Human Behavior*, 35(5), 425–429. <https://doi.org/10.1016/j.evolhumbehav.2014.05.008>
- Shattuck, P. T., Orsmond, G. I., Wagner, M., & Cooper, B. P. (2011). Participation in Social Activities among Adolescents with an Autism Spectrum Disorder. *PLOS ONE*, 6(11), e27176.
<https://doi.org/10.1371/journal.pone.0027176>

- Sheskin, M., Chevallier, C., Lambert, S., & Baumard, N. (2014). Life-history theory explains childhood moral development. *Trends in Cognitive Sciences*, 18(12), 613–615. <https://doi.org/10.1016/j.tics.2014.08.004>
- Sibley, C. G., Wilson, M. S., & Duckitt, J. (2007). Effects of Dangerous and Competitive Worldviews on Right-Wing Authoritarianism and Social Dominance Orientation over a Five-Month Period. *Political Psychology*, 28(3), 357–371. <https://doi.org/10.1111/j.1467-9221.2007.00572.x>
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(8), 921–929. <https://doi.org/10.1097/CHI.0b013e318179964f>
- Slepian, M. L., & Ames, D. R. (2016). Internalized Impressions. *Psychological Science*, 27(2), 282–288. <https://doi.org/10.1177/0956797615594897>
- Sloboda, D. M., Hart, R., Doherty, D. A., Pennell, C. E., & Hickey, M. (2007). Age at Menarche: Influences of Prenatal and Postnatal Growth. *The Journal of Clinical Endocrinology & Metabolism*, 92(1), 46–50. <https://doi.org/10.1210/jc.2006-1378>
- Smith, J. M., & Harper, D. (2003). *Animal Signals*. OUP Oxford.
- Snyder, J. K., Fessler, D. M. T., Tiokhin, L., Frederick, D. A., Lee, S. W., & Navarrete, C. D. (2011). Trade-offs in a dangerous world: women's fear of crime predicts preferences for aggressive and formidable

- mates. *Evolution and Human Behavior*, 32(2), 127–137.
<https://doi.org/10.1016/j.evolhumbehav.2010.08.007>
- Song, R., Over, H., & Carpenter, M. (2016). Young children discriminate genuine from fake smiles and expect people displaying genuine smiles to be more prosocial. *Evolution and Human Behavior*, 37(6), 490–501. <https://doi.org/10.1016/j.evolhumbehav.2016.05.002>
- Sperber, D., & Baumard, N. (2012). Moral Reputation: An Evolutionary and Cognitive Perspective. *Mind & Language*, 27(5), 495–518.
<https://doi.org/10.1111/mila.12000>
- Spisak, B. R., Dekker, P. H., Krüger, M., & Vugt, M. van. (2012). Warriors and Peacekeepers: Testing a Biosocial Implicit Leadership Hypothesis of Intergroup Relations Using Masculine and Feminine Faces. *PLOS ONE*, 7(1), e30399. <https://doi.org/10.1371/journal.pone.0030399>
- Spisak, B. R., Homan, A. C., Grabo, A., & Van Vugt, M. (2012). Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *The Leadership Quarterly*, 23(2), 273–280.
<https://doi.org/10.1016/j.leaqua.2011.08.006>
- Stearns, S. C. (2000). Life history evolution: successes, limitations, and prospects. *Naturwissenschaften*, 87(11), 476–486.
<https://doi.org/10.1007/s001140050763>
- Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences*, 8(2), 60–65. <https://doi.org/10.1016/j.tics.2003.12.003>

- Stewart, L. H., Ajina, S., Getov, S., Bahrami, B., Todorov, A., & Rees, G. (2012). Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology: General*, 141(4), 715–727. <https://doi.org/10.1037/a0027950>
- Stockley, P., & Bro-Jørgensen, J. (2011). Female competition and its evolutionary consequences in mammals. *Biological Reviews*, 86(2), 341–366. <https://doi.org/10.1111/j.1469-185X.2010.00149.x>
- Sussman, A. B., Petkova, K., & Todorov, A. (2013). Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology*, 49(4), 771–775. <https://doi.org/10.1016/j.jesp.2013.02.003>
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127(1), 105–118. <https://doi.org/10.1016/j.cognition.2012.12.001>
- Tennie, C., Frith, U., & Frith, C. D. (2010). Reputation management in the age of the world-wide web. *Trends in Cognitive Sciences*, 14(11), 482–488. <https://doi.org/10.1016/j.tics.2010.07.003>
- Thornhill, R., & Gangestad, S. W. (1994). Human Fluctuating Asymmetry and Sexual Behavior. *Psychological Science*, 5(5), 297–302. <https://doi.org/10.1111/j.1467-9280.1994.tb00629.x>
- Thornhill, R., & Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and

- women. *Evolution and Human Behavior*, 27(2), 131–144.
<https://doi.org/10.1016/j.evolhumbehav.2005.06.001>
- Thornhill, R., Gangestad, S. W., Thornhill, R., Gangestad, S. W., Thornhill, R., Gangestad, S. W., ... Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, 3(12), 452–460.
[https://doi.org/10.1016/S1364-6613\(99\)01403-5](https://doi.org/10.1016/S1364-6613(99)01403-5)
- Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216.
<https://doi.org/10.1016/j.evolhumbehav.2011.09.004>
- Tingley, D. (2014). Face-Off: Facial Features and Strategic Choice. *Political Psychology*, 35(1), 35–55. <https://doi.org/10.1111/pops.12041>
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005a). Inferences of Competence from Faces Predict Election Outcomes. *Science*, 308(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005b). Inferences of Competence from Faces Predict Election Outcomes. *Science*, 308(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, 66(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>

- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating Faces on Trustworthiness After Minimal Time Exposure. *Social Cognition*, 27(6), 813–833. <https://doi.org/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 Sciences*, 12(12), 455–460. <https://doi.org/10.1016/j.tics.2008.10.001>
- Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2013). Is cooperativeness readable in static facial features? An inter-cultural approach. *Evolution and Human Behavior*, 34(6), 427–432. <https://doi.org/10.1016/j.evolhumbehav.2013.08.002>
- Valentine, K. A., Li, N. P., Penke, L., & Perrett, D. I. (2014). Judging a Man by the Width of His Face. *Psychological Science*, 25(3), 806–811. <https://doi.org/10.1177/0956797613511823>
- Van de Vyver, J., Houston, D. M., Abrams, D., & Vasiljevic, M. (2016). Boosting Belligerence How the July 7, 2005, London Bombings Affected Liberals' Moral Foundations and Prejudice. *Psychological Science*, 27(2), 169–177. <https://doi.org/10.1177/0956797615615584>
- van Honk, J., Montoya, E., Bos, P., Vugt, M. van, & Terburg, D. (2012). New evidence on testosterone and cooperation. 485. Retrieved from <http://dare.ubvu.vu.nl/handle/1871/42054>
- 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. <https://doi.org/10.1016/j.cognition.2008.07.002>

- Van Vugt, M. (2006). Evolutionary Origins of Leadership and Followership. *Personality and Social Psychology Review*, 10(4), 354–371. https://doi.org/10.1207/s15327957pspr1004_5
- Van Vugt, M., & Kurzban, R. K. (2007). Cognitive and social adaptations for leadership and followership: Evolutionary game theory and group dynamics. In J. P. Forgas, W. von Hippel, & M. G. Haselton (Eds.), *Evolution and the Social Mind: Evolutionary Psychology and Social Cognition* (pp. 229–244). United Kingdom: Psychology Press Ltd. Retrieved from <https://kar.kent.ac.uk/4533/>
- Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences*, 111(32), E3353–E3361. <https://doi.org/10.1073/pnas.1409860111>
- Viitanen, T. (2014). Parental divorce and other determinants of interpersonal trust: Evidence from HILDA panel data. *Australian Journal of Labour Economics*, 17(1), 35.
- von Rueden, C. (2014). The Roots and Fruits of Social Status in Small-Scale Human Societies. In J. T. Cheng, J. L. Tracy, & C. Anderson (Eds.), *The Psychology of Social Status* (pp. 179–200). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4939-0867-7_9
- von Rueden, C., Gurven, M., & Kaplan, H. (2010). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal*

- Society of London B: Biological Sciences*, rspb20102145.
<https://doi.org/10.1098/rspb.2010.2145>
- von Rueden, C., Gurven, M., Kaplan, H., & Stieglitz, J. (2014). Leadership in an Egalitarian Society. *Human Nature (Hawthorne, N.Y.)*, 25(4), 538–566. <https://doi.org/10.1007/s12110-014-9213-4>
- Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The Tuning of Human Neonates' Preference for Speech. *Child Development*, 81(2), 517–527. <https://doi.org/10.1111/j.1467-8624.2009.01412.x>
- Vullioud, C., Clément, F., Scott-Phillips, T., & Mercier, H. (2017). Confidence as an expression of commitment: why misplaced expressions of confidence backfire. *Evolution and Human Behavior*, 38(1), 9–17. <https://doi.org/10.1016/j.evolhumbehav.2016.06.002>
- Wang, J. J., & Apperly, I. A. (2017). Just one look: Direct gaze briefly disrupts visual working memory. *Psychonomic Bulletin & Review*, 24(2), 393–399. <https://doi.org/10.3758/s13423-016-1097-3>
- Watkins, C. D., DeBruine, L. M., Little, A. C., Feinberg, D. R., & Jones, B. C. (2012). Priming concerns about pathogen threat versus resource scarcity: dissociable effects on women's perceptions of men's attractiveness and dominance. *Behavioral Ecology and Sociobiology*, 66(12), 1549–1556. <https://doi.org/10.1007/s00265-012-1408-2>
- Watkins, C. D., & Jones, B. C. (2012). Priming men with different contest outcomes modulates their dominance perceptions. *Behavioral Ecology*, arr221. <https://doi.org/10.1093/beheco/arr221>

- Waynforth, D. (1998). Fluctuating asymmetry and human male life-history traits in rural Belize. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1405), 1497–1501. <https://doi.org/10.1098/rspb.1998.0463>
- Webster, G. D., Graber, J. A., Gesselman, A. N., Crosier, B. S., & Schember, T. O. (2014). A Life History Theory of Father Absence and Menarche: A Meta-Analysis. *Evolutionary Psychology*, 12(2), 147470491401200202. <https://doi.org/10.1177/147470491401200202>
- West, S. (2004). *Portraiture*. OUP Oxford.
- White, A. E., Kenrick, D. T., & Neuberg, S. L. (2013). Beauty at the Ballot Box Disease Threats Predict Preferences for Physically Attractive Leaders. *Psychological Science*, 24(12), 2429–2436. <https://doi.org/10.1177/0956797613493642>
- Wibral, M., Dohmen, T., Klingmüller, D., Weber, B., & Falk, A. (2012). Testosterone Administration Reduces Lying in Men. *PLOS ONE*, 7(10), e46774. <https://doi.org/10.1371/journal.pone.0046774>
- Willis, J., & Todorov, A. (2006). First Impressions Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, 17(7), 592–598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Wilson, J. P., & Rule, N. O. (2015a). Facial Trustworthiness Predicts Extreme Criminal-Sentencing Outcomes. *Psychological Science*, 26(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>

- Wilson, J. P., & Rule, N. O. (2015b). Facial Trustworthiness Predicts Extreme Criminal-Sentencing Outcomes. *Psychological Science, 26*(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>
- Wilson, V., Lefevre, C. E., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Personality and facial morphology: Links to assertiveness and neuroticism in capuchins (*Sapajus [Cebus] apella*). *Personality and Individual Differences, 58*, 89–94. <https://doi.org/10.1016/j.paid.2013.10.008>
- Woodburn, J. (1982). Egalitarian Societies. *Man, 17*(3), 431–451. <https://doi.org/10.2307/2801707>
- Zebrowitz, L. A., Franklin, R. G., & Palumbo, R. (2015). Ailing Voters Advance Attractive Congressional Candidates. *Evolutionary Psychology, 13*(1), 16–28. <https://doi.org/10.1177/147470491501300102>

Appendix

1. Neighborhood Deprivation Negatively Impacts Children's Prosocial Behavior, Safra*, L., Tecu*, T., Lambert*, S., Sheskin, M., Baumard, N., Chevallier, C., *Frontiers in Psychology* (2016)
2. Climate is not a good candidate to account for variations in aggression and violence across space and time, Mell, H., Safra, L., Baumard, N., Jacquet, P.O., *commentary in Behavioral and Brain Sciences* (2017)
3. Measuring Social Motivation Using Signal Detection and Reward Responsiveness, Chevallier, C., Tonge, N., Safra, L., Kahn, D., Kohls, G., Miller, J., Schultz, R., *PloS one* (2016)
4. Childhood Environmental Harshness Predicts Coordinated Health and Reproductive Strategies: A Cross-sectional Study Of A Nationally Representative (France) Sample., Mell, H., Safra, L., Algan, Y., Baumard, N., Chevallier, C., *in revisions for Evolution and Human Behavior*
5. Social information impairs reward learning in depressive subjects: behavioral and computational evidence, Safra, L., Chevallier, C., Blakemore, S.-J., Palminteri, S., *in preparation*
6. Increased susceptibility to social influence depends on grater environmental harshness and unpredictability experienced during childhood, Jacquet, P.O., Safra, L., Baumard, N., Chevallier, C., *preregistered report in preparation*



Neighborhood Deprivation Negatively Impacts Children's Prosocial Behavior

Lou Safra^{1*}, Teodora Tecu^{2†}, Stéphane Lambert^{3†}, Mark Sheskin³, Nicolas Baumard³ and Coralie Chevallier^{1*}

¹ Laboratoire de Neurosciences Cognitives, INSERM U960, Département d'Études Cognitives, Ecole Normale Supérieure – PSL Research University, Paris, France, ² Faculty of Philosophy, University of Bucharest, Bucharest, Romania, ³ Institut Jean-Nicod, CNRS UMR8129, Département d'Études Cognitives, Ecole Normale Supérieure – PSL Research University, Paris, France

Children show stronger cooperative behavior in experimental settings as they get older, but little is known about how the environment of a child shapes this development. In adults, prosocial behavior toward strangers is markedly decreased in low socio-economic status (SES) neighborhoods, suggesting that environmental harshness has a negative impact on some prosocial behaviors. Similar results have been obtained with 9-year-olds recruited from low vs. high SES schools. In the current study, we investigate whether these findings generalize to a younger age group and a developing country. Specifically, we worked with a sample of thirty-nine 6- to 7-year-olds in two neighborhoods in a single city in Romania. Using a “Quality Dictator Game” that offers greater resolution than previous measures, we find that children living in the harsher neighborhood behave less prosocially toward a stranger than children living in the less harsh neighborhood.

Keywords: prosociality, poverty, deprivation, behavioral ecology, SES, dictator game

OPEN ACCESS

Edited by:

Ann Dowker,
University of Oxford, UK

Reviewed by:

Tobias Krettenauer,
Wilfrid Laurier University, Canada
Antonio S. Silva,
University College London, UK

*Correspondence:

Lou Safra
lou.safra@gmail.com
Coralie Chevallier
coralie.chevallier@gmail.com

† These authors have contributed
equally to this work and are first
authors.

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 31 May 2016

Accepted: 26 October 2016

Published: 14 November 2016

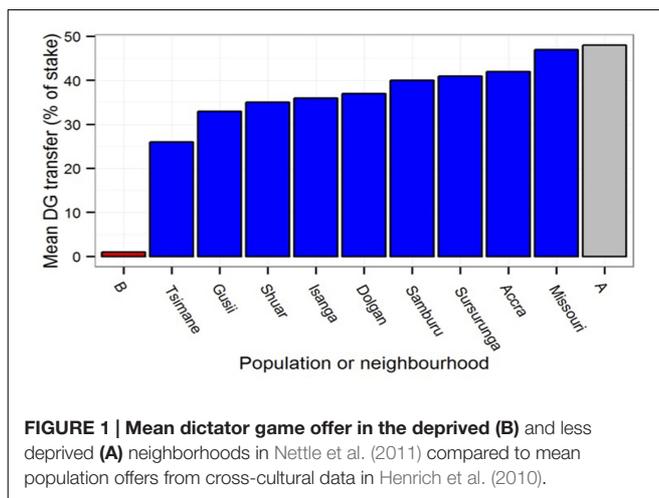
Citation:

Safra L, Tecu T, Lambert S,
Sheskin M, Baumard N and
Chevallier C (2016) Neighborhood
Deprivation Negatively Impacts
Children's Prosocial Behavior.
Front. Psychol. 7:1760.
doi: 10.3389/fpsyg.2016.01760

INTRODUCTION

Prosocial behaviors and motivations emerge early in development, with children in their second year already motivated to provide information to others, to spontaneously pick up objects to help others, and to comfort others in distress (e.g., Eisenberg and Miller, 1987; Warneken and Tomasello, 2006, 2009; Dunfield et al., 2011). Although more costly forms of prosociality, like spontaneous sharing, typically emerge later (Smith et al., 2013; Sheskin et al., 2014), costly prosociality can be seen during the preschool years in cooperative contexts (Hamann et al., 2011).

Less is known about how this development is shaped by a child's environment. In adults, several studies have demonstrated that a behavioral ecology approach can predict part of the variability observed in the prosocial behavior of different individuals. In a study comparing prosocial behaviors in a very deprived and a more affluent neighborhood of Newcastle-upon-Tyne (UK), Nettle et al. (2011) found that participants from the very deprived neighborhood gave substantially less in a dictator game than participants from the more affluent neighborhood. Strikingly, the difference in prosocial behavior in the dictator game in different neighborhoods of this *single* city was an order of magnitude larger than the largest differences found in previous research on differences across cultures (Henrich et al., 2010). This suggests that environmental harshness *within* a single culture—and not only differences *across* cultures—calibrates prosocial motivations; indeed, environmental harshness within a culture may be far more important than any cross-cultural differences (see **Figure 1**).



Consistent with these results, in another study of 20 neighborhoods in London (UK), a strong negative effect of neighborhood income deprivation on altruistic behavior was found, with letters dropped in the poorest neighborhoods having 91% lower odds of being returned than letters dropped in the wealthiest neighborhoods (Holland et al., 2012; see also Silva and Mace, 2014, 2015).

Importantly, research with adults does *not* show decreased prosocial behavior in all situations. In the studies reported by Nettle et al. (2011) for instance, there was no difference across neighborhoods in the likelihood to help in face-to-face interactions (when the experimenter dropped an object, asked for directions, or needed to make change). A potential explanation for this phenomenon is that people in harsher environments are less likely to incur a cost to assist unseen individuals (especially when those individuals are anonymous), but are just as likely as people in more secure environments to assist individuals who are nearby. This pattern of results may be explained by an analysis of the costs and benefits of acting prosocially toward others: many accounts of moral behavior indeed emphasize that prosocial behaviors that improve one's moral reputation can give access to the long-term benefits associated with future cooperation (Barclay, 2011; Baumard et al., 2013). In harsher environments with fewer resources and less resource security, it can be dangerous to invest large amounts in one's cooperative reputation with strangers with whom one might never interact again (Baumard and Chevallier, 2015).

In the developmental literature, several studies have found that the trajectory of prosocial behavior shows both consistency and variation across diverse cultural contexts. In a study of 3- to 14-year-olds across six diverse societies, House et al. (2013) found that *costless* prosocial behavior increased with age in each society, whereas the development of *costly* prosocial behavior showed differences starting around middle childhood (when children's behavior started to tend toward the adult behavior prevalent in their culture). In a study of 4- to 15-year-olds across seven diverse societies, Blake et al. (2015) found that an aversion to receiving *less* than another child

unfairly emerged by middle childhood in each culture, but that aversion to receiving *more* than another child unfairly emerged in a minority of the cultures, and only later in development. Such results emphasize that the emergence of certain features of prosocial behavior is influenced by the environment.

The most directly relevant developmental study is one in which (similar to the study of adults within one city by Nettle et al. (2011), children within a single culture showed differences depending on local environmental harshness. Benenson et al. (2007) had 4-, 6-, and 9-year-old English children from low and high socio-economic status (SES) schools play a dictator game (with 10 stickers and an unknown recipient). They found that 9-year-olds (but not 4- and 6-year-olds) from high SES schools behaved more prosocially than their lower SES counterparts.

In the current study, we set out to replicate and extend the results by Benenson et al. (2007) and Nettle et al. (2011). As in previous research, we compared prosocial behavior from two samples within a single city, and the samples were recruited from neighborhoods that contrasted in deprivation level. Importantly, the participating schools were matched in terms of facilities, distance to the city center, and number of teachers. In contrast with both previous studies in which participants lived in a developed country (England), our research investigated whether similar results would be obtained in a developing country (Romania). A second extension of our study is that we gathered socio-demographic data on each child and their family to ensure that we would have a fine-grained measure of children's social environment. Finally, we used a different method, the "Quality Dictator Game" (adapted from Sheskin et al., 2016), that might be more adapted to young children, potentially allowing us to detect differences at younger ages than previous research.

The Quality Dictator Game investigates how children allocate windfall resources that vary in quality. The Quality Dictator Game allows for an additional analysis that is not possible in studies using a pool of identical resources (e.g., 10 stickers in Benenson et al., 2007). Whereas studies that use identical resources can only measure variation in the *number* of resources allocated by different participants, the Quality Dictator Game can measure variations in the *value* of allocated resources. Specifically, children first allocate four toys of varying quality between oneself and another child, then complete a distractor task, then rank 12 toys (including four toys identical to the ones used in the allocation task). This design produces a "transfer score" for each child based on the value of the toys she kept (subtracted from score) compared to the value of the toys she transferred to the other child (added to score). Taking children's individual preferences into account to calculate this transfer score thus provides a fine-grained measure of children's prosocial behavior.

Previous research on the impact of SES and neighborhood deprivation on prosocial behavior has generally used income as a criterion to define both low/high SES and deprived/affluent neighborhood. In our study, we made sure that both SES and neighborhood status were consistent: we thus included children in the deprived group only if their parents reported earning

less than the Romanian minimum monthly wage and if they lived in a deprived neighborhood. The reverse criteria were used in the non-deprived group. We predicted that children from deprived environments would have lower transfer scores than children from non-deprived environments. We had no specific predictions regarding gender, IQ, or ethnicity, and included them to control for potential effects of these variables on transfer scores.

MATERIALS AND METHODS

Participants

Children were recruited from two schools, both situated within the same city of Slatina (Romania). The schools are located 2.2 km away from one-another and are both about 1 km away from the city center. They are comparable in terms of number of students, qualified teachers, and facilities. They differ in SES: one school (School A) is located in a mostly middle-SES neighborhood and the other (School B) is situated in a very deprived neighborhood (Table 1). We excluded from the analysis children from School A whose parents' monthly income was lower than 850 lei and children from School B whose parents' combined income was higher than 850 lei. The threshold was fixed *a priori* to 850 lei as it was the minimum monthly wage for an employee in Romania at the time we conducted the study, See Monitorul Oficial, Partea I nr. 776 din 12.12.2013. Parents were asked what their combined income including social aids was, with only two response options: "less than 850 lei," "more than 850 lei." This value is approximately \$200.

All the results found by contrasting the two schools were confirmed using parental income (above or below the minimum wage) as the grouping variable (see Appendix 1 in Supplementary Material).

We tested 41 children aged 6–7 years ($M = 6.9$ years, $SD = 0.43$ years, range = 6.1–7.9): 18 children were girls (44%), and 20 identified with the Roma minority (49%). A minimum

target of 20 participants per group was pre-planned based on the number of 6- to 8-year-olds in the low-SES school; the exact number was determined by scheduling constraints and by the number of parental consent forms we received. One child from School A and one from School B were excluded, as they did not meet our pre-determined parental income criterion. Our remaining sample thus contained 19 children from School A ($M = 6.9$ years, $SD = 0.40$, range = 6.1–7.4), with seven girls (37%) and 16 children declaring being Roma (84%), and 19 children from School B ($M = 6.9$ years, $SD = 0.40$, range = 6.1–7.9) with 10 girls (53%) and four children declaring being Roma (21%). There was no significant difference between schools in term of gender (Fisher's t -test: $p = 0.32$), but children from School B were significantly more likely to declare belonging to the Roma minority (Fisher's t -test: $p < 0.001$).

Ethics

The study was approved by the schools' management team and by the School Inspectorate. Parents signed a written informed consent form for them and their children to participate in the study and for their anonymized data to be included in the analysis. Children provided verbal assent at the start of the procedure. This research is the result of a collaboration between a French and a Romanian university, both of which require no formal approval from an ethics committee for non-invasive research. All study procedures were consistent with the Declaration of Helsinki.

Procedure

Each child was tested individually in a quiet room, close to their own classroom. The child and the experimenter sat across from each other at an empty table. The experimenter introduced the activity by saying: "Today we will play a few games and I will ask you to help me do some tasks." The experiment consisted of three parts: a toy distribution task, the administration of the Raven's Color Progressive Matrices test, and the toy ranking task. Additionally, parents completed a questionnaire regarding the participating child's family and life conditions.

Toy Distribution Task

The toy distribution phase began by showing the four toys to be distributed: a yellow whistle, a white ping-pong ball, a pencil, and an arrow sticker. A pretest with a different group of children and a larger set of toys (see Appendix 2 in Supplementary Material) showed that the whistle and ping-pong ball were generally ranked among the highest value toys and that the pencil and arrow card were generally ranked among the lowest value toys. The experimenter explained that any toys put on a blue mat in front of the child would go to the child herself, whereas any toys put on a yellow mat opposite the child would go to another child "who I will see next week, who you don't know, and who you are not going to meet." We then asked the child: "So, where do you want to put the toys?" After the child allocated the four toys, they were put in two envelopes, one for the child and one for the "other child," and the envelopes and the placemats were put to the side for the remainder of the study.

TABLE 1 | Differences between School A and School B.

	School A	School B
School values	<ul style="list-style-type: none"> ● Integrative education ● Peace ● Team work ● Cooperation ● Self-development and affirmation 	<ul style="list-style-type: none"> ● Competence ● Initiative ● Competition ● Innovative spirit ● Sincerity ● Tolerance
Number of classrooms and teachers	<ul style="list-style-type: none"> ● 18 class-rooms ● 22 teachers 	<ul style="list-style-type: none"> ● 22 class-rooms ● 25 teachers
Material resources	<ul style="list-style-type: none"> ● TVs, copy machines, printers, video, projector, scanners, digital cameras ● Library (11,500 volumes) 	<ul style="list-style-type: none"> ● TVs, copy machines, printers, video, projector, scanners, digital cameras ● Library (7,804 volumes)

The data was collected from each school's Secretary Department and it is valid for the 2013–2014 school year.

Raven's Color Progressive Matrices Test

The Color Progressive Matrices test (Raven et al., 2003) is a non-verbal test assessing reasoning skills and providing an accurate estimate of IQ (Mackintosh, 1998). It comprises 36 items presented in increasing order of difficulty within each of three 12-item sets. The Raven's Color Progressive Matrices test provides a single raw score that is then converted to a percentile based on normative data. This test was included in our procedure so that we could control for individual differences in reasoning skills, and to provide a distractor task in between allocating toys in the toy distribution task, and judging the value of those toys in the toy ranking task. This long and challenging distractor task was chosen to limit the possibility that children would recognize that 4 out of 12 toys in the ranking task that were identical to the one they distributed in the toy distribution task, and have their ratings modified by endowment effects (Kahneman et al., 1991). The test took 15–30 min to administer.

Toy Ranking Task

The toy ranking task began by asking for the child's assistance in sorting 12 toys between "cool" and "not so cool" piles, on a green and an orange rectangle placing mats that were placed side-by-side. The set included four toys that were identical copies of the ones that the child had previously allocated in the toy distribution part of the study, as well as eight other toys (a balloon, a wooden building brick, a green plastic frog, a white rubber, a red paper flower, a green rubber band, a yellow car card, and a colorful spring). After sorting the toys into the two groups, the child was asked to rank-order them in each group. This was accomplished first for the "cool" pile, and then for the "not so cool" pile, by choosing first the "best" toy in the group, then the "next best" toy, then the "next best" toy, and so on. At the end of the toy ranking task, children were thanked for their participation and given the toys they had chosen to keep for themselves in the toy distribution task.

Questionnaire for Parents

One of the child's parents completed a questionnaire (see Appendix 3 in Supplementary Material) consisting of 19 items distributed between two dimensions: general information about the family structure, age, ethnicity, and educational level of the parents, and general information regarding the child's life conditions (i.e., nutrition, number of rooms of the house, home utilities, means of transport to reach school, family monthly household income). The estimated time to complete the questionnaire was 5–7 min. Questionnaires were completed with the help of the experimenter when the parent was illiterate.

Calculating the Transfer Score

The data about each child's preferences collected in the toy ranking task allowed us to calculate a "transfer score" that measures the relative value of the toys allocated to each person for each child. The transfer score for each child was calculated in two steps. First, we assigned each of the 12 toys a value based on the child's rank-ordering of the toys. The best toy got 12, the

second 11, etc. Thus, the value is 13 minus the rank. We then calculated the child's transfer score by adding the values of any toys transferred to the other child and subtracting the values of any toys kept for self. For example, here is the calculation for a child who had ranked the ping-pong ball third and allocated it to self, ranked the whistle fourth and allocated it to self, ranked the pencil seventh and allocated it to other, and ranked the sticker tenth and allocated it to other (see **Figure 2**): score = $-(13 - 3) - (13 - 4) + (13 - 7) + (13 - 10) = -10 - 9 + 6 + 3 = -10$. A negative transfer score indicates taking an advantage for oneself, a positive score indicates giving an advantage to the other child, and a score of 0 means a perfectly equal distribution.

RESULTS

Comparisons between Schools

We first compared the environments experienced by children from School A and School B based on the data provided by the children's parents in the Parental Questionnaire (Supplementary Material).

Distance from School

Children from School A and School B lived at a similar small distance (in minutes) from their school [School A: $M = 11.47$, $SD = 6.41$; School B: $M = 10.95$, $SD = 5.98$, independent t -test: $t(36) = 0.26$, $p = 0.80$].

Family Structure

Children from School B had younger parents [School A: $M = 34.8$, $SD = 4.5$; School B: $M = 30$, $SD = 4.12$, independent t -test: $t(36) = 3.38$, $p < 0.002$] and more siblings [School A: $M = 0.47$, $SD = 0.61$; School B: $M = 1.8$, $SD = 1$, independent t -test: $t(36) = 4.95$, $p < 0.001$]. No parents in School A were divorced or separated while 32% of the parents in School B were. Parents in School A were significantly less likely to be separated or divorced than parents in School B (Fisher's t -test: $p = 0.020$, odds ratio, OR = 4.95).

Nutrition and Material Living Conditions

Children in School B lived in more crowded houses, with more people per room [School A: $M = 1.84$, $SD = 0.66$; School B: $M = 2.27$, $SD = 0.46$, independent t -test: $t(36) = 2.33$, $p < 0.026$]. All children in School A ate meat at least once a week, while all children from School B ate meat at most once a month. All parents in School A declared having both electricity and a washing machine at home, while 79% of the parents in School B declared having no electricity and no washing machine at home (Fisher's t -test: $p < 0.001$, OR = 21.59). All parents in School A owned a refrigerator, while 32% of children in School B had no refrigerator at home (Fisher's t -test: $p = 0.020$, OR = 4.95).

Employment

Mothers in School A were significantly more likely to have a job at the time of the experiment (89%) than mothers in School B,



FIGURE 2 | Sample of toy ranking. The star (right) is ranked highest and the wooden block (left) is ranked lowest.

who all declared being a stay-at-home mother (Fisher's t -test: $p < 0.001$, $OR = 27.25$). All fathers in School A were employed, ran a business or worked as a free-lancer, while 84% of the fathers in School B declared being unemployed. Fathers in School A were significantly more likely to work than fathers in School B (Fisher's t -test: $p < 0.001$, $OR = 24.29$).

Parental Education Level

Mothers in School A were significantly more likely to have at least a high-school level of education (100%) than mothers in School B (37%) (Fisher's t -test: $p < 0.001$, $OR = 14.74$). Fathers in School A were also more likely to have at least a high-school level of education (100%) than fathers in School B (32%) (Fisher's t -test: $p < 0.001$, $OR = 27.25$).

Overall, this descriptive analysis confirms that children from School A and School B live in drastically different environments in terms of deprivation: children in School B have younger parents, with a lower education level, a lower combined income, and a higher chance of being unemployed. Children in School B also have more siblings, they live in more crowded houses, and they experience important material poverty and poorer access to food items like meat. Note that all the families in our sample ate meat at least occasionally, so that rare frequency of access to meat was not due to some families choosing to be vegetarian but rather due to meat being an expensive food item.

This descriptive analysis thus confirms that School A is situated in a middle SES neighborhood and School B is situated in a low SES neighborhood.

Toy Distribution Task

Transfer Score

We followed the same analysis plan as Sheskin et al. (2016), on which our experiment is based, and compared the transfer scores of the children from two schools. Because two outliers were identified in our dataset, we used robust linear regression to analyze the transfer scores (outliers: $N = 1$ in School A with transfer score of +20; $N = 1$ in School B with a transfer score of +23). The mean transfer score was -9.53 in children from School A ($N = 19$, $SD = 11.95$, range = -29 to $+20$) and -15.75 in children from School B ($N = 20$, $SD = 12.95$, range = -30 to $+23$). This analysis revealed that children from

the school situated in the very deprived neighborhood transferred significantly less to a stranger than children from the school situated in the less deprived neighborhood [$t(37) = -2.72$, $p = 0.010$; **Figure 3**].

We then ran a robust linear regression on children's transfer score, taking school, IQ, ethnicity, and gender as factors. In this analysis, the effect of school was still present as a trend [$t(34) = -1.84$, $p = 0.075$], while gender [$t(34) = 1.27$, $p = 0.210$], IQ [$t(34) = -0.99$, $p > 0.250$], and self-declared ethnicity [$t(34) = 1.06$, $p > 0.250$] were not significant predictors of children's transfer scores in this regression. Importantly, even if the two schools significantly differed in their ethnic compositions [$\chi^2(1, N = 39) = 11.29$, $p < 0.001$] and in their mean IQ [$t(37) = 2.16$, $p = 0.038$], no multicollinearity problem was evidenced in the regression taking school, gender, self-declared ethnicity, and IQ as factors (maximal variance inflation factor = 1.77, see Hair et al., 2006). In summary, even after controlling for IQ, gender and self-declared ethnicity, the most deprived neighborhood was associated with lower transfer score than the less deprived neighborhood.

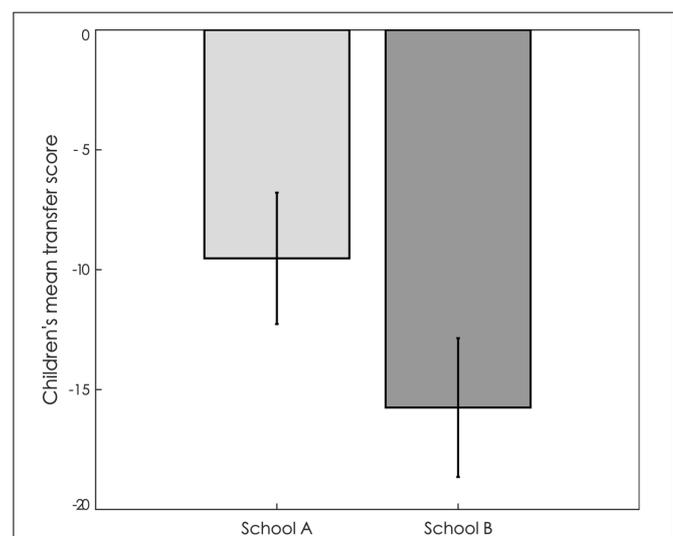


FIGURE 3 | Mean transfer score in children from School A and School B. Error bars represent standard errors to the mean.

Average Transfer Score

The two schools differed in their ranking of one of the four toys used in the task (the whistle). Average ranking of the whistle was lower in School A than in School B [$t(37) = -2.24, p = 0.031$; no significant difference for the other three toys: all $t(37) < 0.72$, all $p > 0.250$]. It is possible that this difference affects the comparison of the transfer score between the two schools. To rule out this possibility, we used the average value of each toy to compute an *average transfer score* for each child. This analysis is less reflective of individual differences in toy valuation, but ensures that significant differences between School A and School B are due to differences in toy distribution rather than toy ranking (since the same average ranking is used across all children in the study).

The conclusions from this additional analysis are identical to the conclusions from our pre-planned analysis. Using the average values of the toys (arrow card: $M = 1.97, SD = 1.40$; ping-pong ball: $M = 7.26, SD = 2.27$; pencil: $M = 5.21, SD = 2.18$; whistle: $M = 7.67, SD = 2.35$), we found that the mean average transfer score was -10.12 in children from School A ($N = 19, SD = 11.84, \text{range} = -22.10 \text{ to } +22.10$) and -15.84 in children from School B ($N = 20, SD = 12.23, \text{range} = -22.10 \text{ to } +22.10$). A robust linear regression revealed that the Average transfer scores of children from School B were significantly lower than those from School A [$t(37) = -5.33, p < 0.001$]. Importantly, as for the individual transfer score, this effect was still present as a trend after controlling for children's IQ, self-reported ethnicity and gender [$t(34) = -1.71, p = 0.096$]. Children from the school situated in the very deprived neighborhood thus transferred less to a stranger than children from the school situated in the less deprived neighborhood, whether or not we used individual value scores or the same value for each child.

Number of Toys Given

It is also possible to supplement our pre-planned analysis of transfer scores (the defining feature of the Quality Dictator Game), with a more standard analysis of how many toys were kept and how many were transferred. Thus, we compared the average number of toys given away by children from School A ($M = 1.37, SD = 1.16, \text{range} = 0-4$) and from School B ($M = 0.65, SD = 1.09, \text{range} = 0-4$). A robust linear regression on the number of toys given revealed that children from School A gave more toys on average than children from School B [$t(37) = -2.12, p = 0.041$]. Therefore, this more standard analysis confirmed the results obtained in the previous analyses.

DISCUSSION

Consistent with what Nettle et al. (2011) found with English adults and what Benenson et al. (2007) found with English 9-year-olds, we showed that variation in deprivation within a single city in a developing country influences the prosocial behavior of 6- to 7-year-olds toward an anonymous stranger. The relationship between neighborhood deprivation and children's prosociality

toward a stranger held when controlling for children's IQ, self-declared ethnicity, and gender.

Furthermore, two *post hoc* analyses converge with the pre-planned analyses. First, calculating children's transfer scores based on the average ranking of the toys did not change the results, suggesting our results were not driven by children valuing the toys they distributed in the toy distribution task differently when ranking them later among others (e.g., because of an endowment effect). Second, a comparison of the number of toys given showed a trend toward children from School A giving more toys than children from School B, showing that children from the middle SES neighborhood gave away more toys than children from the more deprived neighborhood.

Our study contributes to a better understanding of the impact of a harsh social environment on the development of prosocial behaviors toward strangers. However, it is important to note that the effects of environmental harshness on prosociality may not be linear. Both our results and the results of Benenson et al. (2007) are based on children living in low- and middle-SES environments; different results might be found when comparing middle- and high-SES environments. In fact, recent results by Miller et al. (2015) suggest that upper-SES children may behave *less* prosocially than middle-SES children. In their study, 4-year-old children from middle-SES backgrounds shared more of 10 tokens with a sick child than did high-SES children. One candidate underlying mechanism could be that competitiveness is higher in upper- vs. middle-SES children (e.g., Knight and Kagan, 1977; for a similar argument in teenagers, Buunk et al., 2013). Another possibility is that the impact of the child's social environment on prosocial motivation depends on the context. Children living in a harsh social environment might thus behave less prosocially in some contexts and more prosocially in others. In line with this idea, the impact of social status on prosocial behavior varies in adults. For instance, *noblesse oblige* can lead higher status people to behave more prosocially to defend their status, and higher resources associated with higher status can make prosocial behaviors less costly for the individual. Yet, being high status also gives leverage over others and decreases the costs of behaving less prosocially so that depending on the context, high status will lead to lower or higher prosociality (for a review, see Kafashan et al., 2014).

Another possible explanation of Miller et al.'s (2015) results is that the recipient of their dictator game was described as "a sick child" rather than an unknown child (as in our experiment and Benenson et al.'s 2007 study). It could be the case that SES correlates positively with prosociality toward a stranger, but that the correlation goes in the opposite direction when the recipient is not anonymous, or when empathy is involved as in Miller et al. (2015). We know of no other study where the recipient elicits empathy, but in a study by Chen et al. (2013), 4-year-old children from rural China played a dictator game with four stickers with the recipient being a friend or an unknown child. Similar to Miller et al. (2015), lower SES children gave more stickers than higher SES children in the friend condition but not in the unknown condition. In both Benenson et al.'s (2007) and Chen et al.'s (2013) studies, SES did not impact prosocial behavior by the age of 4 when the recipient was unknown, but the results of

Chen et al. (2013) in the friend condition and of Miller et al. (2015) in the “sick child” condition suggest that lower SES may be associated with higher prosocial behavior in some cases. Future research could further study the development of the correlation between SES and prosocial behavior by systematically varying the identity of the recipient and by testing low, middle, and high SES children.

A number of limitations of our study should also be acknowledged: in particular, we only focused on very low and middle SES children; we had—as in previous studies—only one measure of prosocial motivation; and our sample size was too small to disentangle the effects of the different factors that together constitute a harsh social environment such as material factors (e.g., material poverty or parental income) vs. social factors (e.g., parenting style or aggressive interactions) or biological factors (e.g., toxins; see Duncan et al., 1994; Bradley and Corwyn, 2002; Evans and English, 2002; for a review: Brooks-Gunn and Duncan, 1997; Evans, 2004). It will therefore be important for future research to assess whether the effect of social background on prosociality toward strangers is robust and to test the impact of different social factors associated with harsh social environments. Further investigations would be particularly interesting to identify the pathways by which environmental harshness impacts prosociality. For instance, prosocial behavior may be directly influenced by children’s level of resources but also by the type of social interactions they are usually exposed to (e.g., competitive vs. cooperative ones) as well as by the development of their cognitive abilities.

Finally, the impact of childhood environmental harshness has been shown to extend into adulthood in other domains such as health and non-social cognition (see e.g., Case et al., 2005; Luo and Waite, 2005; Cohen et al., 2010). Future research may thus investigate whether childhood environmental harshness has a similarly long-lasting impact on prosociality, and if so, how childhood deprivation interacts with current levels of resources. This next step would be all the more informative that studies on the effect of current socio-economic status on prosociality in adults have yielded mixed results. Specifically, two articles by Piff and his colleagues have reported a positive association between low SES and prosociality (Piff et al., 2010, 2012). In these two sets of studies, American college students reporting a lower subjective SES behaved more prosocially in a dictator game (Piff et al., 2010), drivers of more expensive cars behaved less ethically than drivers of cheaper cars (Piff et al., 2012; replicated by Morling et al., 2014), and

male students of higher SES reported a higher likelihood of behaving unethically in hypothetical scenarios (Piff et al., 2012; replicated by Lyons et al., 2012 and by Konigsberg et al., 2013).

In sharp contrast, a growing number of studies, including large-scale cross cultural experiments, have found that deprivation has a negative impact on prosociality (see e.g., Wilson et al., 2009; Nettle et al., 2011; Holland et al., 2012; Silva and Mace, 2014, 2015; Gomes and McCullough, 2015). In a recently published set of eight studies analyzing large and representative international samples including several thousands of participants, Korndörfer et al. (2015) found positive effects of higher SES on prosociality: based on self-reports, higher SES individuals were more likely to make charitable donations and contribute a higher percentage of their family income to charity, to volunteer, and to be helpful. Higher SES individuals were also more trusting and trustworthy in an economic game when interacting with a stranger than lower SES individuals. Getting a better understanding of how different factors associated with childhood social environment impact children and adult prosociality will advance our understanding of the causes of the great variations we observe in prosocial behaviors and motivations, and possibly help us find ways to use this knowledge to promote them.

AUTHOR CONTRIBUTIONS

SL, MS, TT, NB, and CC contributed to the design of the study; TT did the data collection and anonymized the data; LS, SL, TT, MS, and CC contributed to the data analysis; and LS, SL, TT, MS, NB, and CC contributed to the writing of the paper.

ACKNOWLEDGMENT

This study was supported by the Institut d’Études Cognitives (ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01760/full#supplementary-material>

REFERENCES

- Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. *J. Theor. Biol.* 281, 47–55. doi: 10.1016/j.jtbi.2011.04.023
- Baumard, N., André, J. B., and Sperber, D. (2013). A mutualistic approach to morality: the evolution of fairness by partner choice. *Behav. Brain Sci.* 36, 59–78. doi: 10.1017/S0140525X11002202
- Baumard, N., and Chevallier, C. (2015). The nature and dynamics of world religions: a life-history approach. *Proc. R. Soc. B* 282:20151593. doi: 10.1098/rspb.2015.1593
- Benenson, J. F., Pascoe, J., and Radmore, N. (2007). Children’s altruistic behavior in the dictator game. *Evol. Hum. Behav.* 28, 168–175. doi: 10.1016/j.evolhumbehav.2006.10.003
- Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., et al. (2015). The ontogeny of fairness in seven societies. *Nature* 528, 258–261. doi: 10.1038/nature15703
- Bradley, R. H., and Corwyn, R. F. (2002). Socioeconomic status and child development. *Annu. Rev. Psychol.* 53, 371–399. doi: 10.1146/annurev.psych.53.100901.135233
- Brooks-Gunn, J., and Duncan, G. J. (1997). The effects of poverty on children. *Future Child* 7, 55–71. doi: 10.2307/1602387

- Buunk, A. P., Stulp, G., and Ormel, J. (2013). Parental social status and intrasexual competitiveness among adolescents. *Evol. Psychol.* 12, 1022–1037.
- Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *J. Health Econ.* 24, 365–389. doi: 10.1016/j.jhealeco.2004.09.008
- Chen, Y., Zhu, L., and Chen, Z. (2013). Family income affects children's altruistic behavior in the Dictator Game. *PLoS ONE* 8:e80419. doi: 10.1371/journal.pone.0080419
- Cohen, S., Janicki-Deverts, D., Chen, E., and Matthews, K. A. (2010). Childhood socioeconomic status and adult health. *Ann. N. Y. Acad. Sci.* 1186, 37–55. doi: 10.1111/j.1749-6632.2009.05334.x
- Duncan, G. J., Brooks-Gunn, J., and Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Dev.* 65, 296–318. doi: 10.2307/1131385
- Dunfield, K. A., O'Connell, L., Kuhlmeier, V. A., and Kelley, E. A. (2011). Examining the diversity of prosocial behaviour: helping, sharing, and comforting in infancy. *Infancy* 16, 227–247. doi: 10.1111/j.1532-7078.2010.00041.x
- Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91–119. doi: 10.1037/0033-2909.101.1.91
- Evans, G. W. (2004). The environment of childhood poverty. *Am. Psychol.* 59, 77–92. doi: 10.1037/0003-066X.59.2.77
- Evans, G. W., and English, K. (2002). The environment of poverty: multiple stressor exposure, psychophysiological stress, and socioemotional adjustment. *Child Dev.* 73, 1238–1248. doi: 10.1111/1467-8624.00469
- Gomes, C. M., and McCullough, M. E. (2015). The effects of implicit religious primes on dictator game allocations: a preregistered replication experiment. *J. Exp. Psychol. Gen.* 144, e94–e104. doi: 10.1037/xge0000027
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2006). *Multivariate Data Analysis*, Vol. 6. Upper Saddle River, NJ: Pearson Prentice Hall.
- Hamann, K., Warneken, F., Greenberg, J. R., and Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature* 476, 328–331. doi: 10.1038/nature10278
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science* 327, 1480–1484. doi: 10.1126/science.1182238
- Holland, J., Silva, A. S., and Mace, R. (2012). Lost letter measure of variation in altruistic behaviour in 20 neighbourhoods. *PLoS ONE* 7:e43294. doi: 10.1371/journal.pone.0043294
- House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., et al. (2013). Ontogeny of prosocial behavior across diverse societies. *Proc. Natl. Acad. Sci. U.S.A.* 110, 14586–14591. doi: 10.1073/pnas.1221217110
- Kafashan, S., Sparks, A., Griskevicius, V., and Barclay, P. (2014). "Prosocial behavior and social status," in *The Psychology of Social Status*, eds J. T. Cheng, J. L. Tracy, and C. Anderson (New York, NY: Springer), 139–158.
- Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: the endowment effect, loss aversion, and status quo bias. *J. Econ. Perspect.* 5, 193–206. doi: 10.1257/jep.5.1.193
- Knight, G. P., and Kagan, S. (1977). Development of prosocial and competitive behaviors in Anglo-American and Mexican-American children. *Child Dev.* 48, 1385–1394. doi: 10.2307/1128497
- Konigsberg, L., MacGregor, S., Johnson, M., Massey, A., and Daubman, K. (2013). *Higher SES Predicts Unethical Behavior in Men*. Available at: <http://www.PsychFileDrawer.org/replication.php?attempt=MTUw> (accessed February 23, 2013).
- Korndörfer, M., Egloff, B., and Schmukle, S. C. (2015). A large scale test of the effect of social class on prosocial behavior. *PLoS ONE* 10:e0133193. doi: 10.1371/journal.pone.0133193
- Luo, Y., and Waite, L. J. (2005). The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, S93–S101. doi: 10.1093/geronb/60.2.S93
- Lyons, J., Born, K., Pearsen, J., Fox, M., Collins, C., Trosch, E., et al. (2012). *Higher SES Predicts Unethical Decision Making*. Available at: <http://www.PsychFileDrawer.org/replication.php?attempt=MTMx> (accessed September 24, 2012).
- Mackintosh, N. J. (1998). *IQ and Human Intelligence*. New York, NY: Oxford University Press.
- Miller, J. G., Kahle, S., and Hastings, P. D. (2015). Roots and benefits of costly giving: children who are more altruistic have greater autonomic flexibility and less family wealth. *Psychol. Sci.* 26, 1038–1045. doi: 10.1177/0956797615578476
- Morling, B., Jacobs, T., Chasten, A., Katz, S., Brooks, E., Maurer, T., et al. (2014). *Car Status and Stopping for Pedestrians*. Available at: <http://www.PsychFileDrawer.org/replication.php?attempt=MTky> (accessed June 02, 2014).
- Nettle, D., Colléony, A., and Cockerill, M. (2011). Variation in cooperative behaviour within a single city. *PLoS ONE* 6:e26922. doi: 10.1371/journal.pone.0026922
- Piff, P. K., Kraus, M. W., Côté, S., Cheng, B. H., and Keltner, D. (2010). Having less, giving more: the influence of social class on prosocial behavior. *J. Pers. Soc. Psychol.* 99, 771–781. doi: 10.1037/a0020092
- Piff, P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., and Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4086–4091. doi: 10.1073/pnas.1118373109
- Raven, J., Raven, J. C., and Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.
- Sheskin, M., Bloom, P., and Wynn, K. (2014). Anti-equality: social comparison in young children. *Cognition* 130, 152–156. doi: 10.1016/j.cognition.2013.10.008
- Sheskin, M., Nadal, A., Croom, A., Mayer, T., Nissel, J., and Bloom, P. (2016). Some equalities are more equal than others: quality equality emerges later than numerical equality. *Child Dev.* 87, 1520–1528. doi: 10.1111/cdev.12544
- Silva, A. S., and Mace, R. (2014). Cooperation and conflict: field experiments in Northern Ireland. *Proc. R. Soc. B Biol. Sci.* 281:20141435. doi: 10.1098/rspb.2014.1435
- Silva, A. S., and Mace, R. (2015). Inter-group conflict and cooperation: field experiments before, during and after sectarian riots in Northern Ireland. *Front. Psychol.* 6:1790. doi: 10.3389/fpsyg.2015.01790
- Smith, C. E., Blake, P. R., and Harris, P. L. (2013). I should but I won't: why young children endorse norms of fair sharing but do not follow them. *PLoS ONE* 8:e59510. doi: 10.1371/journal.pone.0059510
- Warneken, F., and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science* 311, 1301–1303. doi: 10.1126/science.1121448
- Warneken, F., and Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends Cogn. Sci.* 13, 397–482. doi: 10.1016/j.tics.2009.06.008
- Wilson, D. S., O'Brien, D. T., and Sesma, A. (2009). Human prosociality from an evolutionary perspective: variation and correlations at a city-wide scale. *Evol. Hum. Behav.* 30, 190–200. doi: 10.1016/j.evolhumbehav.2008.12.002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Safra, Tecu, Lambert, Sheskin, Baumard and Chevallier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

related outcomes like substance use, accidental injury, and unintentional drowning (Steinberg 2013). All of these phenomena are particularly pronounced in adolescent males (Scheidt et al. 1995), in whom patterns of brain development are somewhat delayed relative to those of females (Lenroot & Giedd 2006), and who are also much more likely to engage in violent aggression (Snyder 2012).

The authors have made a strong case that rates of violent aggression rise with geographic proximity to the equator, which is used as a proxy for climate. But the model may require amending to incorporate the mediating or moderating role that the proportion of adolescents across regions may play in geographic variation in violent aggression.

As it happens, more equatorial countries also contain large – sometimes much larger – proportions of adolescents. Almost without exception, those nations with median ages less than the global median age of 29 are equatorial nations of Africa, Asia, and Central and South America. The youngest countries in the world include equatorially proximate African nations like Uganda, Niger, and Mali. By contrast, the oldest countries include Japan, Germany, Monaco, and other European and Asian countries closer to the poles (Central Intelligence Agency 2016b).

So-called “youth bulges” in countries proximal to the equator could explain increased violence in these countries without reference to life history theories or climate. It has been observed that the proportion of individuals within a society who are between the ages of 15 and 24 is predictive of the prevalence of various forms of violence in that society, including homicide, domestic armed conflict, terrorism, and rioting (Bricker & Foley 2013; Mesquida & Wiener 1999; Urdal 2006). In the United States, a significant proportion of changes in violent crime over time can be explained by fluctuations in the proportion of adolescents (Phillips 2006). Globally, the relationship between violence and the proportion of a country’s population composed of adolescents has also been found across several investigations (Cincotta & Leahy 2011; Pampel & Gartner 1995; Urdal 2006).

Van Lange and colleagues might argue that countries like Guatemala, Belize, and Honduras are among the most violent countries in the world because of their equatorial climates, which result in the population of these regions adopting relatively fast life history strategies, characterized by “short-term planning, greater risk taking, a focus on immediate gratification for short-term benefits, and more aggression” (sect. 3.1, para 3). But one could just as easily argue that the reason these quintessentially adolescent traits are pervasive in these countries is that their populations are disproportionately composed of adolescents. These are among the world’s youngest countries, and are the three youngest nations in the Americas. Nearly a quarter of the population of Guatemala, for example, is between the ages of 15 and 24 (Central Intelligence Agency 2016b).

It is possible, then, that the patterns the authors have observed do not reflect climate-induced variation in life history strategies, but are instead an artifact of geographic fluctuations in the proportion of youths, owing to reasons that are unrelated to climate. Alternately, it is possible that fluctuations in the proportion of youths could be incorporated into the CLASH model. Youth bulges are thought to emerge during the stage of a nation’s development when infant mortality has been successfully reduced, yet fertility rates remain relatively high (Bricker & Foley 2013), but no generally accepted explanation exists for why youth bulges currently cluster around the equator.

Any attempt to incorporate the relationship between youth and violence into the CLASH model should reflect the fact that the relationship between youth bulges and violent aggression may not be a simple one. Variables like access to education and jobs are critical influences on the behavior of youths within a society (Bricker & Foley 2013). Likewise, the effect of youth bulges on violent aggression may be mitigated by protective cultural factors like collectivism, which can transform large youth populations into civic opportunities (Pampel & Gartner 1995).

Understanding the role of adolescence in societal variation in aggression therefore requires considering not only how many adolescents a society contains, but also how those adolescents are faring: Are they civically engaged? Educated? Impoverished? Optimistic about their future prospects (Bricker & Foley 2013; Hart et al. 2004; Pampel & Gartner 1995)?

The essential fact remains that a model of cultural variation in violent aggression that does not consider the role of adolescence remains an incomplete model.

Climate is not a good candidate to account for variations in aggression and violence across space and time

doi:10.1017/S0140525X16001060, e91

Hugo Mell,^a Lou Safra,^a Nicolas Baumard,^b and Pierre O. Jacquet^{a,b}

^a*Evolution and Social Cognition Group, Laboratoire de Neurosciences Cognitives (LNC), Département d’Etudes Cognitives, INSERM U960, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France;*

^b*Evolution and Social Cognition Group, Institut Jean Nicod, Département d’Etudes Cognitives, CNRS UMR8129, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France.*

mell@magbio.ens.fr lou.safra@gmail.com

nbaumard@gmail.com pierre.jacquet@inserm.fr

http://iec-esc.ens.fr/

Abstract: We agree with Van Lange et al. that climate is likely to affect individuals’ social behavior in many ways. However, we suspect that its impact on physiology and psychology is so remote that its predictive power disintegrates almost completely through the causal chain underlying aggression and violence.

Using data about current and past patterns of aggression and human sociality, we show that the causal role of climate vanishes once one switches from a worldwide perspective to a more local one, and that it becomes quasi-irrelevant once a historical dimension is considered. Evolutionary models in biology provide explanations of variations in traits that are generalizable across both space and time. We believe that this criterion of relevance is, however, not met by CLASH.

We start our demonstration by testing whether climate predicts interpersonal violence during a restricted period within geographic Europe (Fig. 1A,B), an area that is similar in size and culture to the United States. Except for Russia, all major European countries with available climatic and homicide data for the 2008–2012 period were included (35 countries, sources: World Bank (The World Bank Group 2016a; United Nations Office on Drugs and Crimes [UNODC] 2016). Following methodological recommendations from studies that inspired the target article (Burke et al. 2015; Hsiang et al. 2013), we ran a series of correlations between interpersonal violence (measured by the homicide rate [Burke et al. 2015]) and either yearly average temperature or seasonal variations in temperature (i.e., the difference between the average temperature for the three summer months and the three winter months). We included only countries where CLASH was applicable (average year temperature under 24°C). Contrary to CLASH’s predictions, both year-by-year and period-wise analyses revealed that interpersonal violence did not vary with temperature (all *r* values < .12, all *t*(34) values < 0.67, all *p* values > .25) (Fig. 1A) and, more surprisingly, increased with seasonality (all *r* values > .39, all *t*(34) values > 2.49, all *p* values < .018) (Fig. 1B).

We then tested whether climatic variables predicted homicide rates in elapsed time periods (Fig. 1C,D) using historical records from two geographically distant countries of the temperate zone: Japan (1924–2004 [Statistics Bureau, Ministry of

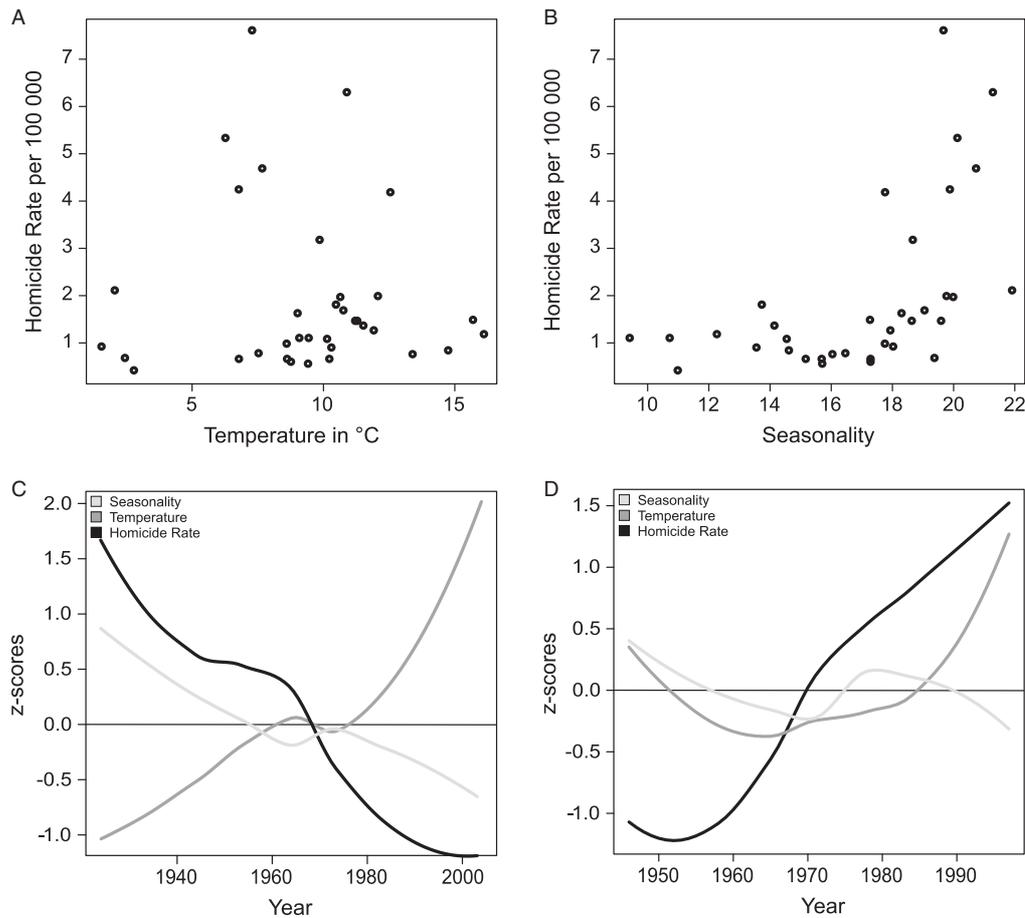


Figure 1 (Mell et al.). (A,B) Homicide rates averaged for the 2008–2012 period for countries (N=35) constituting geographic Europe, as a function of (A) temperature and (B) seasonality. (C,D) Year-by-year evolution of homicide rate, temperature, and seasonal variation for (C) Japan and (D) Great Britain.

Internal Affairs and Communications 2016a; 2016b; The World Bank Group 2016a]) and Great Britain (1951–1999 [Richards 1999; The World Bank Group 2016a]). For representational purposes, all three variables were expressed in units of standard deviations of their respective means over the whole period (*z*-score transformed). Figure 1C and D represent the trend for each variable across the time record, with ranges and standard deviations differing between Great Britain (homicide rate per 100,000: 0.58–2.06, SD = 0.42; temperature: 7.55°C–9.52°C, SD = 0.47; seasonality: 7.54°C–14.01°C, SD = 1.28) and Japan (homicide rate per 100,000: 0.97–4.14, SD = 0.98; temperature: 10.11°C–12.71°C, SD = 0.56; seasonality: 17.5°C–22.95°C, SD = 1.09). The impact of temperature and seasonal variations on homicide rates was the strict opposite of CLASH’s predictions for Japan (average temperature: $r = -.52$, $t(79) = -5.43$, $p < .001$; seasonality: $r = 1.99$, $p = .049$) (Fig. 1C), whereas no effect was observed for Great Britain (both r values $< .18$, both $t(45)$ values < 1.26 , both p values $> .215$) (Fig. 1D). Figure 1D illustrates the importance of considering historical data to avoid spurious correlations when trying to find determinants of trait variations: during the 1990s, homicide rate and temperature positively covaried in Great Britain, but it is seen by going back further in time that the rise in homicides preceded the temperature increase and that the relationship was actually reversed in the 1950s.

One could argue that testing countries with relatively homogeneous temperate climates is inappropriate because small climatic variance might not include the critical threshold at which more dramatic levels of violence occur. If true, then deviations from

the mean could be pure noise. However, we believe that this is not the case. Figure 2 indeed illustrates the difference in murder rates measured for the year 2013 (Federal Bureau of Investigation [FBI] 2013) in all 50 states of the United States and in the five boroughs of New York City (Pediacies NYC) and shows that variations observed at a very local scale (city) can be of similar magnitude as variations observed at a very global scale (continent). We doubt that it is reasonable to posit that climate differences act as a major predictor, at the expense of other explanatory variables such as, for example, differences in income.

In addition to these empirical arguments, we raise a more fundamental concern: CLASH cannot satisfyingly account for major transitions in the evolution of human sociality such as, for instance, the replacement of asocial religions by prosocial ones. Recent work indeed demonstrates that the best explanatory factor of this phenomenon is an increase in affluence (energy capture per capita, urbanization rate, population growth), a variable highly predictive of individuals’ level of resources (Baumard et al. 2015). This is in line with the life history framework, which predicts that individuals enjoying higher levels of resources engage in slower life strategies that are characterized by high investments in long-term goals, including cooperative goals. The early emergence of belief systems promoting prosociality can thus be better understood as the consequence of historical changes in the distribution of resources (Baumard & Chevallier 2015), rather than climate. It becomes especially striking when one considers that prosocial religions appeared in different

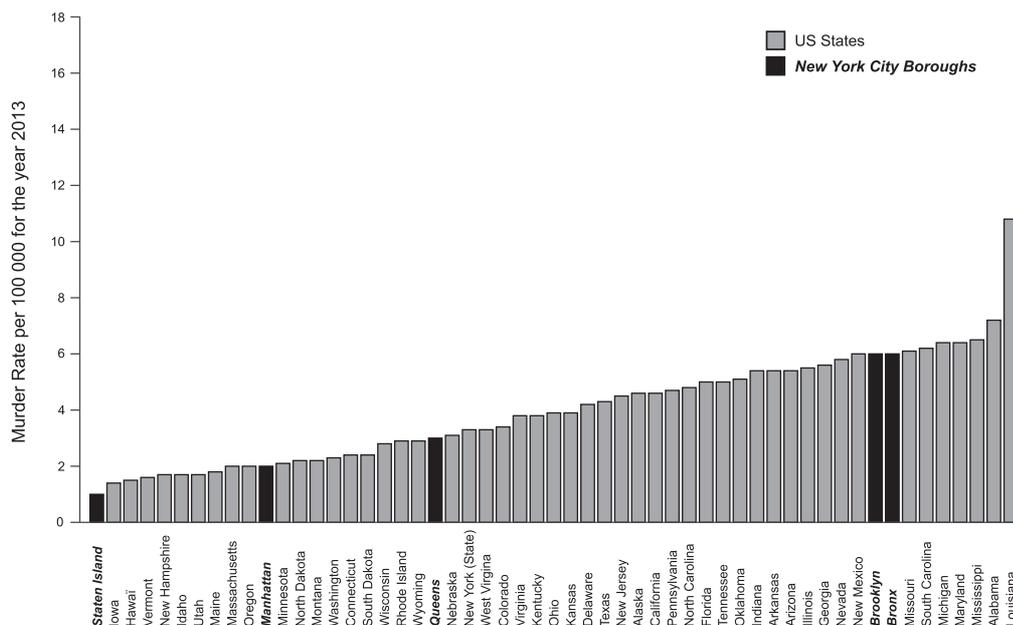


Figure 2 (Mell et al.). Murder rates measured for the year 2013 in all 50 states of the United States and in the five boroughs of New York City.

civilizations located in arid, semi-arid, and tropical zones of the world (Eastern Mediterranean, Mesopotamia, Northern India) long before they emerged in more temperate areas.

In sum, between-individual differences in life strategies – whose acceleration *eventually* leads to greater violence – are more likely to depend on ecological dimensions whose fluctuations matter more than climate for survival and reproduction in complex social worlds. All things being equal, climate differences ought to be part of the general explanation, but we doubt that they should constitute the core feature of evolutionary models of aggressive and violent behaviors.

The paradoxical effect of climate on time perspective considering resource accumulation

doi:10.1017/S0140525X16001072, e92

Gábor Orosz,^{a,b} Philip G. Zimbardo,^c Beáta Bóthe,^{a,d} and István Tóth-Király^{a,d}

^aInstitute of Psychology, Eötvös Loránd University, Budapest, Hungary;

^bInstitute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary;

^cDepartment of Psychology, Stanford University, Stanford, California;

^dDoctoral School of Psychology, Eötvös Loránd University, Budapest, Hungary.

orosz.gabor@ppk.elte.hu drzimbardo@gmail.com

bothe.beata@ppk.elte.hu toth-kiraly.istvan@ppk.elte.hu

https://www.researchgate.net/profile/Orosz_Gabor

<https://psychology.stanford.edu/zimbardo>

https://www.researchgate.net/profile/Beata_Bothe

https://www.researchgate.net/profile/Istvan_Toth-Kiraly

Abstract: Considering purely climate, southern countries are less harsh and more predictable than northern countries. From a historical perspective, freezing winters resulting in fewer available resources contribute to the development of strong future orientation. The paradox is that future orientation contributes to accumulation of resources in the long run, making individuals' immediate living conditions less harsh, leading to slower life strategies.

Predictability and climate can be seen from different points of view. On the basis of Life History Theory (e.g., Hill 1993; Kaplan & Gangestad 2005), it is claimed that fast life strategies (related to reproducing earlier, having higher mortality and morbidity rates, having higher levels of violence) are adopted when life events are unpredictable and harsh (Ellis et al. 2009; Griskevicius et al. 2011), and slow life strategies (the opposite) are adopted when life is more predictable and less harsh. Van Lange et al. connect this to climate and claim that in warmer areas close to the equator, life is more unpredictable and harsher (from the perspective of climate as a result of viruses and natural disasters). Contrary to this argument, if one focuses on climate and its seasonal variation close to the equator, the climate is less extreme and there is less seasonal variation. From the perspective of climate arises the question: What can be more predictable and less harsh than constant warmth with little variation? Contrary to the authors' claim, it is plausible that areas close to the equator are more predictable, and considering purely climate, seasonal changes in areas farther from the equator can result in harsh life conditions such as freezing winters, unexpected summer droughts, and floods.

From a historical perspective, in those regions where the weather was comfortably warm all year, crops could be harvested twice a year and food and shelter were available all year. However, in the past, a second harvest in the winter was not possible in northern regions and it was more difficult to find shelter, which made living conditions harsher and less predictable. In short, they had fewer exposed resources during a certain part of the year. These conditions could have motivated northerners to become more future oriented, that is, to think about the forthcoming winter and its possible negative consequences and, thus, to accumulate and save resources (Ashkanasy et al. 2004). These northerners were forced to accumulate resources to cover the periods when resources were scarce. For them, future orientation was the key to survival (Zimbardo & Boyd 2008). Throughout history, this accumulating behavior driven by future orientation allowed northern societies to reach a higher level of economic development, whereas in southern regions, where resources were available all year long, societies were less focused on the future.

As accumulation of resources reached a point when there were more than enough resources for one harsh winter, the perceived availability of resources changed. If the pantry is full all year, the

RESEARCH ARTICLE

Measuring Social Motivation Using Signal Detection and Reward Responsiveness

Coralie Chevallier^{1,2*}, Natasha Tonge¹, Lou Safra², David Kahn^{1,3}, Gregor Kohls⁴, Judith Miller^{1,3}, Robert T. Schultz^{1,3,5}

1 Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, PA, United States of America, **2** Laboratoire des Neurosciences Cognitives (LNC), INSERM U960, ENS-PSL, Paris, France, **3** Center for Cognitive Neuroscience, University of Pennsylvania, Philadelphia, PA, United States of America, **4** Child Neuropsychology Section, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany, **5** Departments of Pediatrics and Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States of America

* coralie.chevallier@ens.fr



CrossMark
click for updates

Abstract

Background

Recent trends in psychiatry have emphasized the need for a shift from categorical to dimensional approaches. Of critical importance to this transformation is the availability of tools to objectively quantify behaviors dimensionally. The present study focuses on social motivation, a dimension of behavior that is central to a range of psychiatric conditions but for which a particularly small number of assays currently exist.

Methods

In Study 1 (N = 48), healthy adults completed a monetary reward task and a social reward task, followed by completion of the Chapman Physical and Social Anhedonia Scales. In Study 2 (N = 26), an independent sample was recruited to assess the robustness of Study 1's findings.

Results

The reward tasks were analyzed using signal detection theory to quantify how much reward cues bias participants' responses. In both Study 1 and Study 2, social anhedonia scores were negatively correlated with change in response bias in the social reward task but not in the monetary reward task. A median split on social anhedonia scores confirmed that participants with high social anhedonia showed less change in response bias in the social reward task compared to participants with low social anhedonia.

Conclusions

This study confirms that social anhedonia selectively affects how much an individual changes their behavior based on the presence of socially rewarding cues and establishes a tool to quantify social reward responsiveness dimensionally.

OPEN ACCESS

Citation: Chevallier C, Tonge N, Safra L, Kahn D, Kohls G, Miller J, et al. (2016) Measuring Social Motivation Using Signal Detection and Reward Responsiveness. *PLoS ONE* 11(12): e0167024. doi:10.1371/journal.pone.0167024

Editor: Mathias Pessiglione, "INSERM", FRANCE

Received: January 4, 2016

Accepted: November 8, 2016

Published: December 1, 2016

Copyright: © 2016 Chevallier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data, analysis scripts and experimental scripts are available on OSF: <https://osf.io/9uuuc/>.

Funding: This study was funded by a grant from the Pennsylvania Department of Health (SAP # 4100047863) awarded to RTS. CC and LS were also supported by Fondation Roger de Spoelberch, ANR-11-0001-02 PSL*, ANR-10-LABX-0087. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: RTS has received speaking fees from Shire Pharmaceuticals, and consulting

fees from Seaside Therapeutics and Roche Pharmaceuticals. None of these activities overlap with the study described in this paper.

Background

One of the great challenges in understanding the biological cause of psychiatric disorders has been that diagnostic categories based on clinical consensus do not align with findings from neuroscience and genetics [1]. In response to this challenge, it has become clear that it is necessary to go beyond traditional diagnostic boundaries by characterizing individual phenotypes dimensionally [2]. One dimension of behavior that is particularly well-suited to dimensional approaches is social motivation. Social motivation can be described as a set of biological mechanisms biasing the individual to preferentially orient to the social world and to treat social interactions as rewarding. Social motivation has been identified as a relevant behavioral dimension for a number of disorders including schizophrenia-spectrum disorders (e.g., [3]), anorexia nervosa (e.g., [4]), depression (e.g., [5]), psychopathy (e.g., [6]), and, perhaps most notably, autism spectrum disorders (ASDs) (e.g., [7]). However, the presence or absence of social motivation deficits does not directly align with any of these diagnostic categories, which suggests that social motivation is best construed in a dimensional framework cutting across multiple conditions.

Research investigating social motivation has been hindered by the paucity of tools to reliably quantify this construct. Most researchers aiming to quantify social motivation as a dimension indeed rely on a handful of self-reports instruments, such as the Chapman Social Anhedonia Scale [8] or the Affiliative Tendency Scale [9], which provide a useful first-pass evaluation of social motivation in typical adults but might be biased in both psychiatric and control populations. Individuals with a psychiatric condition may indeed lack self-reflection skills and struggle to accurately report on their own feelings [10] and control participants may be susceptible to social desirability effects leading them to overestimate their social motivation [11]. Another concern associated with existing self-report instruments is that they do not specifically target social reward responsiveness and instead combine various aspects of social motivation that might be best construed as distinct traits (e.g., social seeking, social attention, reputation management, gregarious instincts, social reward responsiveness, etc.).

In this paper, our goal is to use signal detection theory to create an objective tool tapping social reward responsiveness specifically. With this goal in mind, we adapt a task developed by Pizzagalli et al. to assess non-social reward responsiveness and its relationship to depression [12]. Specifically, previous work on non-social reward responsiveness has demonstrated that unequal frequency of reward is associated with a preference for the response that is associated with the more frequent reward [13,14] and that this response bias constitutes a reliable proxy of reward responsiveness [12]. Pizzagalli et al. (2005) then demonstrated that participants with elevated levels of depression symptoms showed a blunted response bias towards frequently rewarded stimuli, thereby demonstrating that their behavior was less modulated by reinforcement than individuals with low levels of depression symptoms. Here, we present two studies that extend this work to social reward responsiveness: Study 1 includes a monetary rewards task, which essentially replicates Pizzagalli's task, and our social rewards task. We predicted that both a high frequency of social and monetary rewards would bias participants to answer faster and more accurately to the rich stimulus. More importantly, we predicted that individual differences in this response bias would correlate with self-reported social anhedonia but not with self-reported physical anhedonia, which refers to pleasures associated to non-social sensory experiences (e.g., eating, touching pleasant material, moving, experiencing smells and sounds, etc.). In Study 2, we ran a direct replication of the social task with an unrelated sample of healthy participants.

Study 1: Designing a dimensional tool to assess social motivation

Methods

Ethics Statement. This study was reviewed and approved by the Children's Hospital of Philadelphia's Institutional Review Board (IRB 11–008173). After complete description of the study to the subjects, written informed consent was obtained.

Materials. The monetary and social reward tasks were presented on a 22-inch widescreen monitor using E-prime. In the Money experiment, monetary rewards were presented as the plain text “+ 5 cents”, which represented a real 5-cent reward that participants were told would be added to their total reimbursement. In the social experiment, social rewards were in the form of a silent, full-color video clip of an actor providing approval by simultaneously smiling, nodding, and showing a thumbs-up gesture [15]. The rewards were provided in response to correct identification of a line appearing in the center of a circle as being short (11.5mm) or long (13mm).

Following the computer tasks, participants were left alone to complete the Beck Depression Inventory-II [16] and the Chapman Social and Physical Anhedonia Scales [8]. The BDI-II was used to screen out clinically depressed participants who might not have been identified during our phone screening. The Chapman scales were included in order to measure self-reported anhedonia in response to social and physical situations. The Social Anhedonia Scale measures how much participants find social interactions to be rewarding (e.g., the interpersonal pleasure of being with people, talking, exchanging expressions of feelings, doing things with others, competing, loving, and interacting in multiple other ways). Note that due to a computer error, only the first 32 items were administered in Study 1. 12 participants were re-invited to take the full version of the questionnaire within 4 months of their initial visit. Scores on the 32-item version and on the full 40-item version were highly correlated, $r = .86, p = .0002$. The Physical Anhedonia Scale measures how much participants find physical situations to be rewarding (e.g., physical pleasures linked to eating, touching, feeling, temperature, movement, smell and sound). All experimental scripts, analysis scripts, test materials and raw data are available for academic purposes on the Open Science Framework.

Participants. Forty-eight participants (27 female, 21 male, mean age: 24.16 +/- 3.28 years) were recruited from the Philadelphia area using Craigslist and paper flyers. Interested volunteers were screened over the phone to determine eligibility. Individuals who reported an active Axis I disorder or current psychotropic medication use were not invited to participate in the study. All participants received compensation for their time (\$10 per hour) and travel costs, as well as a variable monetary reward depending on task performance (\$5–6). Three participants with a task accuracy of 2 SD below the mean and an additional two participants above the clinical cutoff on the BDI-II were excluded from final analysis, leaving a final sample of 44 participants (19 males, 25 females).

Design and Procedure. Each participant completed both the Social experiment and the Money experiment, order counterbalanced across participants. Each experiment lasted approximately 25 minutes separated by a 10-minute categorization task that was unrelated to the current study. Participants were tested individually and stayed alone in the testing room during the entire duration of the experiment. The experimenter was only present to introduce each experiment and provide feedback during the training phase. The instructions were read aloud to the participants as they were viewing them simultaneously on the screen. Participants were told that their task would be to classify a line as either short or long by pressing the corresponding key and that feedback for correct responses would only occur some of the time. Depending on the experiment, participants were told that feedback for correct responses would be a “thumbs up” video clip (social reward task) or the text “+ 5 cents” (monetary

reward task). The training phase consisted of five practice trials during which no reward and no feedback were provided.

Each experiment consisted of 300 trials separated into 100-trial blocks. Each trial began with the presentation of a fixation cross (500ms), followed by an empty circle (500ms). The short or the long line was then flashed within the circle (100ms) then disappeared to show a plain black screen during which participants could submit their response (see Fig 1). An equal number of short and long lines were presented within each block. Short and long lines were presented in a random order. Participants were given an infinite amount of time to indicate their response using ‘e’ or ‘p’ on the keyboard. Responses were followed by 1750ms of either a blank screen, or the social or monetary reward described above.

An asymmetrical reinforcement ratio was introduced such that one type of line (i.e. short or long) was rewarded more often. The line type that was rewarded more frequently is referred to as the “rich stimulus” and the line type that was rewarded least is referred to as the “lean stimulus”. Within each 100-trial block, participants received positive feedback on 30 trials for accurately identifying the rich line, and on 10 trials for accurately identifying the lean line. An algorithm was programmed so that reward feedback followed a pseudorandom schedule specifying which specific trials were to be rewarded for correct identifications [12,17]. If a participant failed to correctly identify the line on a trial programmed to get a reward, the reward was delayed until a line of the same type appeared and was correctly identified. The long line was randomly assigned to being the rich or the lean stimulus for each task and participant. Participants earned approximately \$6 for the entire study.

Data Analyses. Our analysis plan was based on Pizzagalli et al. (2005) who relies on signal detection theory [18] to quantify reward responsiveness. In signal detection tasks, participants’ responses are classified into one of four categories: hits, false alarms, misses and correct

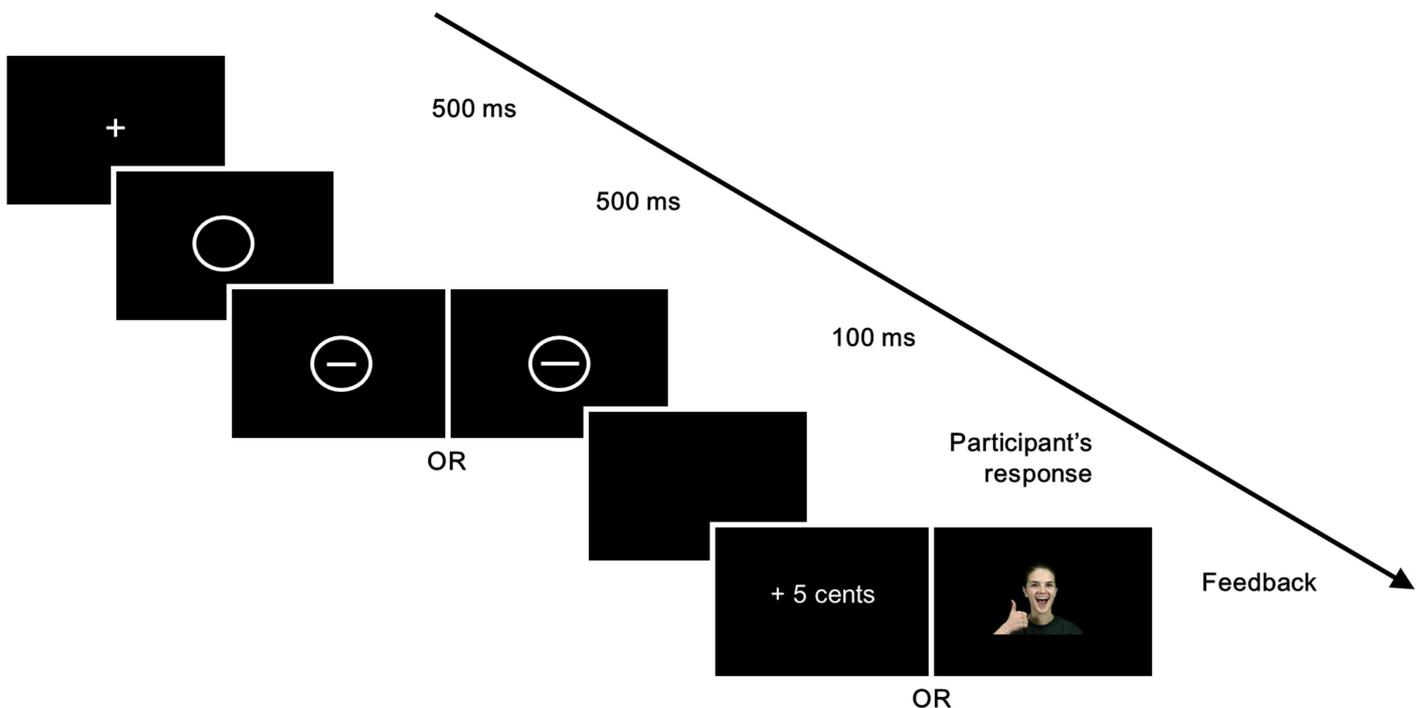


Fig 1. Schematic representation of the tasks. A fixation cross appears for 500ms, followed by an empty circle. A short or a long line is then flashed inside the circle for 100ms. Participants have an infinite amount of time to respond before they receive a reward for some of their correct responses.

doi:10.1371/journal.pone.0167024.g001

rejections. In the experimental condition where the long line is the rich stimulus, hits refer to correct identifications of the long line as long, false alarms refer to identifications of the short line as long, misses refer to identification of the long line as short and correct rejections refer to identification of the short line as short.

Two conceptually distinct metrics can then be computed: 1) ‘discriminability’, which refers to participants’ ability to differentiate the two stimuli, and will be high if there are more hits and correct rejections than misses and false alarms; 2) ‘response bias’, which refers to the likelihood of giving one response more frequently than another and will be positive if there are more hits and false alarms than misses and correct rejections. High discriminability scores indicate increased ease of telling the two target stimuli apart and thus works as an index of task difficulty. Larger positive response bias scores indicate an increased tendency to identify the stimulus line as the more rewarded line and, simultaneously, a decreased tendency to identify the stimulus as the less rewarded line. Put another way, a large response bias indicates that an individual is more responsive to rewards.

Discriminability and response bias were calculated based on previous behavioral models of signal detection [12,19].

Discriminability was computed as:

$$\text{Log}(d) = 1/2 * \log[(\text{Rich}_{\text{correct}} * \text{Lean}_{\text{correct}}) / (\text{Rich}_{\text{incorrect}} * \text{Lean}_{\text{incorrect}})]$$

Response bias was computed as:

$$\text{Log}(b) = 1/2 * \log[(\text{Rich}_{\text{correct}} * \text{Lean}_{\text{incorrect}}) / (\text{Rich}_{\text{incorrect}} * \text{Lean}_{\text{correct}})]$$

$\text{Rich}_{\text{correct}}$ and $\text{Lean}_{\text{correct}}$ correspond to the proportion of correct identifications (hits and correct rejections) to the total number of rich and lean trials respectively, and $\text{Rich}_{\text{incorrect}}$ and $\text{Lean}_{\text{incorrect}}$ correspond to the proportion of false identifications (misses and incorrect rejections) to the total number of rich and lean trials respectively. When accuracy was equal to 1 or 0, we followed the log linear correction procedure described by Hautus et al. [20].

Accuracy, RT and discriminability were analyzed in order to assess overall performance for each experiment separately (Money and Social) using two-way ANOVAs on accuracy, RT and discriminability scores with Condition (Rich, Lean), and Block (1, 2, 3) as within-subjects factors. No main effect or interaction involving gender was found; we therefore present results for both genders combined in all our analyses. As in Pizzagalli et al. (2005), trials with a reaction time shorter than 150 ms, longer than 2500 ms, or ± 3 SDs from the mean (after natural logarithm transformation) were excluded from all analyses (3.6% of the trials overall).

Our main variable of interest was response bias because it reflects participants’ responsiveness to monetary rewards or social rewards. Following Pizzagalli et al. (2005), overall change in response bias (or Δ Response Bias) was conceptualized as the difference between the first block and the last block. We first ran a repeated-measures ANOVA with Block as a within-subject factor to test whether Response Bias changed over time. We then examined the relationship between anhedonia and Δ Response Bias to either social or monetary rewards using correlations, a median-split based on participants’ social anhedonia scores.

Results

Task 1: Social rewards task. Overall task performance (Fig 2). A two-way ANOVA on accuracy scores with Condition (Rich, Lean), and Block (1, 2, 3) as within-subjects factors revealed a main effect of Condition, $F(1,43) = 5.55, p < .023, \eta_p^2 = .114$, a main effect of Block, $F(2,42) = 30.01, p < .001, \eta_p^2 = .588$, and a Condition x Block interaction, $F(2,42) = 9.616, p < .001, \eta_p^2 = .314$. As predicted, accuracy for the rich condition ($M = .777, SD = .090$) was higher

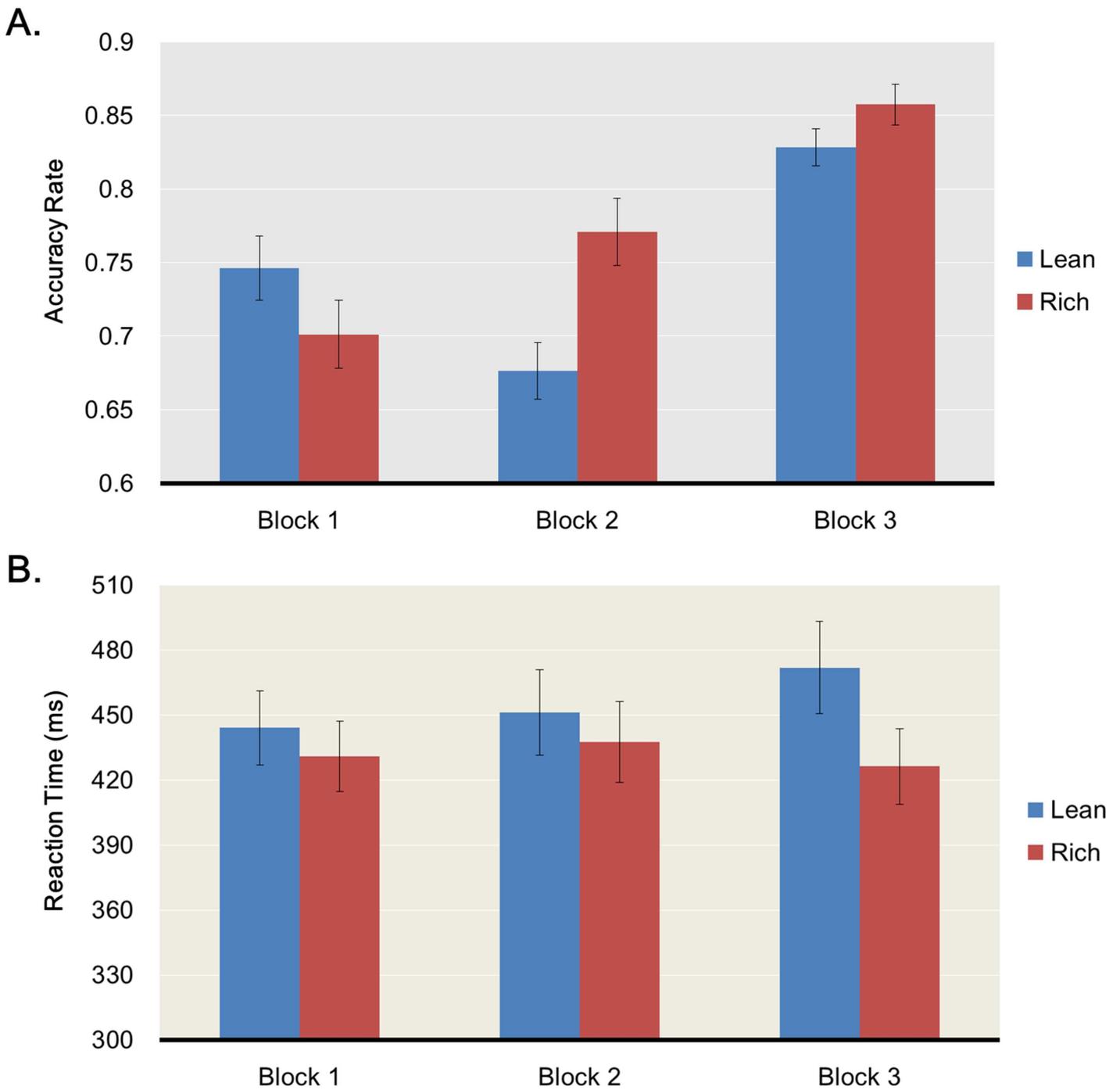


Fig 2. Performance in the social reward task. Accuracy rates (A) and reaction times (B) during the social reward task. Results are split by block (1, 2, 3) and stimulus type (lean in blue, rich in red).

doi:10.1371/journal.pone.0167024.g002

than for the lean condition ($M = .750, SD = .082, t(43) = 2.355, p = .023, d = .32$), and increased over time (Block 1 vs. 2, $t(43) = 0.002, p = .998$; Block 2 vs. 3, $t(43) = -7.83, p < .001$; Block 1 vs. 3, $t(43) = -4.91, p < .001$). The ANOVA on reaction times revealed a main effect of Condition,

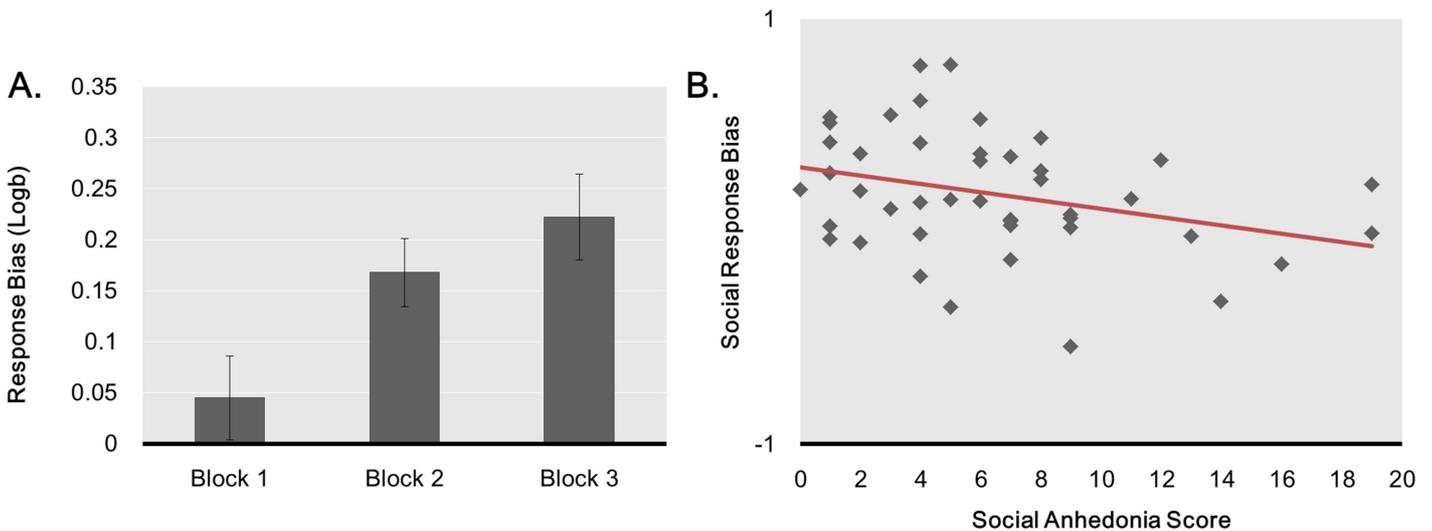


Fig 3. Response bias in the social reward task. Evolution of the response bias across blocks in the social reward task (A) and correlation between change in response bias (i.e. Logb Block 3 – Logb Block 1) and social anhedonia score, $F^2 = .116$.

doi:10.1371/journal.pone.0167024.g003

$F(1,43) = 23.45, p < .001, \eta_p^2 = .353$, no main effect of Block, $F(2,42) = 0.244, p = .785, \eta_p^2 = .011$, and a Condition x Block interaction, $F(2,42) = 5.417, p = .008, \eta_p^2 = .205$. Reaction times for the rich condition ($M = 431.7\text{ms}, SD = 103.2\text{ms}$) were faster than in the lean condition ($M = 455.8\text{ms}, SD = 116.0\text{ms}$), $t(43) = -4.84, p < .001$. Planned comparisons revealed that the Condition x Block interaction was due to participants being faster in Block 3 for the rich condition, $t(43) = 5.78, p < .001$, but not in the other two blocks, Block 1: $t(43) = 1.84, p = .073$, Block 2: $t(43) = 1.67, p = .103$. Finally, the ANOVA on discriminability with Block as a within-subject factor revealed no main effect, $F(2,42) = 1.203, p = .305, \eta_p^2 = .027$.

Response bias (Fig 3). In line with our hypothesis that response bias would increase over time, a repeated-measures ANOVA on response bias with Block as a within-subject factor revealed a main effect of Block, $F(2,42) = 7.77, p = .001, \eta_p^2 = .270$, due to an increase in response bias from Block 1 to Block 2, $t(43) = -3.081, p = .004$, and from Block 1 to Block 3, $t(43) = -3.902, p < .001$. Pearson correlations revealed that Δ Response Bias did not correlate with physical anhedonia scores, $r = -.044, p = .774, n = 44$ but correlated with social anhedonia scores, $r = -.340, p = .024, n = 44$, indicating that participants with stronger self-reported social anhedonia were also the least biased by social rewards. By contrast, social anhedonia scores and physical anhedonia scores strongly correlated, $r = .442, p = .003, n = 44$, suggesting that the specificity of each instrument should be questioned and that Δ Response Bias might be a more accurate measure of social anhedonia than a commonly used self-report measure.

Using a standard categorical approach, we then compared change in response bias in higher and lower social anhedonia participants. Specifically, we looked at the difference in response bias between Block 1 and 2 as well as between Block 1 and 3. Change in response bias was greater for lower social anhedonia participants than for higher social anhedonia participants for the Block 1 vs. Block 2 comparison, $t(42) = 2.409, p = .020$, and for the Block 1 vs. Block 3 comparison, $t(42) = 2.627, p = .012$. Importantly, Participants in the lower social anhedonia group (score range: 0–6, $M = 3.24, SD = 1.92, N = 25$) and a higher (score range: 7–19, $M = 10.47, SD = 3.98, N = 19$) and in the higher social anhedonia group did not differ with respect to age, $t(42) = 0.426, p = .672$, Gender ratio, $\chi^2(1) = 0.548, p = .547$, percentage of outliers, $Z = -0.996, p = .319$, Mann-Whitney U, or number of feedbacks received during the

experiment, Money: $t(42) = 0.206, p = .838$; Social: $z = -.326, p = .744$, Mann-Whitney U. Finally, the two groups did not differ in the rich-to-lean reward ratio, which is the most important variable that drives response bias (Money task: $t(42) = 0.709, p = .483$; Social Task: $t(42) = 1.239, p = .222$).

Task 2: Monetary rewards responsiveness. Overall task performance (Fig 4). A two-way ANOVA on accuracy scores with Condition (Rich, Lean) and Block (1, 2, 3) as within-subjects factors revealed a main effect of Condition, $F(1,43) = 16.15, p < .001, \eta_p^2 = .27$, a main effect of Block, $F(2,42) = 39.17, p < .001, \eta_p^2 = .65$, and a Condition x Block interaction, $F(2,42) = 9.37, p < .001, \eta_p^2 = .308$. As predicted, accuracy for the rich condition ($M = .798, SD = .067$) was higher than for the lean condition ($M = .751, SD = .095$), $t(43) = 4.019, p < .001$, and increased over time (Block 1 vs. 2, $t(43) = -2.277, p = .028$; Block 2 vs. 3, $t(43) = -6.72, p < .001$; Block 1 vs. 3, $t(43) = -7.97, p < .001$). The ANOVA on reaction times revealed a main effect of Condition, $F(1,43) = 25.95, p < .001, \eta_p^2 = .376$, no main effect of Block, $F(2,42) = 0.405, p = .670, \eta_p^2 = .019$, and no Condition x Block interaction, $F(2,42) = .842, p = .438, \eta_p^2 = 0.039$. Reaction times for the rich condition ($M = 417.9\text{ms}, SD = 105.9\text{ms}$) were faster than for the lean condition ($M = 444.9\text{ms}, SD = 118.1\text{ms}$), $t(43) = -5.09, p < .001$. Finally, the ANOVA on discriminability with Block as a within-subject factor revealed no main effect, $F(2,42) = 0.606, p = .548, \eta_p^2 = .014$.

Response bias (Fig 5). In line with our hypothesis that response bias would increase over time, a repeated-measures ANOVA on response bias with Block (1, 2, 3) as a within-subject factor revealed a trend for a main effect of Block, $F(2,42) = 2.955, p = .057, \eta_p^2 = .064$, due to an increase in response bias from Block 1 to Block 3, $t(43) = -2.105, p = .041$ and from Block 1 to Block 2, $t(43) = -7.594, p < .001$. Unlike what we found in Task 1, Δ Response Bias correlated neither with social anhedonia scores, $r = .024, p = .875, n = 44$, nor with physical anhedonia scores $r = .030, p = .849, n = 44$. Similarly, we found no difference between higher and lower social anhedonia participants on Δ Response Bias for Block 1 vs. Block 2, $t(42) = -0.148, p = .883$, or for Block 1 vs. Block 3, $t(42) = -0.432, p = .668$. We had no hypothesis regarding the correlation of Δ Response Bias between the Money and Social task but our exploratory analysis revealed a weak correlation, $r = .28, p = .066$.

Study 2: Direct Replication Study

In order to ensure the robustness of the relationship between social anhedonia and Δ Response Bias we ran the same study in an independent sample of participants.

Methods

Ethics Statement. This study was reviewed and approved by the local ethics committee (Comité de protection des personnes Ile de France III). After complete description of the study to the subjects, written informed consent was obtained.

Participants. 26 French participants (14 females and 12 males) completed the Social experiment described in Study 1. As in Study 1, participants were left alone after the computer task to complete the French version of the Beck Depression Inventory [21] and the Chapman Anhedonia Scales [22] Participants earned approximately 10€ for the entire task.

Analyses. No difference between males and females was found in any of the variables of interest, we therefore present results for both genders combined in all our analyses. As in Study 1, trials with a reaction time shorter than 150 msec, longer than 2500 msec or ± 3 SDs from the mean (after natural logarithm transformation) were excluded from all analyses (2.1% of the trials overall). Discriminability and response bias were calculated using the same formulae as in Study 1.

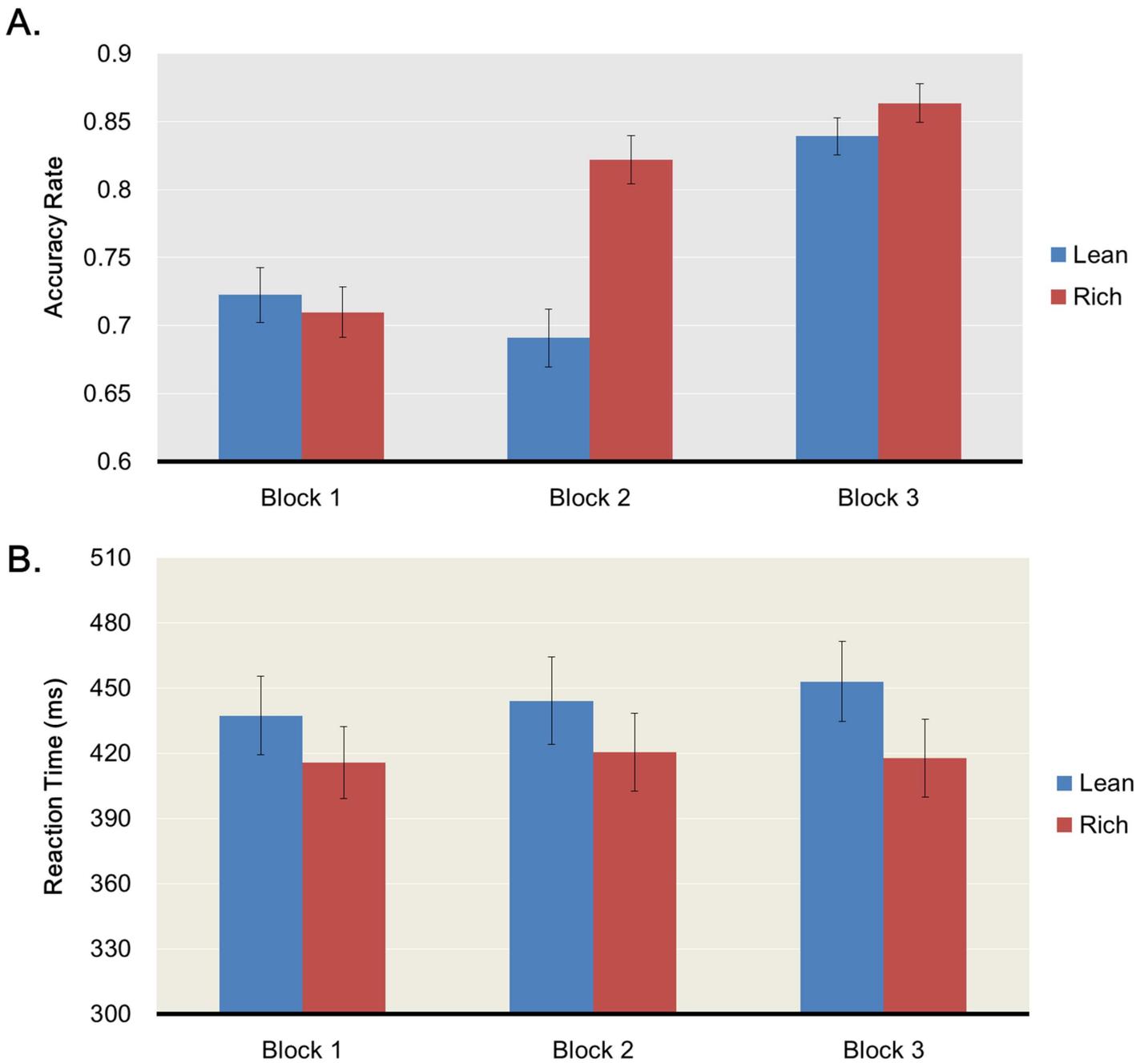


Fig 4. Performance in the monetary reward task. Accuracy rates (A) and reaction times (B) during the monetary reward task. Results are split by block (1, 2, 3) and stimulus type (lean in blue, rich in red).

doi:10.1371/journal.pone.0167024.g004

Results

Overall task performance. A two-way ANOVA on accuracy scores with Condition (Rich, Lean), and Block (1, 2, 3), as within-subjects factors revealed a main effect of Condition, $F(1,25) = 11.29, p < .01, \eta_p^2 = .31$, a main effect of Block, $F(2,24) = 8.11, p < .01, \eta_p^2 = .40$, but no Condition x Block interaction, $F(2,24) = 1.41, p = .26, \eta_p^2 = .10$. As predicted, accuracy for

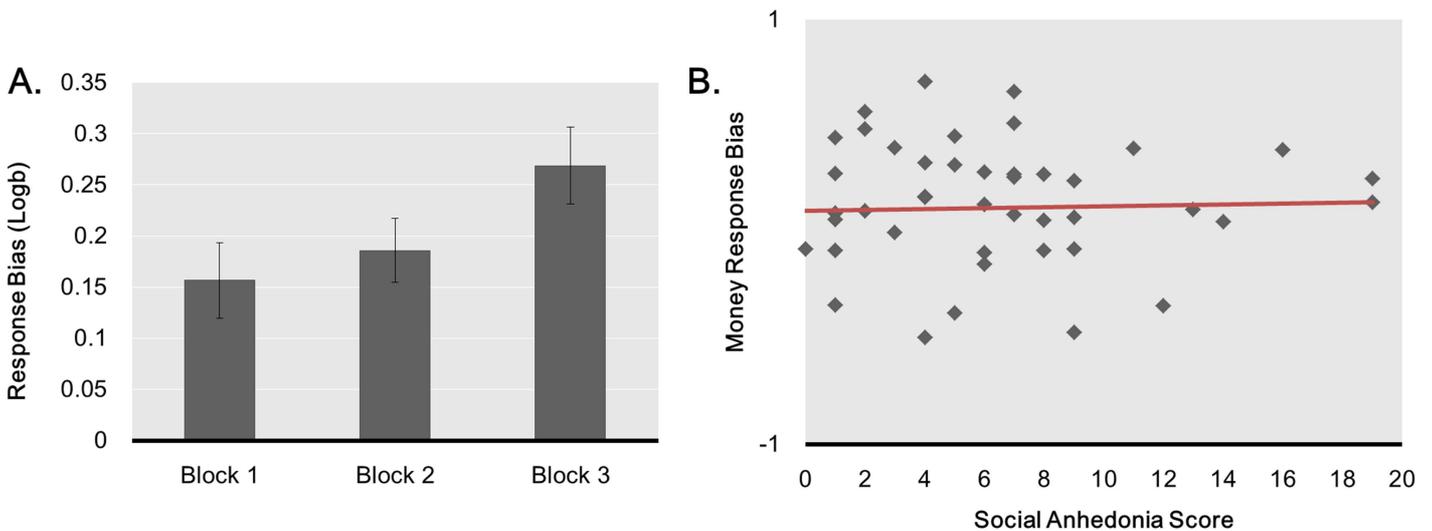


Fig 5. Response bias in the monetary reward task. Evolution of the response bias across blocks in the monetary reward task (A) and correlation between change in response bias (i.e., Logb Block 3–Logb Block 1) and social anhedonia score, $R^2 = .058$.

doi:10.1371/journal.pone.0167024.g005

the rich condition ($M = .853$, $SD = .101$) was higher than for the lean condition ($M = .760$, $SD = .153$), $t(25) = 3.36$, $p < .01$. The ANOVA on reaction times revealed a main effect of Condition, $F(1,25) = 13.20$, $p < .01$, $\eta_p^2 = .346$, no main effect of Block, $F(2,24) = 0.42$, $p = .661$, $\eta_p^2 = .034$, and no Condition x Block interaction, $F(2,24) = 1.47$, $p = .249$, $\eta_p^2 = 0.109$. Reaction times for the rich condition ($M = 466.9\text{ms}$, $SD = 124.5\text{ms}$) were faster than for the lean condition ($M = 491.1\text{ms}$, $SD = 134.6\text{ms}$), $t(25) = -3.63$, $p < .01$. Finally, the ANOVA on discriminability with Block as a within-subject factor revealed a main effect of Block, $F(2,24) = 11.024$, $p < .001$, $\eta_p^2 = .408$.

Response bias. A repeated-measures ANOVA on response bias with Block (1, 2, 3) as a within-subject factor revealed no main effect of Block, $F(2,24) = 0.808$, $p = .452$, $\eta_p^2 = .119$. Even though response bias is computed to be independent from discriminability, we checked for the presence of any significant correlation between these two variables. These analyses revealed no significant correlation between discriminability and response bias: $r = -.194$, $p = .343$, $n = 26$, as well as no significant correlation between discriminability and Δ Response Bias: $r = .092$, $p = .653$, $n = 26$. As found in Experiment 1 of Study 1, Δ Response Bias did not correlate with physical anhedonia scores $r = -.105$, $p = .609$, $n = 26$ but correlated with social anhedonia scores, $r = -.462$, $p = .017$, $n = 26$.

We then compared change in response bias in higher and lower social anhedonia participants by creating a median-split based on participants' scores in the Chapman social anhedonia scale. We separated our sample into a lower (range: 2–7, $n = 13$) and a higher (range: 8–27, $n = 13$) social anhedonia Group. Lower social anhedonia ($M = 5.46$, $SD = 1.85$) and Higher social anhedonia ($M = 14.15$, $SD = 5.67$) subjects did not differ with respect to Gender ratio, $\chi^2(1) = 0.002$, $p = .962$, percentage of outliers, $Z = -.515$, $p = .621$, Mann-Whitney U, number of feedbacks received during the experiment: $Z = .506$, $p = .673$, Mann-Whitney U, or rich-to-lean reward ratio, $t(24) = 1.239$, $p = .222$. The two groups differed on age, with the Higher social anhedonia group being older than Lower social anhedonia, $t(24) = 2.24$, $p = .034$. However, age did not correlate with Δ Response Bias, $r = -.303$, $p = .132$, $n = 26$, and Δ Response Bias still correlated with Social Anhedonia Scores when controlling for Age, $r = -.395$, $p = .039$, $n = 26$.

We then looked at the difference in response bias between Block 1 and 2 as well as between Block 1 and 3. Change in response bias was not significantly greater for lower social anhedonia participants than for higher social anhedonia participants for the Block 1 vs. Block 2 comparison, $t(24) = 0.505, p = .618$, but was significantly greater for lower social anhedonia participants than for higher social anhedonia participants for the Block 1 vs. Block 3 comparison, $t(24) = 2.180, p = .039$.

Discussion

Lack of reliable and objective measures available to assess social motivation is likely to be an important barrier in understanding the biological and genetic roots of a number of psychiatric disorders in which disrupted social motivation is commonly reported [7]. In this study, social rewards presented in the context of a signal detection task were found to positively bias performance in two independent healthy adult samples, which replicates and extends prior findings using monetary rewards [12]. Most importantly, individual differences in the strength of this social bias correlated with self-reported social motivation as measured with the social anhedonia scale. In other words, the behavior of participants who reported taking less pleasure in social interactions was less affected by social rewards.

Dimensional approaches to psychiatry have recently been encouraged as a necessary paradigm shift in order to further our understanding of the biological bases of a range of disorders. Indeed, genetic approaches rely on careful characterization of aberrant phenotypes to infer the consequence of genetic mutations; and in a similar way, neuroscientists infer the function of specific brain areas by relating brain activity to the subject's behavior. The idea of breaking down clinical symptomatology into biologically relevant units of quantifiable behaviors also aligns with NIMH's recent Research Domain Criteria (RDoC) initiative encouraging research on constructs that cut across diagnostic boundaries. As part of this initiative, working groups have been identifying promising and reliable research paradigms to study three key constructs in the social domain: i) affiliation and attachment (which encompasses social motivation), ii) social communication, iii) perception and understanding of self and others [23]. Interestingly, only three experimental paradigms are listed in the summary of the working groups as potential assays for affiliation and attachment, in comparison to more than 30 tools listed for each of the other two constructs. This highlights the great need to create finely graded dimensional tools for making predictions about genetic, neurobiological and other behavioral dimensions. It is clear that more measures are needed to fully develop a nomological network [24] of social reward and to validate concepts such as social motivation, affiliation and attachment.

One might however question the potential benefit of using behavioral measures rather than psychometric scales. In the context of our task, we found that the Social Anhedonia Scale is correlated with the Physical Anhedonia Scale whereas social reward responsiveness correlates specifically with social anhedonia (but not with physical anhedonia). This suggests that our task taps into a narrower—and hopefully better defined—construct than the one assessed by the Social Anhedonia Scale. Second, as we explained in the introduction, psychometric measures are limited in a number of ways: some people may lack insight and fail to accurately report their own social preferences and social skills, other individuals may be well-aware of their social limitations but may choose to under-report their issues to provide more socially desirable responses [11]. This concern applies to many items in the Chapman scale where there is an answer that is viewed more favorably by society (e.g., answering 'No' to: "I attach very little importance to having close friends", "Playing with kids is a real chore", "I don't really feel very close to my friends"). Given that the social desirability bias is not equally strong in all populations, this feature of psychometric scales assessing social motivation may amplify or

distort group differences when studying participants who are less concerned with their reputation (e.g., [7]). or with societal norms (e.g., [6]).

Our behavioral task also differs from a number of existing paradigms, notably eye-tracking experiments in which it is difficult to tease apart the influence of participants' preference for social stimuli from lower-level factors such as stimulus features [25] or individual differences in the visual system [26]. In designing our study, our goal was to make sure that participants' behavior was truly guided by internal value signals rather than by low level features of the task. Similarly, a recent study by Dubey et al. on social seeking demonstrated that typical adults put in more conscious effort to view social vs. non social stimuli, which can be used as a proxy to measure social motivation [27]. We believe that our protocol, which focuses more specifically on social reward responsiveness rather than on social seeking, nicely complements this instrument, with the added value that participants have to make no conscious decision to react to the social signal. The implicit nature of the task indeed prevents social desirability biases from confounding the results.

There are, however, a number of limitations to the present study. First, it will be important to use this tool in clinical populations to assess its sensitivity to variations in social motivation in various psychopathologies; e.g. Is social reward responsiveness lower in Autism Spectrum Disorders, higher in Williams Syndrome, lower in depression, etc.? Another interesting empirical question will be to determine whether social reward responsiveness varies with participants' social cognition abilities. It is indeed possible that individuals who are confused by social interactions do not respond to social rewards in the same way as individuals who are more socially gifted.

Ultimately, advancing the validation of adequate psychological tools that can reliably quantify subtle individual differences in social functioning is an important challenge to advance our knowledge of the genetics and neurobiology of psychiatric conditions. Behavioral tools measuring social motivation might also prove necessary for identifying subtypes for various disorders affected by social motivation impairments, including schizophrenia, ASDs, anorexia nervosa and depression and in gaining a better understanding of other clinically complex conditions. Reliable and finely graded dimensional measures will also make more sensitive predictions about clinically relevant outcomes, such as predicting risk and disorder onset, natural history, and differential response to various treatments. Combining diagnostic categories with dimensional description is therefore likely to be a fruitful paradigm, as it has been in most areas of medicine for many years [28]. As Rutter recently pointed out, the use of dimensional approaches will indeed allow us "to portray the true picture of a clinical presentation that involves multiple facets or straddles the boundary between two adjacent categories, without ending up with spurious comorbidity and without the necessity of forcing symptom patterns into a predetermined stereotype." ([28], p. 655).

Author Contributions

Conceptualization: CC NT JM RTS.

Data curation: CC LS NT DK JM.

Formal analysis: CC NT LS.

Funding acquisition: CC RTS.

Investigation: CC NT LS DK JM.

Methodology: CC NT LS RTS.

Project administration: CC NT DK JM.

Resources: GK JM RTS.

Software: CC NT LS.

Supervision: CC GK RTS JM.

Validation: CC NT LS.

Visualization: CC NT LS DK.

Writing – original draft: CC NT LS DK GK JM RTS.

Writing – review & editing: CC NT LS DK GK JM RTS.

References

1. Insel T, Cuthbert B, Garvey M, Heiassen R, Pine DS, Quinn K, et al. Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *Am J Psychiatry*. 2010; 167: 748–751. doi: [10.1176/appi.ajp.2010.09091379](https://doi.org/10.1176/appi.ajp.2010.09091379) PMID: [20595427](https://pubmed.ncbi.nlm.nih.gov/20595427/)
2. Adam D. Mental health: On the spectrum. *Nature*. 2013; 496: 416–418. doi: [10.1038/496416a](https://doi.org/10.1038/496416a) PMID: [23619674](https://pubmed.ncbi.nlm.nih.gov/23619674/)
3. Silvia PJ, Kwapil TR. Aberrant asociality: how individual differences in social anhedonia illuminate the need to belong. *J Pers*. 2011; 79: 1315–1332. doi: [10.1111/j.1467-6494.2010.00702.x](https://doi.org/10.1111/j.1467-6494.2010.00702.x) PMID: [21480908](https://pubmed.ncbi.nlm.nih.gov/21480908/)
4. Watson KK, Werling DM, Zucker N, Platt M. Altered social reward and attention in anorexia nervosa. *Front Psychol*. 2010; 1: 36. doi: [10.3389/fpsyg.2010.00036](https://doi.org/10.3389/fpsyg.2010.00036) PMID: [21887145](https://pubmed.ncbi.nlm.nih.gov/21887145/)
5. Hasler G, Drevets WC, Manji HK, Charney DS. Discovering endophenotypes for major depression. *Neuropsychopharmacology*. 2004; 29: 1765–1781. doi: [10.1038/sj.npp.1300506](https://doi.org/10.1038/sj.npp.1300506) PMID: [15213704](https://pubmed.ncbi.nlm.nih.gov/15213704/)
6. Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA, et al. Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol Psychiatry*. 2007; 61: 1260–1271. doi: [10.1016/j.biopsych.2006.07.021](https://doi.org/10.1016/j.biopsych.2006.07.021) PMID: [17046722](https://pubmed.ncbi.nlm.nih.gov/17046722/)
7. Chevallier C, Kohls G, Troiani V, Brodtkin ES, Schultz RT. The social motivation theory of autism. *Trends Cogn Sci*. 2012;
8. Chapman L, Chapman J, Raulin M. Scales for physical and social anhedonia. *J Oj Abnorm Psychol*. 1976; 85: 374–382.
9. Mehrabian A. The development and validation of measures of affiliative tendency and sensitivity to rejection. *Educ Psychol Meas*. 1970; Available: <http://psycnet.apa.org/psycinfo/1971-02411-001>
10. Happé F. Theory of mind and the self. *Ann N Y Acad Sci*. 2003; 1001: 134–144. PMID: [14625359](https://pubmed.ncbi.nlm.nih.gov/14625359/)
11. Uziel L. Rethinking Social Desirability Scales From Impression Management to Interpersonally Oriented Self-Control. *Perspect Psychol Sci*. 2010; 5: 243–262. doi: [10.1177/1745691610369465](https://doi.org/10.1177/1745691610369465) PMID: [26162157](https://pubmed.ncbi.nlm.nih.gov/26162157/)
12. Pizzagalli DA, Jahn AL, O'Shea JP. Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biol Psychiatry*. 2005; 57: 319–327. doi: [10.1016/j.biopsych.2004.11.026](https://doi.org/10.1016/j.biopsych.2004.11.026) PMID: [15705346](https://pubmed.ncbi.nlm.nih.gov/15705346/)
13. McCarthy DC. Behavioral detection theory: Some implications for applied human research. *Signal Detect Mech Models Appl*. 1991; 239.
14. Macmillan NA, Creelman CD. *Detection theory: A user's guide* [Internet]. Psychology press; 2004.
15. Kohls G, Perino MT, Taylor JM, Madva EN, Cayless SJ, Troiani V, et al. The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*. 2013; 51: 2062–2069. doi: [10.1016/j.neuropsychologia.2013.07.020](https://doi.org/10.1016/j.neuropsychologia.2013.07.020) PMID: [23911778](https://pubmed.ncbi.nlm.nih.gov/23911778/)
16. Beck A, Ward C, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961; 4: 561. PMID: [13688369](https://pubmed.ncbi.nlm.nih.gov/13688369/)
17. Johnstone V, Alsop B. Reinforcer control and human signal-detection performance. *J Exp Anal Behav*. 2000; 73: 275–290. doi: [10.1901/jeab.2000.73-275](https://doi.org/10.1901/jeab.2000.73-275) PMID: [10866352](https://pubmed.ncbi.nlm.nih.gov/10866352/)
18. Green DM, Swets JA. *Signal detection theory and psychophysics*. New York: Wiley; 1966.
19. Tripp G, Alsop B. Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. *J Clin Child Psychol*. 1999; 28: 366–375. doi: [10.1207/S15374424jccp280309](https://doi.org/10.1207/S15374424jccp280309) PMID: [10446686](https://pubmed.ncbi.nlm.nih.gov/10446686/)

20. Hautus MJ. Corrections for extreme proportions and their biasing effects on estimated values of d' . *Behav Res Methods Instrum Comput.* 1995; 27: 46–51.
21. Collet L, Cottraux J. Inventaire abrégé de la dépression de Beck (13 items). *L'encéphale.* 1986; 10: 77–9.
22. Kosmadakis CS, Bungener C, Pierson A, Jouvent R, Widlöcher D. Traduction et validation de l'Echelle Révisée d'Anhédonie Sociale (SAS Social Anhedonia Scale, ML Eckblad, LJ Chapman et al., 1982): étude des validités interne et concourante chez 126 sujets sains. *L'encephale.* 1995; 21: 437–443. PMID: [8674468](#)
23. NIMH. Social Processes: Workshop Proceedings [Internet]. Available: <http://www.nimh.nih.gov/research-funding/rdoc/social-processes-workshop-proceedings.shtml>
24. Meehl P. CLJ. Construct validity in psychological tests. *Psychol Bull.* 1955; 52: 281–302. PMID: [13245896](#)
25. Chevallier C, Parish-Morris J, McVey A, Rump KM, Sasson NJ, Herrington JD, et al. Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Res.* 2015; 8: 620–628. doi: [10.1002/aur.1479](https://doi.org/10.1002/aur.1479) PMID: [26069030](#)
26. Behrmann M, Thomas C, Humphreys K. Seeing it differently: Visual processing in autism. *Trends Cogn Sci.* 2006; 10: 258–264. doi: [10.1016/j.tics.2006.05.001](https://doi.org/10.1016/j.tics.2006.05.001) PMID: [16713326](#)
27. Dubey I, Ropar D, de C Hamilton AF. Measuring the value of social engagement in adults with and without autism. *Mol Autism.* 2015; 6: 1.
28. Rutter M. Research review: Child psychiatric diagnosis and classification: concepts, findings, challenges and potential. *J Child Psychol Psychiatry.* 2011; 52: 647–660. doi: [10.1111/j.1469-7610.2011.02367.x](https://doi.org/10.1111/j.1469-7610.2011.02367.x) PMID: [21434914](#)

Childhood Environmental Harshness Predicts Coordinated Health and Reproductive Strategies: A Cross-sectional Study Of A Nationally Representative (France) Sample.

Hugo Mell^{1,3*}, Lou Safra¹, Yann Algan², Nicolas Baumard³, Coralie Chevallier^{1,3}

¹Laboratoire de Neurosciences Cognitives (LNC), Département d'Etudes Cognitives, INSERM U960, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France.

²Sciences Po, OFCE, 27 Rue Saint-Guillaume, 75007, Paris, France.

³Institut Jean Nicod, Département d'Etudes Cognitives, CNRS UMR8129, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France.

* Corresponding author: Hugo Mell, hugo.mell@ens.fr

Abstract

There is considerable variation in health and reproductive behaviours within and across human populations. Drawing on principles from Life History Theory, psychosocial acceleration theory predicts that individuals developing in harsh environments decrease their level of somatic investment and accelerate their reproductive schedule. Although there is consistent empirical support for this general prediction, most studies have focused on a few isolated life history traits and few have investigated the way in which individuals apply life strategies across reproductive and somatic domains to produce coordinated behavioural responses to their environment. In our study, we thus investigate the impact of childhood environmental harshness on both reproductive strategies and somatic investment by applying structural equation modelling (SEM) to cross-sectional survey data obtained in a representative sample of the French population (n=1015, age: 19-87 years old, both genders). This data allowed us to demonstrate that (i) inter-individual variation in somatic investment (*e.g.* effort in looking after health) and reproductive timing (*e.g.* age at first birth) can be captured by a latent fast-slow continuum, and (ii) faster strategies along this continuum are predicted by higher childhood harshness. Overall, our results support the existence of a fast-slow continuum and highlight the relevance of the life history approach to understand variations in reproductive and health related behaviors.

Keywords: Psychosocial acceleration theory; Childhood adversity; Life History Theory; Reproductive strategies; Health strategies; Structural Equation Modelling.

Introduction

People engage in numerous behaviours that bear a high cost to the individual and to society: smoking, poor engagement in health-promoting behaviours, overconsumption of high calory foods, teen-pregnancy, etc. In OECD countries, for instance, “lifestyle” conditions linked to tobacco use, excessive diets and physical inactivity are now responsible for most years of lost life (Hurst and Sassi, 2008). Therefore, identifying the determinants of health and reproductive decisions is of vital importance. Here, we argue that behavioural diversity for health and reproductive decisions should not come as a surprise and should be construed as the predictable outcome of humans’ evolutionary make-up. Specifically, we investigate the idea that health and reproductive decisions are adjusted during development to the way individuals perceive their environment. Put simply, the hypothesis is that focusing on one’s health or delaying reproduction to invest in other areas of life might be less beneficial in environments where mortality is high than in environments where mortality is low. Childhood environmental harshness may therefore place individuals on a reproductive and health path that is calibrated to their ecology. While previous studies have already highlighted such effects of harshness on reproduction and health behaviours independently, we go further by integrating variation in both domains to take into account the coordinated nature of people’s allocation strategies.

Life History Theory (Roff, 2002; Stearns, 1992) provides a general framework to investigate variation in allocation decisions. It states that the life history strategy of any organism is the product of the interaction between tradeoffs among traits and environmental factors that affect mortality and fertility rates (Stearns, 2000). Drawing on these fundamental insights, evolutionary psychologists started to investigate how adversity events occurring during ontogeny, could be used by individuals as cues to adjust their strategies (Belsky et al., 1991; Chisholm et al., 1993). This led to a fruitful body of research often designated by the name “psychosocial acceleration theory” (Nettle et al., 2012). Specifically, it predicts that individuals living in harsh environment, *i.e.* exposed to high mortality regimes, should exhibit overall “fast” strategies, whereas individuals living in favorable environments, *i.e.* exposed to low mortality regimes, are expected to show overall “slow” strategies (Ellis et al., 2009). The “fast” end of this fast-slow continuum is generally characterized by a shortened period of growth associated with an early onset of reproduction (early sexual maturation and first reproductive event), a higher number of offspring with a lower investment per offspring, lower body maintenance and a reduced lifespan; whereas the slow end of the continuum has the opposite characteristics (Ellis et al., 2009). According to this theory, having a faster

strategy in harsh environments increases an individual's chances to reproduce before dying, whereas a slow strategy in favorable environments would allow for a longer growth period, which in turn, would lead to larger future reproductive benefits.

Although the theoretical link between harshness and fast strategies is not straightforward (Baldini, 2015), empirical studies have repeatedly found that exposure to harsh events during childhood is indeed associated with a fast reproductive strategy later in life. In particular, shorter life expectancy predicts younger ages at first birth both within and across populations (Geronimus et al., 1999; Low et al., 2008; Quinlan, 2010; Walker et al., 2006) as well as earlier sexual debuts and more sexual partners (Simpson et al., 2012). High socioeconomic deprivation and low parental care are also associated with earlier reproduction (Nettle, 2011) and with an earlier onset of puberty (Ellis, 2004; Ellis and Essex, 2007; Tither and Ellis, 2008). Internal features that might serve as cues of harshness, such as lower birthweights, also predict early reproduction in a longitudinal study of the British population, even after controlling for other socioecological variables (Nettle et al., 2009). Hence, people coming from harsher backgrounds develop overall faster reproductive strategies that manifest in a coherent manner for various sexual traits.

In parallel to these effects on reproduction, exposure to harsh environments also influences resource allocation to body maintenance (Del Giudice, 2014a). There is indeed a well-documented social gradient in preventive health behaviours (Stringhini, 2010) and part of the disinvestment in health observed in people with lower SES could be due to initial disparities in life expectancies (Nettle, 2010). Indeed, subjective socioeconomic standing is associated with reported effort in looking after one's health in a cross-sectional sample of the American population, and the effect of subjective socioeconomic position is fully mediated by perceived extrinsic mortality risks (Pepper and Nettle, 2014). Inter-individual differences in risky behaviors such as risky sexual behaviors, alcohol or drug use, which are in part predicted by early exposure to harsh environments in longitudinal studies of adolescent behaviors (Brumbach et al., 2009; Simpson et al., 2012), could also be interpreted as a disinvestment in body maintenance in response to higher extrinsic mortality risks.

Overall, there is therefore consistent empirical support for the application of life history theory principles to the study of allocation strategies in human populations. Nevertheless, most studies have focused on the impact of environmental harshness on a few life history traits restricted to one of the two allocation domains previously highlighted, *i.e.* reproductive or somatic efforts. Yet, based on psychosocial acceleration theory, we actually expect clusters of correlated traits across these domains, reflecting functional suites of

multiple traits that aim towards short-term returns in harsh conditions and long-term returns in favorable environments (Griskevicius et al., 2011; Reale et al., 2010). Indeed, if the fast-slow continuum is a broad axis of variation relevant to human life history strategies, all else being equal, individuals who adopt fast reproductive strategies should also exhibit lesser investments in their embodied capital (Kaplan et al., 2003). To our knowledge, only one study (Brumbach et al., 2009) explicitly assessed life history strategies with traits related to both reproductive and somatic investments in a single sample and showed that exposure to harsh events during adolescence predicted faster strategies across domains in young adulthood. In the current paper, we further test the existence of coordinated fast-slow strategies by analyzing data from a cross-sectional survey of a nationally representative sample of the French population specifically designed to test the existence of such a fast-slow continuum. We used structural equation modelling to test the prediction that part of the variation in reproductive and somatic effort is predicted by individual differences in exposure to harsh events during childhood. Specifically, we predicted that: 1) It is possible to identify a latent construct reflecting individuals' Life History Strategy which influences decisions pertaining to both reproductive and health choices; 2) Individuals' Life History Strategy fall along a fast-slow continuum. 3) Childhood environmental harshness has an influence on Life History Strategy, such that increased harshness leads to faster behaviours on the fast-slow continuum. Finally, we also tested whether specific harsh events were better predictors of individuals' future Life History Strategy.

Methods

Sample and procedure

Our sample consisted of 1015 French males (N=447) and females (N=568), aged 19 to 87 years old (mean: 52.5 ± 14.3 sd) and recruited online by the French polling institute Ipsos. Initially, 11,000 people received an electronic invitation from the institute to take part in our online study. Subjects willing to participate first had to answer a demographic survey which collected information about their gender, age, location, household composition, marital status, socio-professional category, occupational status, annual income and educational status. The quota sampling method was applied to select a fraction of the individuals based on these demographic variables, in order to obtain a representative sample of the French population. The retained subjects were then asked to answer a second survey grouping all the items

pertaining to life history traits and environmental harshness during childhood. Two of our key reproductive variables, namely *age at first birth* and *number of children*, were relevant only for people who had already reproduced. Therefore, among the 1691 participants who completed the questionnaire, we restricted our final study sample to individuals who already had children at the time of the study (N=1063). We also excluded participants who exhibited a number of non-response or absurd answers (*e.g.* number of years spent smoking greater than the participant's age) above three standard deviations (N=48). Our final sample size was 1015 participants.

Variables of interest

Participants were asked to answer questions pertaining to their childhood environment and their adult reproductive and health strategy. We now present a summary of the various areas covered by our questionnaire (full questionnaire available in the Supplementary Information).

Environmental Harshness: The level of environmental harshness experienced during childhood was assessed with a survey consisting of 24 items, reflecting various aspects of childhood environment that previous studies had found to be associated with one or several life history traits in adulthood (Griskevicius et al., 2011; McCullough et al., 2012; Nettle and Cockerill, 2010; Simpson et al., 2012). The first seven items captured general features of the family unit during participants' childhood. Sample items include "Have you ever lived with a stepfather?" and "Were you ever placed in an institution or in a foster family?". A three-item "parental investment" scale was used to assess participants' perception of the parental care they received during childhood, with items such as "My parents always seemed to care about what I was doing.". A "parenting style" scale of three items captured the harshness of parental education, with items such as "Some of the punishments I received when I was a child now seem too harsh to me.". Participants were also asked if they had been the victim of psychological, sexual or physical abuse during childhood and whether these episodes were caused by people in or outside their families. A subsequent index concerned the exposure to other particular familial difficulties. Example item included "Did you live with one or several people who had spent time in prison?". Participants were also asked whether they had suffered a long illness requiring a hospitalization before the age of seven and a "neighborhood stability" scale collected information about the stability of their growing-up environment with two items ("How many times did you move?" and "How many times did you change school?"). Lastly, participants' childhood socioeconomic status (SES) was measured by a

scale developed by Griskevicius and colleagues (Griskevicius et al., 2013) from the following three items: “My family usually had enough money for things when I was growing up.”, “I grew up in a relatively wealthy neighborhood.” and “I felt relatively wealthy compared to the other kids in my school.”. Cronbach’s alphas were superior to 0.8 for all the scales used in the analysis, suggesting good internal consistencies.

Reproductive strategy: Participants’ reproductive strategy was assessed using four items from the literature (Nettle et al., 2009; Simpson et al., 2012): number of children (“How many children have you had?”); age at first birth (calculated from the age of the participant and the reported birth date of their first child); age at first (consented) sexual intercourse; and number of short-term sexual partners. For the last two items, participants could choose not to answer the question by selecting an “I don’t want to answer” response.

Somatic strategy: Participants’ somatic strategy was also assessed using four items previously used in the literature (Pepper and Nettle, 2014): body mass index (BMI), calculated based on reported height and weight following the standard formula used in the biomedical field; general health status (“How is your health in general?”); health effort (“How much effort do you make to look after your health and ensure your safety these days?”); and level of cigarette’s consumption (“In total, during how many years did you smoke daily or almost daily?”). The responses for this last item were divided by the participant’s age to allow for more a meaningful comparison between young adults and older participants.

Analysis

All statistical analyses were carried out in R 3.0.3 (<https://www.r-project.org/>). Since this study aimed to investigate a specific theoretical model that involved a latent construct, namely “life history strategies”, structural equation modeling (SEM) was used as our main multivariate analysis method. Although our variables of interest showed overall low percentages of missing responses (ranging from 0 to 6.5%), multiple imputation techniques were used to preserve sample size and avoid biased estimations of model parameters. Twenty complete datasets were generated by fully conditional specifications for categorical and continuous data using the r package *mice* (Buuren and Groothuis-Oudshoorn, 2011). This package allows the use of different imputation methods depending on the type of variable with missing entries. Predictive mean matching was used for numeric variables, logistic

regression imputation for binary data and proportional odds model for ordered categorical variables with more than two levels.

In order to assess the effect of harshness during childhood on life history strategies later in life, a synthetic harshness measure had to be constructed based on the associated survey items. Drawing on the methodology used by Brumbach and colleagues (2009), environmental harshness was modeled as an emergent variable rather than a reflective latent variable. Indeed, environmental harshness is arguably better conceived as an emergent variable since harsh events during childhood can be thought of as risk factors (like particular genetic variants, smoking and poor diet for cardiovascular diseases) that are not necessarily correlated with one another, but that all contribute to the cumulative probability of developing a particular outcome; in our case a faster or slower strategy. For example, having been exposed to the death of a sibling, hospitalized for a long illness or lived with a stepfather are three events that we can theoretically expect to increase the probability of developing a faster strategy, but that might often occur independently.

Furthermore, we do not expect that all harsh events will have effects of the same magnitude on the cumulative risk of developing a particular life history strategy. Instead, some events might be better accounted for in a general harshness score when they are attributed heavier weights. One simple method to model this type of emergent variables used by Brumbach and colleagues (2009) is to sum individual z -scores for each harshness item. The use of z -transformed scores confers more weight to the most highly dispersed items and therefore reflect the implicit assumption that rare harsh events should be better predictors of fast strategies. Thus, the experience of rare events like losing one's mother or having been the victim of physical abuse will contribute more to an individual's harshness score than more frequent events, such as having changed school a couple of times.

Nonetheless, such an assumption might not hold in all cases and it would be valuable to compute weights of the harshness items based on their predictive power rather than implicitly through the degree of dispersion of their distributions. Such an approach can be implemented in SEMs using unknown weight composites, which capture the collective effects of a set of causes on a response variable (Grace and Bollen, 2008). In this case, the composite score is computed via a set of weights that maximize variance explanation in the dependent variable and hence allows to compare the relative contribution of the hypothesized causes to the overall predictive power of the composite. Thus, after fitting a SEM following the methodology previously used by Brumbach and colleagues (2009; Figure 1A), harshness was

also modelled as an unknown weights composite in a subsequent SEM to gain these inferential benefits (Figure 1B).

Whether harshness was computed as a sum of z-scores or as a composite, it was used as a predictor of the latent variable capturing individuals' general life history strategies. This latent construct was modelled as a unique factor capturing the covariation between all life history indicators (*i.e.* reproductive and somatic items). Yet, one might expect that items within each strategy will show some additional degree of correlation that will not be captured by the single general factor. For example, subjects suffering from hereditary diseases would probably tend to declare a poor health state and higher efforts in looking after their health even though it might not be linked to a faster or slower reproductive strategy. To deal with this issue we elaborated on the single factor model by allowing for correlations between residual errors of items within the same domain (*i.e.* only between somatic and reproductive items respectively). This way the model's implied covariance matrix captures the correlations between items that are not explained by the latent life history factor but that can still be theoretically expected due to various unmeasured causes.

The latent variable reflecting individuals' Life History Strategy was scaled by fixing its variance to 1 in both SEMs. Composite variables also need to be scaled for identification purposes by fixing the coefficient of one of the causal indicator. Therefore, in the second SEM, harshness was scaled by setting the path from *violence in the family* to 1. The latter item measured whether participants had been victim as a child of physical, sexual or psychological abuse caused by people in their family. Its significance was assessed through the partially reduced form of the model, which directly estimates the pathways from the harshness items to the latent variable without the use of a composite (Grace et al., 2010). Finally, since our study sample covers a wide age range (19-87 years old), age was used as an auxiliary variable to control for its effects on life history indicators. SEM models were fitted using the R packages *lavaan* (Rosseel, 2012) and the function *runMI* of the R package *semTools* (semTools Contributors, 2016) was used to combine the results obtained for the 20 imputed datasets. Parameter estimates and standard errors were pooled using Rubin's rules (Rubin, 2004). The MLMV estimator was used for its robustness to departures from normality since this assumption did not hold for all reflective indicators. Hence the dependent variable *health status* with four ordered levels had to be treated as continuous to allow the use of this robust maximum likelihood estimator. The large size of our sample and the absence of floor or ceiling effects in this variable justified such a treatment (see Supplementary Information, S1).

Finally, the chi-square statistics and the related fit indices were pooled using the method described in Li, Meng, Raghunathan, & Rubin (1991).

Results

Correlation matrix and descriptive statistics

Based on psychosocial acceleration theory, we expected correlations between all life history measures and with childhood environmental harshness in a pattern consistent with the relationships implied by the fast-slow continuum. Table 1 reports descriptive statistics and the correlation matrix for the variables included in the first SEM (Figure 1A). An extended table including the individual harshness indicators used in the second SEM (Figure 1B) is available in the Supplementary Information (S2).

The raw correlation matrix shows low but significant correlations among some of the life history indicators and with the global harshness score. Furthermore, the direction of the effect is consistent with the theory for every significant correlation. Hence, to further explore this pattern and to assess the theoretical model presented in the introduction, we fitted the two structural equation models represented in Figure 1 on the data. Fit indices and parameters estimates are reported separately for each model in the following sections.

SEM with harshness as a sum of z-scores

- Model fit

Table 2 reports fit indices for the SEMs. The chi-square test yielded significant p-values for the first SEM. However, the large sample size of our study (N=1015) prevents us from interpreting this statistic as evidence for a discrepancy between the sample and the model-implied covariance matrix. The chi-square statistics is indeed known to be particularly sensitive to sample size, which can lead models fitted on large samples to be systematically rejected (Schermelleh-Engel *et al.*, 2003). We therefore focus on several approximate fit indices, the Comparative Fit Index (CFI), Standardised Root Mean Square Residual (SRMR) and Root Mean Square Error of Approximation (RMSEA), which eliminate the issue of sample size dependency (Kline *et al.*, 2015). The RMSEA value of 0.046, associated with a close-fit test giving a p-value of 0.63, suggests an approximately good fit of the model. CFI and SRMR values of respectively 0.93 and 0.030 are also consistent with a close fitting model. Therefore, the approximate fit indices reveal no strong misspecification for this model.

- Measurement model: the Life History Strategy latent factor

Parameter estimates, standard errors and associated p-values of the relationships of primary focus are reported in Table 3 (see Supplementary Information S3 for an extensive list of model coefficients). All life history variables included in the model loaded significantly on the general life history latent factor except “number of children”. Inspection of the estimated covariance however, shows that “number of children” is not independent of the other reproductive items but correlates with “age at first birth” ($r = -0.36, p < 0.001$). Yet, even though the moderate correlation between “number of children” and “age at first birth” is consistent with the theory, it is not part of the general pattern captured by the life history factor. For all other life history items, the pattern of covariation follows our predictions (Table 3): higher scores on the life history factor are associated with lower BMI (standardized $c = -0.18, p = .01$), higher self-reported efforts in looking after one’s health (standardized $c = 0.21, p = .005$), higher self-reported general health status (standardized $c = 0.15, p < .001$), a lesser proportion of life spent smoking daily (standardized $c = -0.15, p < .001$), later age at first birth (standardized $c = 0.16, p < .001$), later sexual debut (standardized $c = 0.27, p < .001$) and fewer short-term sexual partners (standardized $c = -0.25, p < .001$); and therefore correspond to the theoretical description of a slow strategy. Hence, the latent life history construct in the first SEM is consistent with the proposed fast-slow continuum.

- Structural model: effect of harshness on the life history strategy factor

Concerning the structural part of the first SEM, the global score of harshness during childhood is negatively associated with the latent variable (standardized $c = -0.34, p < 0.001$; Table 3). This relationship confirms the predictions of the theory since higher scores on the harshness index are associated with lower scores on the life history factor, which reflect a faster strategy.

SEM with harshness as a latent composite

- Model fit

For the second model, as expected, the chi-square test yields a significant p-value. SRMR and RMSEA values were closer to zero (respectively 0.028 and 0.026) compared to the first SEM, which indicates a closer fit. On the other hand, the CFI index with a value of 0.83, which is inferior to the soft criterion of 0.90 (Bentler & Bonett 1980), no longer indicates a close fit. The latter discrepancy might be due to the numerous degrees of freedom

introduced by including the various harshness items/scales. Overall, the approximate fit indices still reveal no strong misspecification for this model.

- Measurement model: the Life History Strategy latent factor

Coefficients related to the latent life history factor are very similar to those obtained for the first SEM. Although the particular values of the coefficients slightly vary, the overall pattern of covariation is identical: the variable “number of children” does not significantly correlate with the general factor but the remaining life history indicators covary as predicted by the fast-slow continuum (Figure 1B). Indeed, higher scores on the life history strategy factor still reflect slower strategies characterized by lower BMI (standardized $c = -0.15$, $p = .03$), higher health efforts (standardized $c = 0.24$, $p < .001$), better health status (standardized $c = 0.19$, $p < .001$), a lesser proportion of life spent smoking daily (standardized $c = -0.14$, $p < .001$), later age at first birth (standardized $c = 0.15$, $p < .001$), later sexual debut (standardized $c = 0.24$, $p < .001$) and fewer short-term sexual partners (standardized $c = -0.21$, $p = .003$).

- Structural model: effect of harshness on the life history strategy factor

Childhood harshness measured as a latent composite in the second SEM also predicted faster life history strategies. The use of a composite led to slightly stronger association between these variables (standardized $c = -0.37$, $p < .001$; Figure 1B). However, the examination of the composite weights also reveals that this effect of childhood harshness is mostly driven by the item *violence in the family*. Indeed, only this item, which measured whether participants suffered from physical, sexual or psychological abuse caused by members of their family, contributed significantly to the effect of the composite (partially reduced model: standardized $c = -0.24$, $p < .001$). Marginal contributions to the composite’s effects on the latent life history factor of *death of the mother* (standardized $c = 0.41$, $p = .055$) and *having lived with a stepfather* (standardized $c = 0.54$, $p = .061$) also emerged from this model.

Discussion

Research in human behavioural ecology suggests that exposure to high levels of environmental harshness during ontogeny increases the probability of individuals adopting fast strategies. Previous studies have provided empirical support for this proposal by examining patterns of inter-individual differences often for various measures of either somatic

or reproductive investments (Belsky, 2012; Pepper and Nettle, 2014). The present research further supports these findings by showing that, in a representative sample of the French population, distinct life history variables covary across both allocation domains in a theoretically coherent manner. The latent variable indeed contrasts individuals exhibiting i) traits suggestive of a lesser investment in their soma (smoking, lower self-reported health status, efforts in looking after one's health and higher BMI) and ii) a faster reproductive strategy (earlier sexual debut, age at first birth and higher number of sexual partners), with individuals showing the opposite characteristics. Furthermore, childhood harshness predicted scores reflecting faster strategies, which is consistent with our interpretation of this latent variable as the fast-slow continuum.

Thus, the emerging covariation pattern fits well with the idea of a broad fast-slow axis of life history variation. Nonetheless, it should be noted that the fast-slow continuum captures only a fraction of the variance in individual life history traits and that it will not necessarily be relevant for all of them. Such a result is not unexpected since life history traits are undoubtedly under the influence of multiple unmeasured causes, which can lead individuals' allocation strategies to depart from typical fast or slow combinations of traits. For example, BMI is sensitive to genetic factors (Locke et al., 2015) and its relationship with the intensity of physical activity is not completely linear, e.g., athletes tend to have high BMI but low percentage of body fat (Etchison et al., 2011). In such cases, the associated variance in BMI will not be captured by the fast-slow continuum and might even correlate in opposite directions with other life history indicators.

Similarly, many determinants of fertility might isolate it from the fast-slow continuum in developed countries and explain why *number of children* did not correlate with the latent variable. In our representative sample of a country where the demographic transition has long been completed, mean fertility is indeed close to two children per parent with little variation around this value. Cultural factors such as easy access to contraceptives, universal health care for both the child and the parents, widespread access to wage labor via economic markets for women and highly shared norms about family size might for example explain why fertility is disconnected from the fast-slow continuum (Colleran, 2016; Lawson and Borgerhoff Mulder, 2016). Eventually, several meaningful axes of variation are likely to emerge once one tries to capture finer inter-individual differences in life strategies across human populations and to identify particular socioecological factors that call for more diverse clusters of allocation strategies (Del Giudice, 2014b).

In addition, our analysis also suggests that the calibration of life history strategies might be particularly sensitive to specific events. Indeed, the composite model revealed that when all harshness predictors were considered independently, only *violence in the family* contributed significantly to its effect on the latent variable, with marginal effects of *having lived with a stepfather* and *death of the mother*. These differential effects of harshness items could be interpreted in the light of theoretical models of adaptive developmental plasticity (Frankenhuis and Panchanathan, 2011a, 2011b). These models predict that the reliability (*i.e.* the strength of the association between a cue and a particular state of the environment) of the cues used by an organism to adjust its developmental trajectory should influence the timing and the rigidity of the organism's commitment to a particular life strategy. Therefore, a higher cue reliability compared to other harsh events could be one property of the item *violence in the family*, beyond the fact that it is a particularly strong measure of low parental care. This would be the case if such violent behaviours from the caregivers have been more reliably associated with future harsh conditions over human evolutionary history compared to other types of harsh events. Alternatively, the absence of significant coefficients for rare events such as *death of mother*, *death of father* or *long illness* could be due to the small number of positive realizations in our sample, which might prevent the detection of meaningful effects.

Regarding the influence of childhood harshness on life history strategies, it should also be stated that the correlational nature and the cross-sectional design of the current study do not allow strong inferences on the causal role of early adversity on future life history strategies. Indeed, the influence of the environment experienced later in development on life strategies could not be controlled for. However, several longitudinal studies in adolescents have already found that both early and later environments predict individuals' life strategies (Brumbach et al., 2009; Simpson et al., 2012). Moreover, recent works, studying the effect of external shocks (famine, epidemics, war, etc.) during fetal life and early childhood, have demonstrated that lack of resources has detrimental and durable effects on a range of outcomes later in life: health problems (Lin and Liu, 2014), attention deficits (de Rooij et al., 2010), anti-social behaviours (Neugebauer, 1999), lower educational level (Lavy et al., 2016), or lower probability of being married and getting a job (Almond et al., 2007). Hence, while life-history strategies remain flexible in the face of new information, at least part of the effect of childhood harshness measured here might reflect conditional adaptations to early life conditions.

Thus, despite the caveats mentioned above, the overall pattern measured in this study is consistent with the idea that people form coherent life history strategies that can be partly

captured by a fast-slow continuum and shaped by early experience of harsh events. Such a general pattern in a developed country is not easily explained without adopting an evolutionary developmental perspective (Frankenhuis et al., 2016) and it will therefore be interesting to extend this work. For instance, future research should identify which fast and slow strategies hold or vary across the broader range of situations encountered by humans. One promising direction could be to implement statistical techniques such as SEMs with composite variables in longitudinal designs or capitalizing on relevant natural experiments. This way one could assess the respective contributions to the development of fast strategies of different harsh events measured at various time points during ontogeny.

To conclude, our results support the relevance of adopting an evolutionary framework to explore patterns of individual differences within and across human populations. Our study also highlights the relevance of approaches that consider whole suites of behaviors rather than single outcomes in order to test functional hypotheses related to Life History Theory. More importantly perhaps, this framework puts forward a different way of construing important behavioural obstacles to health improvement in developed countries. Indeed, while vaccination, antibiotics and improved sanitation have greatly increased life expectancy, this process based on technological advances may have reached its limits. Recent works indeed suggests that the maximum lifespan of humans is subject to natural constraints (Dong et al., 2016). By contrast, many years of life are still lost due to lifestyle factors, in particular in middle and lower social classes. Moreover, while the most important health issues in the 20th were due to infectious pathogens, the most important health issues of the 21st century are primarily due to “lifestyle” decisions (dietary risks, high body-mass index, and tobacco smoking). Despite these evidence, behaviour-related causes of health are still ill-understood. For most people, dietary risks, high body-mass index, and tobacco smoking are seen as the result of lifestyle choices over which individuals have control (Hallsworth et al., 2016). Instead, the framework we put forward in this paper suggests that part of the variance observed in these at-risk behaviours can be traced back to evolved mechanisms geared to maximize short-term rewards over long-term investments in an environment that is perceived as dangerous.

Acknowledgments

This study was supported by the Institut d'Études Cognitives (ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*), the Institut national de la santé et de la recherche médicale (INSERM), an ERC consolidator grant (n°8657) and an “Action Incitative” funding from the

École Normale Supérieure. We are grateful to Anne L'Hôte for programming the survey and to Céline Dusautois for her help in preliminary analyses.

Conflict of interests: none.

Bibliography

- Almond, D., Edlund, L., Li, H., Zhang, J., 2007. Long-Term Effects Of The 1959-1961 China Famine: Mainland China and Hong Kong (No. w13384). National Bureau of Economic Research, Cambridge, MA.
- Baldini, R., 2015. Harsh environments and “fast” human life histories: What does the theory say? *bioRxiv* 14647. doi:10.1101/014647
- Belsky, J., 2012. The Development of Human Reproductive Strategies: Progress and Prospects. *Curr. Dir. Psychol. Sci.* 21, 310–316. doi:10.1177/0963721412453588
- Belsky, J., Steinberg, L., Draper, P., 1991. Childhood Experience, Interpersonal Development, and Reproductive Strategy: An Evolutionary Theory of Socialization. *Child Dev.* 62, 647. doi:10.2307/1131166
- Brumbach, B.H., Figueredo, A.J., Ellis, B.J., 2009. Effects of Harsh and Unpredictable Environments in Adolescence on Development of Life History Strategies: A Longitudinal Test of an Evolutionary Model. *Hum. Nat.* 20, 25–51. doi:10.1007/s12110-009-9059-3
- Buuren, S. van, Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* 45. doi:10.18637/jss.v045.i03
- Chisholm, J.S., Ellison, P.T., Evans, J., Lee, P.C., Lieberman, L.S., Pavlik, Z., Ryan, A.S., Salter, E.M., Stini, W.A., Worthman, C.M., 1993. Death, Hope, and Sex: Life-History Theory and the Development of Reproductive Strategies [and Comments and Reply]. *Curr. Anthropol.* 34, 1–24. doi:10.1086/204131
- Colleran, H., 2016. The cultural evolution of fertility decline. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150152. doi:10.1098/rstb.2015.0152
- de Rooij, S.R., Wouters, H., Yonker, J.E., Painter, R.C., Roseboom, T.J., 2010. Prenatal undernutrition and cognitive function in late adulthood. *Proc. Natl. Acad. Sci.* 107, 16881–16886. doi:10.1073/pnas.1009459107
- Del Giudice, M., 2014a. An Evolutionary Life History Framework for Psychopathology. *Psychol. Inq.* 25, 261–300. doi:10.1080/1047840X.2014.884918
- Del Giudice, M., 2014b. A Tower Unto Heaven: Toward an Expanded Framework for Psychopathology. *Psychol. Inq.* 25, 394–413. doi:10.1080/1047840X.2014.925339
- Dong, X., Milholland, B., Vijg, J., 2016. Evidence for a limit to human lifespan. *Nature* 538, 257–259. doi:10.1038/nature19793
- Ellis, B.J., 2004. Timing of Pubertal Maturation in Girls: An Integrated Life History Approach. *Psychol. Bull.* 130, 920–958. doi:10.1037/0033-2909.130.6.920
- Ellis, B.J., Essex, M.J., 2007. Family Environments, Adrenarche, and Sexual Maturation: A Longitudinal Test of a Life History Model. *Child Dev.* 78, 1799–1817. doi:10.1111/j.1467-8624.2007.01092.x
- Ellis, B.J., Figueredo, A.J., Brumbach, B.H., Schlomer, G.L., 2009. Fundamental Dimensions of Environmental Risk: The Impact of Harsh versus Unpredictable Environments on the Evolution and Development of Life History Strategies. *Hum. Nat.* 20, 204–268. doi:10.1007/s12110-009-9063-7
- Etchison, W.C., Bloodgood, E.A., Minton, C.P., Thompson, N.J., Collins, M.A., Hunter, S.C., Dai, H., 2011. Body Mass Index and Percentage of Body Fat as Indicators for Obesity in an Adolescent Athletic Population. *Sports Health* 3, 249–252. doi:10.1177/1941738111404655

- Frankenhuis, W.E., Panchanathan, K., 2011a. Balancing sampling and specialization: an adaptationist model of incremental development. *Proc. R. Soc. B Biol. Sci.* 278, 3558–3565. doi:10.1098/rspb.2011.0055
- Frankenhuis, W.E., Panchanathan, K., 2011b. Individual Differences in Developmental Plasticity May Result From Stochastic Sampling. *Perspect. Psychol. Sci.* 6, 336–347. doi:10.1177/1745691611412602
- Frankenhuis, W.E., Panchanathan, K., Nettle, D., 2016. Cognition in harsh and unpredictable environments. *Curr. Opin. Psychol.* 7, 76–80. doi:10.1016/j.copsyc.2015.08.011
- Geronimus, A.T., Bound, J., Waidmann, T.A., 1999. Health inequality and population variation in fertility-timing. *Soc. Sci. Med.* 1982 49, 1623–1636.
- Grace, J.B., Anderson, T.M., Olff, H., Scheiner, S.M., 2010. On the specification of structural equation models for ecological systems. *Ecol. Monogr.* 80, 67–87.
- Grace, J.B., Bollen, K.A., 2008. Representing general theoretical concepts in structural equation models: the role of composite variables. *Environ. Ecol. Stat.* 15, 191–213. doi:10.1007/s10651-007-0047-7
- Griskevicius, V., Ackerman, J.M., Cantu, S.M., Delton, A.W., Robertson, T.E., Simpson, J.A., Thompson, M.E., Tybur, J.M., 2013. When the Economy Falters, Do People Spend or Save? Responses to Resource Scarcity Depend on Childhood Environments. *Psychol. Sci.* 24, 197–205. doi:10.1177/0956797612451471
- Griskevicius, V., Delton, A.W., Robertson, T.E., Tybur, J.M., 2011. Environmental contingency in life history strategies: The influence of mortality and socioeconomic status on reproductive timing. *J. Pers. Soc. Psychol.* 100, 241–254. doi:10.1037/a0021082
- Hurst, J., Sassi, F., 2008. The Prevention of Lifestyle-Related Chronic Diseases (OECD Health Working Papers No. 32).
- Kaplan, H., Lancaster, J., Robson, A., 2003. Embodied Capital and the Evolutionary Economics of the Human Life Span. *Popul. Dev. Rev.* 29, 152–182.
- Lavy, V., Schlosser, A., Shany, A., 2016. Out of Africa: Human Capital Consequences of In Utero Conditions (No. w21894). National Bureau of Economic Research, Cambridge, MA.
- Lawson, D.W., Borgerhoff Mulder, M., 2016. The offspring quantity–quality trade-off and human fertility variation. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150145. doi:10.1098/rstb.2015.0145
- Li, K.-H., Meng, X.-L., Raghunathan, T.E., Rubin, D.B., 1991. SIGNIFICANCE LEVELS FROM REPEATED p-VALUES WITH MULTIPLY-IMPURED DATA. *Stat. Sin.* 1, 65–92.
- Lin, M.-J., Liu, E., 2014. Does in utero Exposure to Illness Matter? The 1918 Influenza Epidemic in Taiwan as a Natural Experiment (No. w20166). National Bureau of Economic Research, Cambridge, MA.
- Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., et al., 2015. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–206. doi:10.1038/nature14177
- Low, B.S., Hazel, A., Parker, N., Welch, K.B., 2008. Influences on Women’s Reproductive Lives: Unexpected Ecological Underpinnings. *Cross-Cult. Res.* 42, 201–219. doi:10.1177/1069397108317669
- McCullough, M.E., Pedersen, E.J., Schroder, J.M., Tabak, B.A., Carver, C.S., 2012. Harsh childhood environmental characteristics predict exploitation and retaliation in humans. *Proc. R. Soc. B Biol. Sci.* 280, 20122104–20122104. doi:10.1098/rspb.2012.2104

- Nettle, D., 2011. Flexibility in reproductive timing in human females: integrating ultimate and proximate explanations. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 357–365. doi:10.1098/rstb.2010.0073
- Nettle, D., 2010. Why Are There Social Gradients in Preventative Health Behavior? A Perspective from Behavioral Ecology. *PLoS ONE* 5, e13371. doi:10.1371/journal.pone.0013371
- Nettle, D., Coall, D.A., Dickins, T.E., 2009. Birthweight and paternal involvement predict early reproduction in British women: Evidence from the National Child Development Study. *Am. J. Hum. Biol.* NA-NA. doi:10.1002/ajhb.20970
- Nettle, D., Cockerill, M., 2010. Development of Social Variation in Reproductive Schedules: A Study from an English Urban Area. *PLoS ONE* 5, e12690. doi:10.1371/journal.pone.0012690
- Nettle, D., Frankenhuys, W.E., Rickard, I.J., 2012. The adaptive basis of psychosocial acceleration: Comment on Beyond Mental Health, Life History Strategies articles. *Dev. Psychol.* 48, 718–721. doi:10.1037/a0027507
- Neugebauer, R., 1999. Prenatal Exposure to Wartime Famine and Development of Antisocial Personality Disorder in Early Adulthood. *JAMA* 282, 455. doi:10.1001/jama.282.5.455
- Pepper, G.V., Nettle, D., 2014. Perceived Extrinsic Mortality Risk and Reported Effort in Looking after Health: Testing a Behavioral Ecological Prediction. *Hum. Nat.* 25, 378–392. doi:10.1007/s12110-014-9204-5
- Quinlan, R.J., 2010. Extrinsic Mortality Effects on Reproductive Strategies in a Caribbean Community. *Hum. Nat.* 21, 124–139. doi:10.1007/s12110-010-9085-1
- Reale, D., Garant, D., Humphries, M.M., Bergeron, P., Careau, V., Montiglio, P.-O., 2010. Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 4051–4063. doi:10.1098/rstb.2010.0208
- Roff, D.A., 2002. Life history evolution. Sinauer Associates, Sunderland, Mass.
- Rosseel, Y., 2012. lavaan : An R Package for Structural Equation Modeling. *J. Stat. Softw.* 48. doi:10.18637/jss.v048.i02
- Rubin, D.B., 2004. Multiple imputation for nonresponse in surveys, Wiley classics library. Wiley-Interscience, Hoboken, N.J.
- Simpson, J.A., Griskevicius, V., Kuo, S.I.-C., Sung, S., Collins, W.A., 2012. Evolution, stress, and sensitive periods: The influence of unpredictability in early versus late childhood on sex and risky behavior. *Dev. Psychol.* 48, 674–686. doi:10.1037/a0027293
- Stearns, S.C., 2000. Life history evolution: successes, limitations, and prospects. *Naturwissenschaften* 87, 476–486. doi:10.1007/s001140050763
- Stearns, S.C., 1992. The evolution of life histories. Oxford University Press, Oxford ; New York.
- Stringhini, S., 2010. Association of Socioeconomic Position With Health Behaviors and Mortality. *JAMA* 303, 1159. doi:10.1001/jama.2010.297
- Tither, J.M., Ellis, B.J., 2008. Impact of fathers on daughters' age at menarche: A genetically and environmentally controlled sibling study. *Dev. Psychol.* 44, 1409–1420. doi:10.1037/a0013065
- Walker, R., Gurven, M., Hill, K., Migliano, A., Chagnon, N., De Souza, R., Djurovic, G., Hames, R., Hurtado, A.M., Kaplan, H., Kramer, K., Oliver, W.J., Vaegleggia, C., Yamauchi, T., 2006. Growth rates and life histories in twenty-two small-scale societies. *Am. J. Hum. Biol.* 18, 295–311. doi:10.1002/ajhb.20510

Table 1 Descriptive statistics and correlations from self-report data (n=1015).

	BMI	Health status	Health effort	Smoking	Number of children	Age at first birth	Sexual debut	Short-term partners	Harshness	Age
BMI	-									
Health status	-0.20*	-								
Health effort	-0.12*	0.18*	-							
Smoking	0.04	-0.09*	-0.12*	-						
Number of children	0.13*	-0.00	0.01	-0.05	-					
Age at first birth	-0.09*	0.13*	-0.00	-0.05	-0.37*	-				
Sexual debut	-0.03	0.04	0.06*	-0.20*	0.04	0.18*	-			
Short-term partners	0.04	-0.00	-0.10*	0.19*	-0.07*	0.10*	-0.22*	-		
Harshness	0.05	-0.14*	-0.07*	0.12*	0.04	-0.10*	-0.09*	0.04	-	
Age	0.11*	-0.14*	0.21*	0.04	0.17*	-0.15*	0.17*	0.07*	-0.06*	-
Mean	26.04	2.75	68.98	0.21	2.15	25.66	18.57	6.21	-0.06	52.52
SD	5.15	0.70	19.01	0.25	0.96	4.86	3.13	12.42	12.36	14.33
Range	13.6-58.8	1-4	0-100	0-0.91	1-7	14-56	10-61	0-160	-17.2 - 66.6	19-87

* $p < 0.05$

Table 2 Statistical and practical fit indices for the structural equation models.

Model	χ^2	df	<i>p</i>	RMSEA	SRMR	CFI
Model 1	58.6	15	<0.001	0.046	0.030	0.93
Model 2	191	113	<0.001	0.026	0.028	0.83

Table 3 Parameter estimates, standard errors and their significance for measurement and structural coefficients of model 1 and composite weights of model 2.

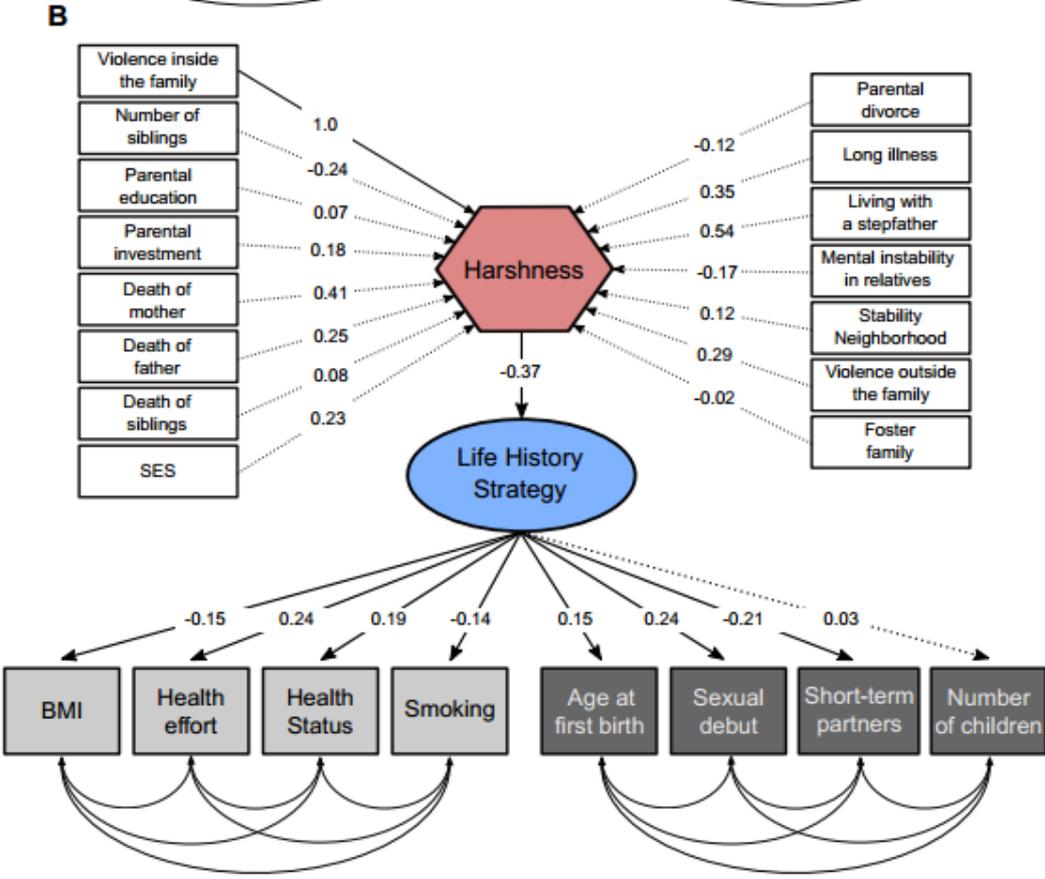
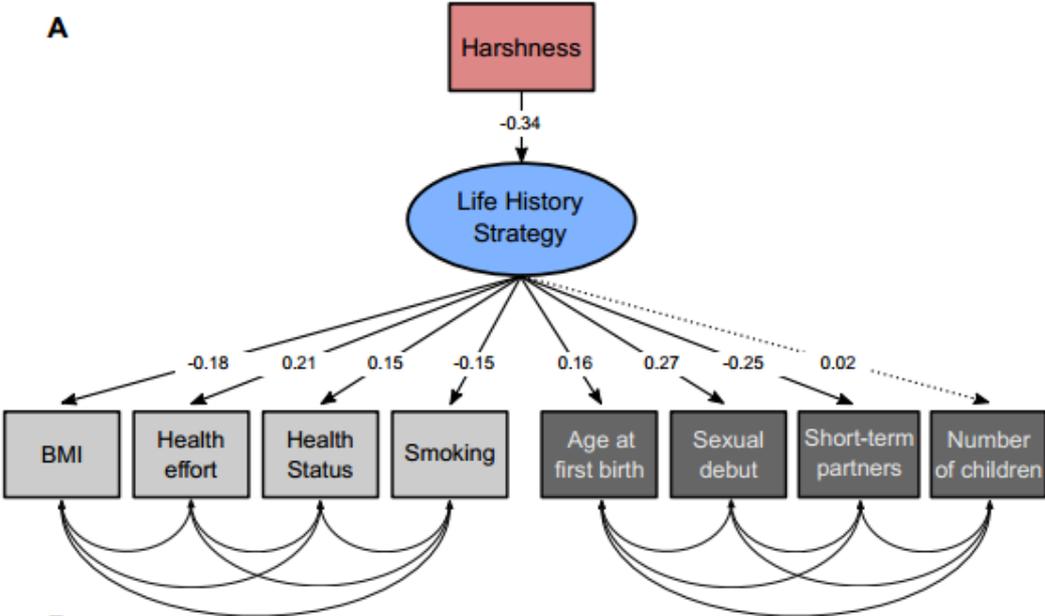
parameter	value	se	Z	p	standardized
LHS → BMI	-0.85	0.34	-2.48	0.01	-0.18
LHS → Health effort	3.76	1.33	2.83	0.005	0.21
LHS → Health status	0.21	0.05	4.16	<0.001	0.15
LHS → Smoking	-0.14	0.02	-6.54	<0.001	-0.15
LHS → Number of children	0.03	0.04	0.62	0.54	0.02
LHS → Age at first birth	0.79	0.23	3.50	<0.001	0.16
LHS → Sexual debut	0.92	0.14	6.48	<0.001	0.27
LHS → Short-term partners	-2.92	0.83	-3.52	<0.001	-0.25
Harshness → LHS	-0.06	0.01	-5.94	<0.001	-0.34
Harshness ^a → LHS	-0.25	0.07	-3.88	<0.001	-0.37
Violence family → LHS ^b	-0.73	0.19	-3.88	<0.001	-0.24
SES → Harshness ^a	0.23	0.22	-1.07	0.29	0.23
Parental investment → Harshness ^a	0.18	0.25	-0.72	0.48	0.18
Stability neighborhood → Harshness ^a	0.12	0.21	0.57	0.57	0.12
Parental education → Harshness ^a	0.07	0.26	0.28	0.78	0.07
Number of siblings → Harshness ^a	-0.24	0.20	-1.19	0.24	-0.24
Death of father → Harshness ^a	0.25	0.20	1.24	0.21	0.25
Death of mother → Harshness ^a	0.41	0.21	1.92	0.06	0.41
Death of siblings → Harshness ^a	0.08	0.21	0.36	0.72	0.08
Living with a stepfather → Harshness ^a	0.54	0.29	1.88	0.06	0.29
Parental divorce → Harshness ^a	-0.12	0.28	-0.41	0.69	-0.12
Mental instability in relatives → Harshness ^a	-0.17	0.19	-0.92	0.35	-0.17
Violence outside the family → Harshness ^a	0.29	0.23	1.27	0.20	0.29
Long illness → Harshness ^a	0.35	0.23	1.53	0.13	0.35

Foster family → Harshness ^a	-0.016	0.17	-0.09	0.93	-0.02
--	--------	------	-------	------	-------

^a Harshness in this case refers to the composite variable in model 2

^b All values for this effect are computed from the partially reduced form of model 2

Fig. 1. Path diagrams of the structural models fitted showing for clarity only the structure of the residual covariances in the measurement model and standardized regression weights for harshness and reflective indicators. Significant paths at the 5% level are represented with a plain arrow **a**. Model 1, harshness is modelled as a weighted sum **b**. Model 2, harshness is modelled as a latent composite.



Social information impairs reward learning in depressive subjects: behavioral and computational evidence

Lou Safra^{1,2,3}, Coralie Chevallier^{1,2,3}, Sarah-Jayne Blakemore⁴, Stefano Palminteri^{1,2,3}

¹*Laboratoire de Neurosciences Cognitives, Institut National de la Santé et de la Recherche Médicale, Paris, France*

²*Département d'Études Cognitives, Ecole Normale Supérieure, Paris, France*

³*Université de Recherche Paris Sciences et Lettres, Paris, France*

⁴*Institute of Cognitive Neuroscience, University College London, London, UK.*

Abstract

Depression has been associated with abnormal reward learning and decision-making. Besides general reward processing deficits, depression is also associated with atypical social cognition. Although explaining these differences is critical to improve the understanding of depression and its treatment, the cognitive mechanisms underlying atypical decision-making in social contexts in depression are still underspecified. In the present paper, we specifically investigate these cognitive processes. To do so, we presented two cohorts of subjects drawn from a general population ($N = 50$ each) with a task involving reward learning in a social context with different levels of social information (absent, partial and complete), while concomitantly assessing depression, anxiety and schizotypy. A general linear model revealed a specific detrimental effect of depression on behavioral performance in the presence of social information, specifically when the subjects were informed about the choices of another player. Importantly, the effect of depression was robust after controlling for participants' anxiety and schizotypy levels. Computational model-based analyses allowed us to further characterize this deficit as a reduction of the integration of self-generated information in social contexts, rather than a social learning deficit *per se*. To conclude, our results shed light on the computational mechanisms underlying the interaction between socio-cognitive and reward learning and decision-making in depressive patients.

Main text

In addition to general mood disorders, depression is associated with atypical decision-making (1). For instance, depressive patients reports higher levels of anhedonia and decisional conflict as well as lower levels of self-efficacy (1). Experimental data have confirmed these decision-making atypicalities and have shown that depression is notably associated with decreased responsiveness to rewards and decreased processing of counterfactual information (2, 3). These studies provide precious insights into cognitive functioning in depression, but atypical decision-making in this condition may not be fully explained by differences in private decision-making processes. Indeed, depression occurs in interpersonal contexts (4) and depressive symptoms have a pervasive effect on decision-making in social contexts. For instance, depressive patients are more ready to modulate their behavior after receiving an advice but not necessarily to match the advised behavior (5). In the clinical domain, depressive patients are known to be less compliant to physicians' recommendations such as medical treatments, which may play a role in the association between depression and worse long-term prognosis of health diseases (6, 7). Understanding the effect of the social context on decision-making in depression can thus provide important insights on how to increase the efficiency of depressive patients care (8). Despite this clinical relevance, the cognitive mechanisms underlying differences in decision-making in a social context are still underspecified.

Three main cognitive processes may explain the difference in social context decision-making associated with depressive symptoms. First, depression may be linked with differences in the way social information is processed (9). In line with this idea, a recent paper by Hofheinz et al. (2017) has shown that depression is associated with differences in advice-taking (5). However, access to others' actions does not only increase the amount of information an individual has access to but also modifies the weight she grants to one's own action. Indeed, weighting of information sources depending on their reliability has been theorized as a necessary feature of efficient decision-making using socially provided information (10, 11). Accordingly, explicit instructions about actions' value has been shown to significantly decrease the reliance on self-experienced outcomes (12, 13).

Differences in the valuation of their own compared to others' action may thus explain atypical social influence in depression. Indeed, depressed participants report lower levels of confidence (14, 15) as well as a higher perceived sense of control for others than for themselves (16). Finally, the presence of others by itself also disrupts cognitive functioning and may explain differences in decision-making in a social context (17–19). For instance, Conty et al. (2010) have shown that performances in a stroop task are significantly decreased when participants are exposed to open eyes, signaling the observation of another individual, compared to a condition in which they see closed eyes (19). It has been proposed that these effects reflect the allocation of attentional resources to the processing of the social stimulus (20). Since depression is associated with lower cognitive

functioning in general (21), it is possible that the mere presence of others exacerbates these differences by capturing already scarcer attentional resources.

The aim of the present paper is to disentangle these different hypotheses. To do so, we asked participants to complete a social version of a two-arms bandit task (22) with three levels of social information: absent, partial and complete.

Results

Experimental protocol

In order to test the influence of depressive symptoms on decision-making in a social context, we relied on online participant samples. These samples provide diverse cohorts of participants, which are particularly suited to investigate the influence of psychiatric traits on cognitive mechanisms, and notably on those involved in decision-making (23–26). We tested two independent cohorts of 50 Amazon MTurkers in which depression scores ranged from totally absent to sure cases of depression (depression score superior to 10, Table 1) (27), with a good internal consistency of the three psychiatric trait scales that were administrated (depression and anxiety: all alphas > .85; schizotypy: both alphas > .75 in the two samples).

To arbitrate between the different mechanisms that can account for differences in decision-making in a social contexts, participants were paired with a virtual partner and performed a probabilistic instrumental learning task in three social contexts: a Private condition in which participants

performed the task individually with no access to their partner's decisions, an Action condition in which the information about their partners' actions was available, and an Action+Outcome condition in which participants had access to their partner's actions and outcomes (Figure 1A).

As expected, participants selected the most rewarded symbol above chance in the three conditions (Private: $M = 0.65 \pm 0.03$, $t(99) = 11.42$, $p < .001$; Action : $M = 0.64 \pm 0.03$, $t(99) = 11.63$, $p < .001$; Action + Outcome: $M = 0.67 \pm 0.03$, $t(99) = 12.36$, $p < .001$; Figure 1A). In addition, a trial-by-trial regression confirmed that participants responded to factors manipulated in each condition: participants' previous choice and reward (Private, Action and Action+Outcome conditions) as well as their partners' choice (Action and Action+Outcome conditions) and reward (Outcome conditions; all $ps < .001$; Figure 1B). Finally, significant effects of post-test ratings on their partner's avatar trustworthiness, dominance and competence on participants' behavior suggested that the provided information was computed as a socially rather than as a non-socially provided information (see Supplementary Materials).

Behavioral results

The effect of depressive symptoms on the rate of correct choices was measured using a mixed logistic regression taking Block (Stable vs Reversal) and Condition (Private vs Action vs Action+Outcome) as within-subject factors and depression scores as well as partner's performance as between-subject variables. Higher partner's percentage of correct choices were

associated with higher participants' rates of correct choices in two social conditions (Action: $b_{Partner*Action} = 0.16 \pm 0.10$, $z = 3.27$, $p = .001$; Action + Outcome: $b_{Partner*Action} = 0.16 \pm 0.11$, $z = 2.78$, $p = .005$) but not in the Private condition ($b_{Partner} = 0.01 \pm 0.07$, $z = 0.14$, $p > .250$; Figure 2A). In addition, a significant effect of depressive symptoms was evidenced such that more depressed participants showed lower rates of correct choices in the Action condition compared to the Private condition ($b_{Dep*Action} = -0.19 \pm 0.10$, $z = -3.81$, $p < .001$; no other significant effect was evidenced: all $ps > .250$; Figure 2A). Importantly, this effect of depressive symptoms was robust to the inclusion of anxiety and schizotypy, two comorbid traits of depression (28, 29), as controlling factors ($b_{Dep*Action} = -0.22 \pm 0.12$, $z = -3.51$, $p < .001$; no other significant effect evidenced: all $ps > .250$).

Finally, we tested whether the difference in the rate of correct choices between the Action and Private conditions could accurately identify the participants classified as sure cases of depression on the depressions scale. Classification analyses revealed that the difference in correct choice identified the sure cases of depression disorders with a good accuracy of $78 \pm 0.5\%$ and both few false negative and false positive cases as testified by the good a sensitivity ($81 \pm 0.7\%$) and specificity ($59 \pm 1.7\%$) of the classifier (Figure 2B).

Computational results

Although the effect of depressive symptoms on learning in the Action condition is robust, the functional mechanisms underlying this effect are not

clear. The effect of depressive symptoms could either be due to differences in the processing of socially provided information or to differences in the weighting of self-provided information when secondhand information is also available. To arbitrate between these two possibilities, we fitted an observational reinforcement learning model inspired by a validated model of social learning (22, 30). In addition to a classical Q-learning algorithm, this model allowed participants to update their choice probability depending on their partner's choice in the Action condition and to update the value attributed to each symbol depending on the partner's reward in the Action+Outcome condition. Compared to the original model developed by Burke et al. (2010) (22), we allowed participants to have different Q-learning parameters in the Private and in the two Social Information conditions (Action and Action+Reward conditions; Figure 3) in order to test for potential differences in the integration of self-generated information in the presence of secondhand information.

Parameter recovery

As we were interested by the modulation of specific parameters by depression scores we first tested whether our task allowed us to successfully retrieve correlation between parameters and individuals' scores (31). To do so, we ran 100 sets of simulations for each parameter, each simulating 100 participants, with the parameter of interest correlating with an arbitrary variable and the other parameters being randomly fixed in a defined range. The simulated data were then fitted using our social learning model. The

manipulated correlations were significantly retrieved in 65 and 90% of the simulations set. Importantly, the recovery of the correlations was specific to the manipulated parameter with false alarms detected in less than 10 of the cases except for the alpha-beta in the private condition where false alarms were found in 36% of the simulations (Figure 4A).

Computational effects of depressive symptoms

We then applied the same analysis to the parameters fitted on participants' actual behavior using depression scores as the independent variable. Higher depression scores were specifically associated with lower learning rates in the Social Information conditions ($b_{Dep} = -0.07 \pm 0.06$, $z = -2.38$, $p = .017$; other $ps > .200$; after controlling for anxiety and schizotypy: $b_{Dep} = -0.01 \pm 0.01$, $z = -1.733$, $p = 0.083$; other effects of depression: all $ps > .250$; Figure 4B). Importantly, data simulated using the fitted parameters accurately recovered the decrease in performance associated with depression scores between the Action and Private condition ($b_{Dep*Action} = -0.01 \pm 0.01$, $t(492) = -3.32$, $p = .001$; Figure 4D) confirming that this effect cannot simply be due to specific task contingencies. In addition, in line with a specific effect of depressive symptoms on learning rates in the Social Information conditions, a trial-by-trial regression revealed that higher depression scores were associated with a decreased sensitivity to the previous reward in the Action condition ($b_{Dep} = -0.11 \pm 0.05$, $z = -4.01$, $p < .001$; Figure 1D) which was found in the simulated data ($b_{Dep} = -0.01 \pm 0.01$, $z = -2.53$, $p = .011$).

Decreased learning rates as a diagnostic tool

Interestingly, high depression scores were not solely associated with decreased learning rates in the Social Information conditions, but also with decreased learning rates in the Social Information conditions *compared* to the Private condition ($b = -0.08 \pm 0.07$, $t(98) = -2.226$, $p = .028$; Figure 4C) indicating that the presence of socially provided information decreased the learning rate of the most depressed participants. Adopting a computational psychiatry approach, we tested whether this difference in learning rates between the Private and Social Information conditions could identify the sure cases of depression detected by the questionnaire be used as a diagnostic tool for depression. The difference in learning rates detected sure cases of depression with good accuracy ($78 \pm 1\%$) and good sensitivity ($84 \pm 1\%$) but a low specificity ($49 \pm 1.7\%$). Comparing this classifier with the correct choice rate-based classifier revealed that the two classifiers were equally accurate at detecting sure cases of depression ($t(198) = 0.222$, $p > .250$) but that the classifier based on computational parameters was more sensitive ($t(198) = 2.29$, $p = .023$) but less specific ($t(198) = -5.05$, $p < .001$) than the classifier based on correct choice rates.

Discussion

In this paper, we show that depressive symptoms are associated with specific impairments in value-based decision-making in a social context. More precisely, more depressed participants are less accurate in the presence of information about others' behavior, a difference due to a

decreased integration of self-generated information when socially provided information is available. Importantly, this effect of depression is specific in two ways: first, it is specific to depression as it is robust to the inclusion of anxiety and schizotypy, two comorbid traits of depression, as control variables (28, 29) and it holds a good diagnostic value to detect depression cases; second, it is specific to the inclusion of secondhand information as no behavioral differences was evidenced when only self-generated information was available. These results support the hypothesis that differences in decision-making in social contexts associated with depressive symptoms are neither due to the disruptive presence of others nor to differences in the integration of socially-provided information but rather to a decreased integration of self-provided information when socially provided information is also available.

This effect of depressive symptoms in the presence of social information may be explained by differences in the perceived reliability of their own and of others' actions. Indeed, depressive patients are known to report a decreased sense of control and lower levels of confidence in their own actions but also a higher sense of control in others' actions (14–16). More generally, negative perception of self and negative comparison to others are core symptoms of depression (32). Therefore, it is possible that the most depressed participants perceived their partner's behavior as more reliable, thus underweighting the information they acquired through their own experience. In line with this idea, this decreased integration of self-generated information when socially provided information is available

closely matches the modulation of decision making in the presence of instructional information found in the general population (12, 13). However, although the influence of partners' avatar evaluations on task completion suggests that our task is indeed perceived as a social task, it is possible that the influence of social information on behavior actually reflects a more general effect of any secondhand information, be it socially or asocially provided.

Indeed, it is possible that the presence of additional information more generally disrupts the integration of self-provided information. Indeed, the virtual partner's behavior in our task was also subjected to the task contingencies, with the probability of obtaining the reward after making the right choice at around 75% across blocks. The partner's actions and outcomes may thus have been conflicting with the participants' experience in some trials, resulting in a higher demand in cognitive resources. Under this view, the under-integration of participants' own experience for decision-making may be linked with the decreased cognitive functioning associated with depression (21). However, why would, in this case, secondhand information be favored compared to firsthand information is still an open question. Further experiments are needed to address the question of whether the observed effect is due to differences in the perceived reliability of the information or to the availability of additional information, as well as to assess the impact of the information source (i.e., social vs non social). In particular, in a dimensional approach to psychiatric conditions, it would particularly interesting to investigate which trait of

depression drives these behavioral differences and notably the respective effect self perception or cognitive conflict resolution deficits.

In addition, our experiment provides new evidence that private reinforcement learning atypicalities in depression are not clear (33), and suggests that the differences in learning rates associated with depressive symptoms may only arise in special contexts. Our results are thus in line with the idea that reward perception, action valuation and reward sensitivity should be dissociated and investigated separately to better understand anhedonic traits in depression (34, 35). However, it is worth noting that we did not control for participants' actual diagnosis and treatment, which may be problematic since medication interacts with decision-making in depression (36). Therefore, our results would benefit from being replicated in carefully controlled population and notably by contrasting a non-clinical population with depressive patients. This replication would allow us to further measure the diagnostic value of our behavioral task. Indeed, in the present study, we only tested its ability to detect depression cases identified by a self-rated scale (27) but it would be particularly interesting to measure whether such an objective measure of behavior is better than existing self-assessments at detecting clinically diagnosed cases of depression.

To summarize, our study exemplifies how computational approaches can provide new insights into cognitive processes differences associated with clinical conditions. Indeed, computational modeling demonstrated that the effect of depression was selective of the way individual information was

processed. At this point it is worth noting that these conclusions were only allowed after a careful testing of the ability of our task to precisely identify which model parameter would be influenced by depression (31). In addition to a better understanding of the cognitive mechanisms differences associated with clinical conditions, computational psychiatry can provide important information to practitioners. Indeed, instead of an 'imitation impairment' suggested by behavioral results, our computational modeling approach reveals that depression is associated with differences in the way private information is processed when social information is also available. In addition, we found that the processing of socially provided information is not decreased by depression symptoms. Therefore, our results suggest that decision-making impairments in depression, and notably concerning medical treatment compliance, may be resolved by providing more secondhand information about the recommended actions and their outcomes.

Finally, and more generally, our study also provides new insights into the cognitive mechanisms underlying decision-making in a social context. Indeed, in line with theoretical work on social learning (10, 11), it shows that the observation of others' behavior does not only increase the amount of available information to the observer but also influences the processing her own experience. Therefore, our results extend previous findings on the influence of instructions on decision-making (12, 13) and suggest that the information weighting mechanism applies more widely to the access of any secondhand information. However, the factors guiding these information-

weighting mechanisms is still an open question. According to social learning theories (10, 11), and as suggested by our results the effect of depressive symptoms, weighting differences of first- and second-hand information may be guided by the perceived reliability of the different information sources. Our results thus open the path to a new type of investigations of observational learning, that integrates into the study of observational learning, the analysis of the weighting of firsthand and secondhand information as well as of the influence of different information sources (37, 38).

Materials and Methods

Experimental design

Participants performed the probabilistic instrumental learning task described in the Result section (Fig. 1A). The task was programmed on Qualtrics and was composed of six blocks. In each block, participants had to choose between two cues that were associated with different reward probabilities (20-80% or 30/70%). For instance, in a 20/80% pair, the most rewarded cue provided a positive outcome (+1 point) 80% of the times and a negative outcome (-1 point) 20% of the time, while the less rewarded cue provided a negative outcome 80% of the time and a positive outcome 20% of the time. Two given cues were always presented together and only in one block. Participants made their choice by pressing left or right (Q or P keys) and were given no explicit information on reward probabilities, which

they had to learn through trial and error. In addition, they were encouraged to accumulate as many points as possible, with their final amount of points being translated into bonus money at the end of the experiment.

Participants were paired with another player at the beginning of the experiment with which they played in turn at each trial. Importantly, in order to ensure that all the partners behaved similarly, the other player was in fact a robot, simulated using a reinforcement learning algorithm with an optimal set of parameters for this task ($\alpha = 0.5$, $\beta = 10$). Since depression is associated with differences in social cognition (39), the other player was represented by a neutral avatar, chosen to be generally perceived as neither dominant or submissive nor trustworthy or untrustworthy (40), while participants had to choose their own avatars in a set of 16 faces at the beginning of the task. Participants performed this task in three different social contexts with different amounts of social information: a Private condition in which they did not have access to their partner's behavior, an Action condition in which participants could see their partner's behavior but not their outcomes and an Action+Observation in which participants could observe their partner's decisions and outcomes. Importantly, participants performed each condition twice, one in which the cues reward probabilities were fixed and assigned a 30/70% contingency, and one in which they reversed at the middle of the block and were assigned to more extreme contingencies (20/80%). Finally, at the end of the experiment, participants rated their partner's avatar on three personality traits (trustworthiness,

dominance and competence) and completed the Hospital Anxiety and Depression Scale (27) as well as the Peters et al. Delusions Inventory (41).

Participants

Two independent cohorts of 50 American participants were recruited via Amazon Mechanical Turk to participate in this online study (Table 1). Each participant received a fixed amount 4\$ for completing the 40-minute task to which was then added the bonus money they earned during the experiment. Participant received a description of the study and signed an informed consent before starting the experiment.

Statistical analyses

Percentage of correct choices

The percentage of choices of the most rewarded cue was extracted for each block and used as a dependent variable. A mixed logistic regression with both random intercept and random slopes was conducted on correct choice rates taking participants' ID as a random factor, block type ('Stable' vs 'Reversal'), condition ('Private', 'Action' vs 'Action+Outcome') as within-subject variables and depression, anxiety and schizotypy scores as well as partner's performance as between-subject variables.

Identification of depression cases

Out of sample tests were used to assess the diagnostic value of our task. 50 participants were randomly extracted from the entire sample and used to optimize a classifier of sure cases of depression using either the difference in correct choice rates between the Action and the Private conditions or the difference in learning rates between the Private and Social Information conditions. The classifier and the associated optimal cut-off was tested on the 50 remaining participants. This operation was repeated 100 times in order to estimate the average accuracy, sensitivity and sensibility of the classifiers.

Trial-by-trial analyses

For each condition, the chosen actions and the information obtained at one trial were used to predict the choice at the subsequent trial. More precisely, in addition to the matching of the action with the choice at the preceding trial, a variable corresponding to the rationality of choosing an action based on the previous trial was computed such that this variable when participants chose the action rewarded at the previous trial or did not chose the action punished at the previous trial and equal to zero when participant persevered in a punished action or switched after receiving a reward. These two variables were used to predict participants' choice in the Private condition. In addition, in the Action condition, an additional variable corresponding to the matching between participant's choice and their

partner's choice at the preceding trial was also included. Finally, a similar combination of action and outcomes based on partner's actions and outcomes was also included in the Action+Outcome condition. The regressions were computed for each participant individually and the extracted marginal effects were then analyzed using structural equation modeling.

Computational analyses

Computational model

A social Q-learning model similar to the one presented by Burke et al. (2010) (22) was used to fit our behavioral data. The only change we made is the inclusion of different learning rates and temperature variables in the Private and Social Information conditions (see Result section for details). We optimized the model parameters by minimizing the negative log-posterior probability. This model was found to fit the data better than a simple Q-learning model without social influence (log-posterior probabilities: 85%; model attribution: 90%). Importantly, the purely individual Q-learning model fit the robot's behavior better than the social version of the model (log-posterior probabilities: 99%; model attribution: 99%). The obtained parameters were then analyzed using structural equation modeling to account for correlations between the different dependent variables.

Simulated data analysis

Finally, we assessed the ability of the model to recover the observed behavioral effect of depressive symptoms using model simulations (31). For each participant, we simulated behavioral data for each condition based on their best fitting parameters. Importantly, a simulated partner was also generated, such that the simulated data were completely independent of the contingencies actually experienced by the participants. This procedure was repeated 100 times, to avoid any effect of participant's and partner's history of choice and outcomes. Analysis of the recovered percentage of correct choices was ran on the averaged rates of correct choices across the 100 simulations using a linear mixed regression taking the exact same predictors as the mixed general linear model used for analyzing participants' percentage of correct choices. Similarly, trial-by-trial analyses were ran on each simulation independently and the marginal effects were averaged across the 100 simulations to be analyzed using the same structural equation model as the one used on actual behavioral data.

Figures

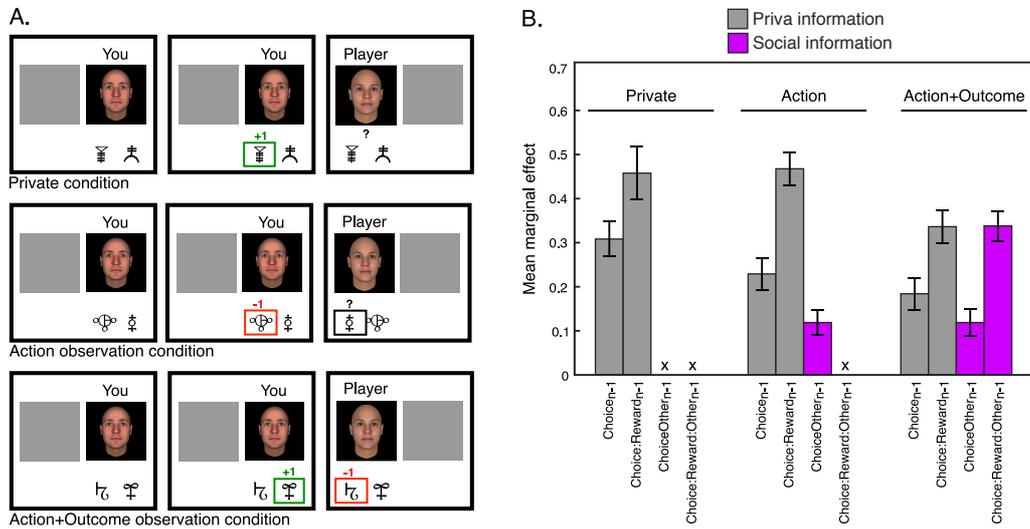


Figure 1. Social learning task

(A) Experimental protocol. In each condition, participants played in turn with a simulated partner with the same pair of symbols. After each choice, participants received a reward or a punishment. In the Private condition, participants did not see the choice or the outcome of their partner. In the Action observation condition, the choice of their partner was displayed at each trial. In the Action+Observation condition, both the choice and the outcome of the partner were displayed. **(B) Results of the trial-by-trial analysis.** Participants accurately integrated private information (choice and outcome; in gray) in the three conditions. In addition, in the Action observation condition, participants' choices were also predicted by the partner's choice and, in the Action+Outcome condition both the partner's choice and outcome predicted the participant's choice (in pink).

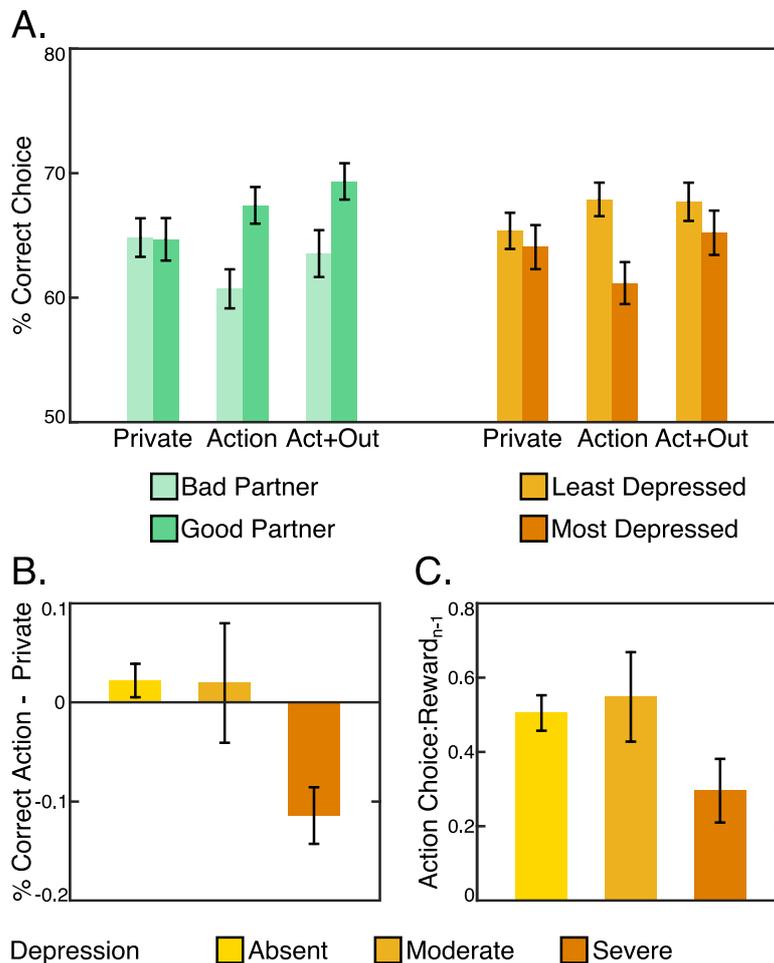


Figure 2. Behavioral results

(A) Effect of partner’s behavior and depression on social learning. Both the partner’s behavior and the participant’s depression score influenced the correct choice rate in the Action condition, such that participants paired with a robot with lower rates of correct choices (‘Bad robot’) had also lower rates of correct choices than those paired with a robot with higher rates of correct choices (‘Good robot’; green). Similarly, participants with higher scores of depressions had lower rates of correct choices in the Action condition (orange; median split). **(B) Effect of depression on the difference in the rate of correct choices between the Action observation and Private conditions.** The rates of correct choices of the participants identified as sure cases of depression (‘severe’) were significantly lower in the Action than in the Private condition. **(C) Effect of**

depression on the integration of private information in the Action condition. Analysis of the trial-by-trial regression revealed that in the Action observation condition, the choice of the participants identified as sure cases of depression ('severe') was less predicted by the outcome obtained at the previous trial than those of the other participants, revealing that they integrated less the individually provided information in this condition. Error bars correspond to 95% confidence intervals.

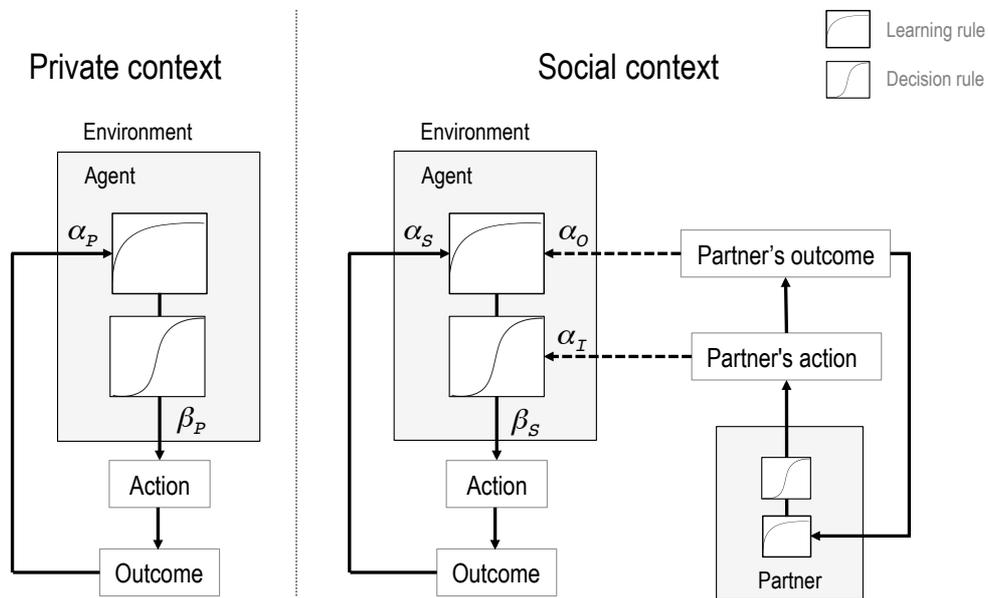


Figure 3. Social reinforcement learning model

A social reinforcement learning model was fitted on participants' behavior. In the Private information context, the model corresponded to a classical Q-learning model. On the contrary, in the Social information context, a supplemented Q-learning model was fitted such that social information was integrated to the decision process. Following Burke et al. (22), choice probability was updated based on the partner's action in the Action observation decision and the option value was updated when both the partner's action and outcome were presented.

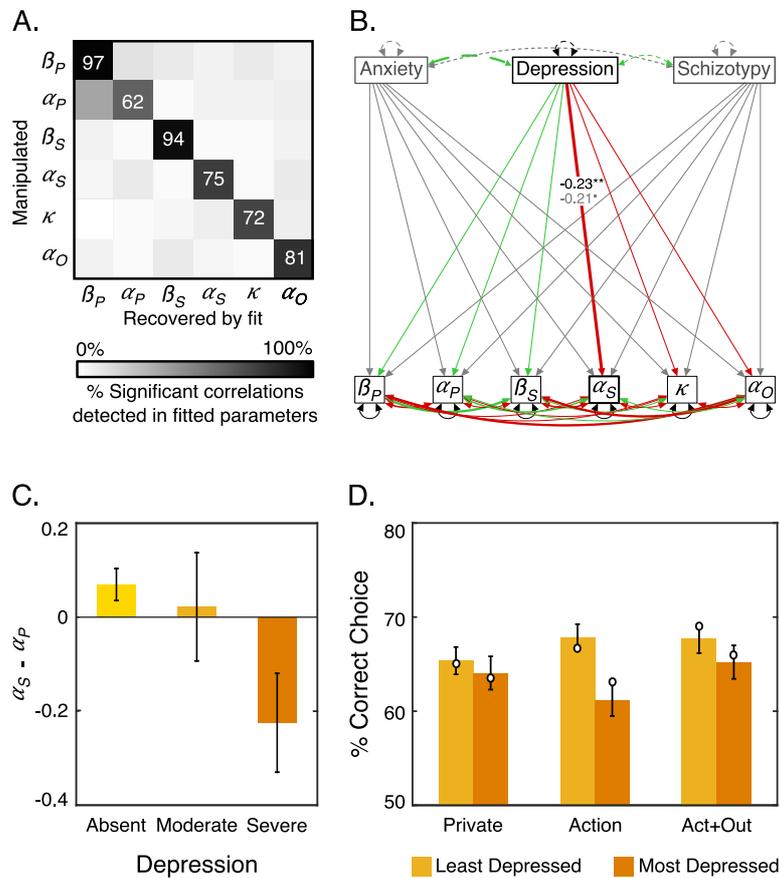


Figure 4. Computational effects of depression
(A) Parameter recovery. A sensitivity-specificity test on our tasks was conducted to assess whether this task was suited to accurately recover correlations between the model parameters and individual variables. The diagonal corresponds to the correlation accurately recovered, the other cases correspond to cross-contamination of the different parameters (number of correlation indicated from white for zero correlations to black for 100 correlations). The test indicated that our task was highly sensitive and specific for detecting correlations between model parameters and individual variables. **(B) Effect of depression on the model parameters.** Depression was specifically associated with a decrease in the individual learning rate in the Social information conditions (results obtained by structural equation modeling). **(C) Effect of depression on the difference in the individual learning rates between the Private and Social**

information conditions. Participants identified as sure cases of depression presented a decrease in their individual learning rates between the Private and Social information conditions such that they integrated less the individually generated information when presented with socially-provided information. **(D) Comparison of participants' behavior and model simulations of the rate of correct choices.** The simulation (points) accurately recovered the decrease in performance in the Action condition associated with higher scores of depression (bars correspond to the data). Error bars correspond to 95% confidence intervals.

Table 1. Descriptive statistics for age, gender, depression, anxiety and schizotypy scores

	Age	Sex ratio (% women)	Depression score	Anxiety score	Schizotypy score
First sample (N = 50)	33.02 ± 1.25 [22 – 62]	28%	5.46 ± 0.63 [0 – 19]	6.40 ± 0.58 [0 – 15]	2.50 ± 0.39 [0 – 13]
Second sample (N = 50)	33.76 ± 1.63 [19 – 61]	42%	4.96 ± 0.63 [0 – 16]	6.30 ± 0.64 [0 – 20]	2.14 ± 0.33 [0 – 8]

For each sample, the mean of each demographic variable is presented with its 95% confidence interval

References

1. A. van Randenborgh, R. de Jong-Meyer, J. Hüffmeier, Decision making in depression: differences in decisional conflict between healthy and depressed individuals. *Clin. Psychol. Psychother.* **17**, 285–298 (2010).
2. M. P. Paulus, A. J. Yu, Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* **16**, 476–483 (2012).
3. N. Eshel, J. P. Roiser, Reward and Punishment Processing in Depression. *Biol. Psychiatry.* **68**, 118–124 (2010).
4. J. C. Coyne, Toward an Interactional Description of Depression. *Psychiatry.* **39**, 28–40 (1976).
5. C. Hofheinz, M. Germar, T. Schultze, J. Michalak, A. Mojzisch, Are Depressed People More or Less Susceptible to Informational Social Influence? *Cogn. Ther. Res.*, 1–13 (2017).
6. M. R. DiMatteo, H. S. Lepper, T. W. Croghan, Depression Is a Risk Factor for Noncompliance With Medical Treatment: Meta-analysis of the Effects of Anxiety and Depression on Patient Adherence. *Arch. Intern. Med.* **160**, 2101–2107 (2000).
7. R. C. Ziegelstein et al., Patients With Depression Are Less Likely to Follow Recommendations to Reduce Cardiac Risk During Recovery From a Myocardial Infarction. *Arch. Intern. Med.* **160**, 1818–1823 (2000).

8. N. K. Arora, C. A. McHorney, Patient preferences for medical decision making: who really wants to participate? *Med. Care.* **38**, 335–341 (2000).
9. C. C. Ruff, E. Fehr, The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15**, 549–562 (2014).
10. K. N. Laland, Social learning strategies. *Learn. Behav.* **32**, 4–14 (2004).
11. A. Bandura, *Social learning theory* (Prentice-Hall, Oxford, England, 1977), vol. viii.
12. J. Li, M. R. Delgado, E. A. Phelps, How instructed knowledge modulates the neural systems of reward learning. *Proc. Natl. Acad. Sci.* **108**, 55–60 (2011).
13. M. M. Walsh, J. R. Anderson, Modulation of the feedback-related negativity by instruction and experience. *Proc. Natl. Acad. Sci.* **108**, 19048–19053 (2011).
14. E. R. Stone, C. L. Dodrill, N. Johnson, Depressive Cognition: A Test of Depressive Realism Versus Negativity Using General Knowledge Questions. *J. Psychol.* **135**, 583–602 (2001).
15. T. Szu-Ting Fu, W. Koutstaal, L. Poon, A. J. Cleare, Confidence judgment in depression and dysphoria: The depressive realism vs. negativity hypotheses. *J. Behav. Ther. Exp. Psychiatry.* **43**, 699–704 (2012).

16. D. J. Martin, L. Y. Abramson, L. B. Alloy, Illusion of control for self and others in depressed and nondepressed college students. *J. Pers. Soc. Psychol.* **46**, 125–136 (1984).
17. S. Kajimura, M. Nomura, When we cannot speak: Eye contact disrupts resources available to cognitive control processes during verb generation. *Cognition.* **157**, 352–357 (2016).
18. J. J. Wang, I. A. Apperly, Just one look: Direct gaze briefly disrupts visual working memory. *Psychon. Bull. Rev.* **24**, 393–399 (2017).
19. L. Conty, D. Gimmig, C. Belletier, N. George, P. Huguet, The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition.* **115**, 133–139 (2010).
20. L. Conty, N. George, J. K. Hietanen, Watching Eyes effects: When others meet the self. *Conscious. Cogn.* **45**, 184–197 (2016).
21. P. L. Rock, J. P. Roiser, W. J. Riedel, A. D. Blackwell, Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol. Med.* **44**, 2029–2040 (2014).
22. C. J. Burke, P. N. Tobler, M. Baddeley, W. Schultz, Neural mechanisms of observational learning. *Proc. Natl. Acad. Sci.* **107**, 14431–14436 (2010).
23. D. Medin, B. Ojalehto, A. Marin, M. Bang, Systems of (non-)diversity. *Nat. Hum. Behav.* **1**, 0088 (2017).
24. C. M. Gillan, M. Kosinski, R. Whelan, E. A. Phelps, N. D. Daw, Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife.* **5**, e11305 (2016).

25. C. M. Gillan, N. D. Daw, Taking Psychiatry Research Online. *Neuron*. **91**, 19–23 (2016).
26. D. N. Shapiro, J. Chandler, P. A. Mueller, Using Mechanical Turk to Study Clinical Populations. *Clin. Psychol. Sci.* **1**, 213–220 (2013).
27. A. S. Zigmond, R. P. Snaith, The Hospital Anxiety and Depression Scale. *Acta Psychiatr. Scand.* **67**, 361–370 (1983).
28. J. T. W. Wigman et al., Evidence That Psychotic Symptoms Are Prevalent in Disorders of Anxiety and Depression, Impacting on Illness Onset, Risk, and Severity—Implications for Diagnosis and Ultra–High Risk Research. *Schizophr. Bull.* **38**, 247–257 (2012).
29. D. A. Regier, D. S. Rae, W. E. Narrow, C. T. Kaelber, A. F. Schatzberg, Prevalence of anxiety disorders and their comorbidity with mood and addictive disorders. *Br. J. Psychiatry.* **173**, 24–28 (1998).
30. M. Puskaric, B. von Helversen, J. Rieskamp, How social and non-social information influence classification decisions: A computational modelling approach. *Q. J. Exp. Psychol.* **70**, 1516–1534 (2017).
31. S. Palminteri, V. Wyart, E. Koechlin, The Importance of Falsification in Computational Cognitive Modeling. *Trends Cogn. Sci.* **21**, 425–433 (2017).
32. L. B. Alloy, J. S. Albright, C. M. Clements, in *Social Processes in Clinical and Counseling Psychology*, J. E. Maddux, C. D. Stoltenberg, R. Rosenwein, Eds. (Springer New York, 1987;

http://link.springer.com/chapter/10.1007/978-1-4613-8728-2_8), pp. 94–112.

33. C. Chen, T. Takahashi, S. Nakagawa, T. Inoue, I. Kusumi, Reinforcement learning in depression: A review of computational research. *Neurosci. Biobehav. Rev.* **55**, 247–267 (2015).
34. D. Chung *et al.*, Valuation in major depression is intact and stable in a non-learning environment. *Sci. Rep.* **7** (2017), doi:10.1038/srep44374.
35. Q. J. Huys, D. A. Pizzagalli, R. Bogdan, P. Dayan, Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* **3**, 12 (2013).
36. M. M. Herzallah *et al.*, Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Front. Integr. Neurosci.* **7** (2013), doi:10.3389/fnint.2013.00067.
37. R. L. Kendal, I. Coolen, Y. van Bergen, K. N. Laland, B.-A. in the S. of Behavior, Ed. (Academic Press, 2005; <http://www.sciencedirect.com/science/article/pii/S006534540535008X>), vol. 35, pp. 333–379.
38. L. A. Wood, R. L. Kendal, E. G. Flynn, Whom do children copy? Model-based biases in social learning. *Dev. Rev.* **33**, 341–356 (2013).
39. M. J. Weightman, T. M. Air, B. T. Baune, A Review of the Role of Social Cognition in Major Depressive Disorder. *Front. Psychiatry.* **5** (2014), doi:10.3389/fpsy.2014.00179.

40. A. Todorov, R. Dotsch, J. M. Porter, N. N. Oosterhof, V. B. Falvello, Validation of data-driven computational models of social perception of faces. *Emot. Wash. DC.* **13**, 724–738 (2013).
41. E. R. Peters, S. A. Joseph, P. A. Garety, Measurement of Delusional Ideation in the Normal Population: Introducing the PDI (Peters et al. Delusions Inventory). *Schizophr. Bull.* **25**, 553–576 (1999).

Supplementary Materials

Impact of face evaluations on social learning

A effect impact of trustworthiness evaluations was found, such that, compared the Private condition, participants who perceived the partner's avatar as more trustworthy had higher correct choice rates in Action ($b_{Action*Trust} = 0.19 \pm 0.20, z = 1.71, p = .087$) and in the Action+Outcome conditions ($b_{Action*Trust} = 0.27 \pm 0.22, z = 2.07, p = .038$; Figure S1). This effect of the social evaluation of the partner's avatar suggests that participants actually processed the information in a social context.

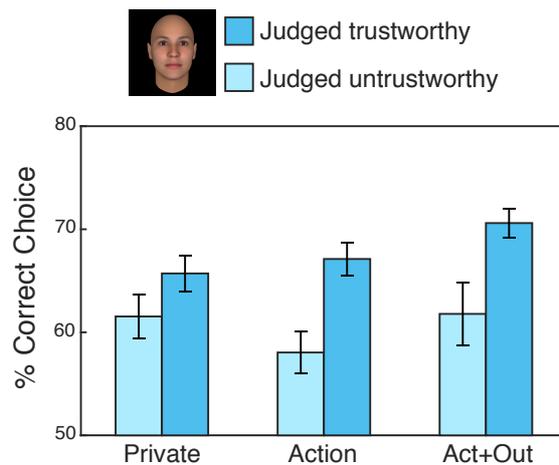


Figure S1. Effect of avatar's perceived trustworthiness on social learning

Participants who rated the partner's avatar as trustworthy had higher rates of correct choices in the Action and Action+Outcome observation conditions than those who rated the avatar as untrustworthy.

– Registered report –

Increased susceptibility to social influence depends on greater environmental harshness and unpredictability experienced during childhood

Pierre O. Jacquet^{1,2}, Lou Safra¹, Nicolas Baumard², Coralie Chevallier¹

¹ Laboratoire de Neurosciences Cognitives (LNC), Département d'Etudes Cognitives, INSERM U960, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France.

² Institut Jean Nicod, Département d'Etudes Cognitives, CNRS UMR8129, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France.

Corresponding Author:

Pierre O. Jacquet,

Laboratoire de Neurosciences Cognitives, INSERM U960, Département d'Etudes

Cognitives, Ecole Normale Supérieure

29, rue d'Ulm

75005, Paris, France

Phone : +33(0)1.44.32.26.40

E-Mail address: pierre.jacquet@inserm.fr

Abstract

Recent animal models suggest that susceptibility to social influence is biased by how much the environment penalizes exploratory behaviours and novelty seeking. Environments bearing significant fitness costs should logically increase individuals' susceptibility to social influence at the expense of exploration and novelty seeking. The possibility that similar variations might occur in humans has however received little attention. By combining classical psychophysics and computational modelling, we aim to test the independent and joint contribution of childhood environmental harshness and unpredictability on humans' susceptibility to social influence. Preliminary results collected on-line (N=122) in a social preference task showed that susceptibility to social influence and childhood environment were positively linked. This relation was relatively dimension-specific: greater harshness resulted in a marginal increase in social influence later in adulthood, while a greater unpredictability significantly tuned people's preference towards social information. Crucially, those participants who had faced more adverse environments – i.e., a high level of harshness coupled with a high level of unpredictability – were more prone to rely on social information. Given the important implications that our findings might have for current views on the mechanisms mediating the social transmission of cultural values, we plan to test the robustness of this effect by running a replication of this experiment in a bigger sample (~ 220 healthy adults recruited on-line).

Introduction

The beliefs of our conspecifics about unfamiliar people represent an immediately available information that often influences our own social preferences (e.g., who to trust, mate with, etc.)¹⁻⁶. An intriguing fact which has received little – if any – attention is that we are not all equally likely to match our social beliefs to those of our conspecifics. At the population scale, individual variations in social influence may lead to variations in the emergence and the maintenance of behavioural standards that may in certain situations bear non-negligible costs for human societies (i.e., low cooperation, low trust, discrimination, prejudices and stereotypes, etc.)⁷. What are the factors driving this inter-individual variability? Do these variations reflect random noise or do they have a functional (adaptive) explanation? The present study aims to investigate these important gaps in knowledge by combining classical psychophysics and computational modeling.

Social information (behaviours, beliefs and attitudes that can be observed from peers) is often thought to be strategically exploited by observers depending on a number of ‘external factors’, such as the content of the information itself or the characteristic of the agents from which it originates (e.g., reputation)⁸. From an evolutionary viewpoint, the question of whether to use social information or not amounts to identifying the optimal trade-off between the cost and the benefit of individual exploration by trial-and-error (e.g., searching for options that provide the best, albeit more delayed and uncertain rewards) and simply mimicking behaviours

observed in conspecifics (e.g. copying available, albeit potentially maladaptive options with the aim of getting a lower but more immediate reward)⁸. These trade-offs should vary across individuals as a function of environmental parameters that are known to increase or decrease the fitness costs of exploration and social information use⁹. Although this theoretical model originally applies to foraging decisions in non-human animals, it could be particularly useful to account for human decisions in the social domain. For example, judging an unfamiliar person as trustworthy at first sight increases the motivation to have an interaction with that person (i.e., a typical exploratory behaviour) but simultaneously increases the risk of being cheated or exploited. In an environment where the probability of being harmed, cheated or exploited is high, it might thus be more advantageous on average to follow others' advice regarding the approachability of unfamiliar people rather than to gather information alone through direct social interaction. In sum, whether our social world is cast in an adverse ecology or not shifts the cost-benefit trade-off of individual exploration and social information use, just as in the foraging domain.

In humans, early exposure to environmental stressors alters individuals' resource allocation decisions of toward an acceleration of life strategies which, at the psychological level, triggers a constellation of traits characterized by a sensitivity for immediate instead of long-term benefits. This 'present-oriented' psychology is pervasive and affects behaviour in various domains such as investment in health, parenting, reproduction, economic decision-making or

cooperation¹⁰⁻¹⁴. Our starting assumption is that a greater susceptibility to social influence can also be viewed as a ‘present-oriented’ psychological trait. Indeed, following other people preferences is a way to immediately benefit from payoffs resulting from well-learned, minimally optimal behaviours. However relying on social information too much also has an opportunity cost: the observer can miss more optimal, albeit more delayed and uncertain opportunities that might have occurred had she relied on a more exploratory strategy. We thus aim to test the hypothesis that variation in the susceptibility to social influence depends on the level of environmental adversity that individuals experienced during their life.

Converging evidence suggest that environmental stress experienced early in life is mainly carried by two fundamental dimensions – harshness and unpredictability – which determine the adaptive calibration of an individual’s psychology towards a ‘present-oriented’ or a ‘future-oriented’ strategy. Harshness defines the rates at which external factors – morbidity-mortality – cause disability and death at each age in a population, while unpredictability defines the rates at which harshness varies over time and space¹⁵. An important cue for experienced harshness during childhood is socio-economic status (SES). Indeed, lower levels of SES have been shown to correlate with virtually all forms of morbidity and mortality^{16,17}, indicating that people with low SES levels experienced a greater exposure to disease, disability and death. Childhood unpredictability has been much less studied in humans, but several studies demonstrated that it could be driven by

frequent changes or ongoing inconsistency in several dimensions of childhood environments, i.e., family instability and household inconsistency, frequent residential change, etc.¹⁸⁻²³.

Overall, high levels of harshness and unpredictability are thought to uniformly trigger present-oriented psychological responses. There are some circumstances however in which the two dimensions might differently act on behaviour. A notable difference concerns parental investment. If both dimensions undermine the quantity and quality of efforts individuals put in rearing their offspring, empirical data suggest that unpredictability has a significantly greater impact on parental functioning and subsequent child outcomes. This has important consequences as parental investment is the main source of information that young children exploit to infer risks and opportunities in their current environment and adaptively calibrate their phenotype to its predicted future states. The very reason lies in the fact that harshness imposes a tractable and consistent cost on the individuals' fitness, while unpredictability increases its temporal variance. By definition, in stochastic conditions there are no available cues that allow the individual to consistently predict what the environment will look like in the future. Parental investment is thus of little help in monitoring the cost imposed by unpredictable conditions on offspring survival. The most adaptive strategy is therefore to focus behaviours on the 'here and now'. Instead, the effects of harshness can be counter-acted, at least partly, by an increase in parental investment, i.e., a

strategy that is more likely to ensure children survival through shielding them from predictable sources of morbidity and mortality threats.

This line of observations allows us to draw specific predictions about the specific effect that harshness and unpredictability (as well as their combination) might have on the individuals' susceptibility to social influence. First, variations in experienced harshness and unpredictability should explain a significant part of the variance in social influence independently: Individuals exposed to harsh or unpredictable environments early in life should be more likely to be influenced by peers to guide their behaviours. Second, unpredictability should be a better predictor of social influence than harshness, because the effect of the latter on the development of a 'present-oriented' psychology can be partially counter-acted by other external forces like the quantity and quality of parental investment. Third, the combination of a high level of harshness with a high level of unpredictability is expected to have a greater effect on social influence than the two dimensions taken in isolation. Beyond a certain threshold, we indeed predict that the deleterious effect of harshness on individuals' fitness cannot be counteracted by any alternative phenotypes. For example, it may prove more advantageous on average to follow one's conspecifics' advice regarding the approachability of unfamiliar persons rather than to gather this information alone through direct social interaction if the probability of being harmed, cheated or exploited is consistently high (harshness), or varies in time from high to low in a stochastic manner (unpredictability).

To test these hypotheses we adapted a well-validated social preference task made in a peer pressure context⁵ (Figure 1), asking participants to rate unfamiliar faces on the approachability dimension before and after being confronted to the rating of a simulated group of conspecifics. Social information consists in displaying the most frequent rating provided by a group of peers. This rating can deviate positively or negatively from the participants' rating to a moderate or high extent (disagreement conditions), or not deviate at all (agreement condition). A computational model of choice is used to analyze the weight attributed by participants to social information during post-test ratings. Behavioural inter-individual variability is analyzed in light of perceived childhood environmental adversity and of its two dimension, i.e., harshness and unpredictability.

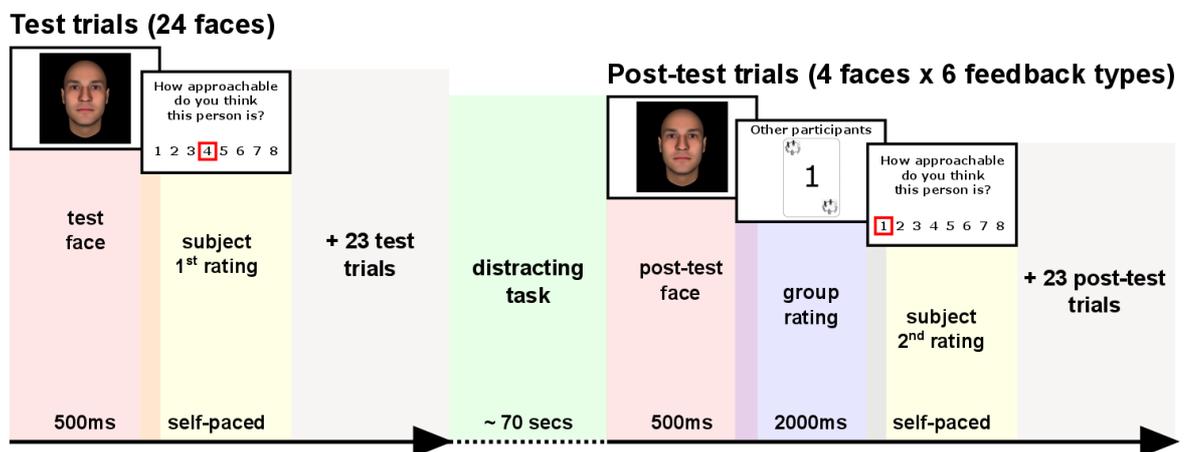


Figure 1. Experimental procedure. In a first phase (test trials) participants rate a series of twenty-four computerized faces (randomized by participants) on the

approachability dimension by doing a mouse click on one of the 8 possible values of the scale. After the completion of the first phase, participants perform a distracting task lasting approximately 70 secs. Differences between the distracting task in the pilot and in the target study are detailed in the Material and Methods section. Participants then complete the third and last phase of the experiment (post-test trials), which consists in presenting the twenty-four faces seen in the first phase for a second time (the order of the post-test series is randomized by participants). After each face presentation, participants are exposed to the rating provided by the (simulated) group of peers for a 2000ms duration. Right after this feedback period they are asked to rate the face for a second time. In the present example, the comparison of the participant's first and second rating indicates that she was influenced by the information displayed by the group: the rating was adjusted from 4 to 1, that is, to the most frequent rating produced by other participants.

Results of the pilot study

Influence of social information on participants' approachability ratings

We started our analyses by testing whether participants, independently of early-life environmental stress, exhibited a general bias towards social information during the rating of faces approachability. We thus compared the social influence index (SII) averaged across the moderate (the group rating could differ of +2 or -2 points on the

scale from the participants rating) and high magnitude (the group rating could differ of +3 or -3 points on the scale from the participants rating) sub-dimensions of social disagreement 0 (one-tailed t-test), a value indicating that in post-test trials participants stuck on average to their initial choices and did not follow the group. A positive SII therefore indicated that participants adjusted their ratings in the direction of the group rating, and a negative SII indicated that they adjusted their ratings away from the group. A t-test revealed that the SII obtained in social disagreement trials ($.69 \pm .54$) differed from 0 in a positive way ($t > 14.02$, $p < .0001$), which showed that the participants' rating adjustment was influenced by social feedback (Figure 2.A). In order to test whether the bias towards social information was observed independently of the magnitude of the disagreement between the participants' ratings and the group ratings, the same comparison of SII with 0 was done with trials split into moderate and high disagreements, respectively. Results showed that exposure to high disagreements elicited a greater SII than exposure to moderate disagreements ($.88 \pm .70$ vs $.50 \pm .56$, $t > 6.39$, $p < .0001$, one-tail), in spite of the fact that both SII obtained in both conditions differed from zero (all $t_s > 9.91$, all $p_s < .0001$) (Figure 1.A).

We further checked whether the bias towards social information varied as a function of the valence dimension of the disagreement (Figure 2.B). Here, a mean rating adjustment (MRA) for both positive (the group rated the face as more approachable than the participants) and negative (the group rated the face as less

approachable) disagreement trials was calculated. At this stage, a positive MRA indicated that participants adjusted their ratings positively in post-test trials, i.e., they rated the faces as more trustworthy than they initially did), and a negative MRA indicated that they adjusted their ratings negatively in post-test trials, i.e., they rated the faces as less trustworthy than they initially did. A near-zero MRA indicates that, on average, participants stuck to their initial ratings in post-test trials. Finally, we calculate an MRA for trials displaying an agreement with the group. We thus expected positive and negative MRAs to be observed in the positive and negative sub-dimensions of the social disagreement, respectively. A near-zero MRA should be obtained in the agreement condition. In line with these predictions, we found a negative MRA ($-.80 \pm .70$) differing from zero for the negative disagreement sub-dimension ($t < 12.59, p < .0001$), and a positive MRA ($.58 \pm .75$) differing from zero for the positive disagreement sub-dimension ($t > 8.51, p < .0001$). In the agreement condition, we obtained a value ($-.17 \pm .60$) that was closer to zero when compared to the MRA obtained in the negative (two-tailed t-test, $t > 7.47, p < .0001$) and positive (two-tailed t-test, $t > 8.47, p < .0001$) disagreement sub-dimensions. However, it significantly differed from 0 ($t < 3.15, p = .002$), indicating a *by-default* decrease of approachability ratings in post-test trials.

To characterize the mechanism leading to social information use in the present task, we fitted participants' behaviour using a simple model of choice which hypothesizes that the decision of adjusting a rating after the integration of social

information is formed on the basis of a comparison between the faces presented in test trials and the social information (group rating) presented in post-test trials (see Materials and Methods). The model consists of two free parameters, fitted to each participant's behaviour: 1) a social information use index δ corresponding to the adjustment of an initial rating in post-test trials, measured as the signed fraction of disagreement between the initial rating and the subsequent group rating, and 2) an internal noise magnitude σ corresponding to the standard deviation of the post-test rating. We obtain maximum-likelihood estimates of these two parameters from each participant's behaviour (see Materials and Methods) and then compare the predictions it makes with participants' performance in terms of social influence index (SII) and mean rating adjustment (MRA), depending of the disagreement dimension under consideration (i.e., magnitude or valence). We then tested the adequacy of the model by checking whether the effects that revealed by the analyses of the participants' performance were also obtained with the model's predictions (Figure 2).

First, the model-predicted SII for the general social disagreement condition ($.70 \pm .54$) differed from 0 in a positive way ($t > 14.41$, $p < .0001$), so were the model-predicted SII for the high ($t > 14.39$, $p < .0001$) and moderate ($t > 14.37$, $p < .0001$) disagreement sub-dimensions. Note that the modelled SII for high ($.84 \pm .65$) and moderate ($.56 \pm .43$) disagreement trials differed between each other ($t > 13.79$,

$p < .0001$) in the same way as did their observed behavioural counterparts (Figure 2.A).

Furthermore, the model predicted a negatively signed MRA ($-.71 \pm .54$) for the negative disagreement sub-dimension ($t < 14.43$, $p < .0001$), a positively signed MRA ($-.69 \pm .54$) for the positive disagreement sub-dimension ($t > 14.14$, $p < .0001$), and a near-0 MRA ($-.02 \pm .08$) for the agreement condition ($t < 3.01$, $p = .003$). As with the real participants performances, model-predicted MRA obtained in the negative ($t < 14.06$, $p < .0001$, one-tail) and positive (t-test, $t > 14.35$, $p < .0001$, one-tail) disagreement sub-dimensions both differed from the predicted MRA obtained in the agreement condition (Figure 2.B).

The adequacy of the model was further evidenced by the amount of inter-individual variance of the observed SII and MRA that was captured by their model-predicted equivalent, 1) SII: general disagreement: $R^2 = .99$, $p < .0001$; moderate disagreement: $R^2 = .59$, $p < .0001$; high disagreement: $R^2 = .87$, $p < .0001$; 2) MRA: negative disagreement: $R^2 = .54$, $p < .0001$; positive disagreement: $R^2 = .57$, $p < .0001$; agreement: $R^2 = .07$, $p < .01$.

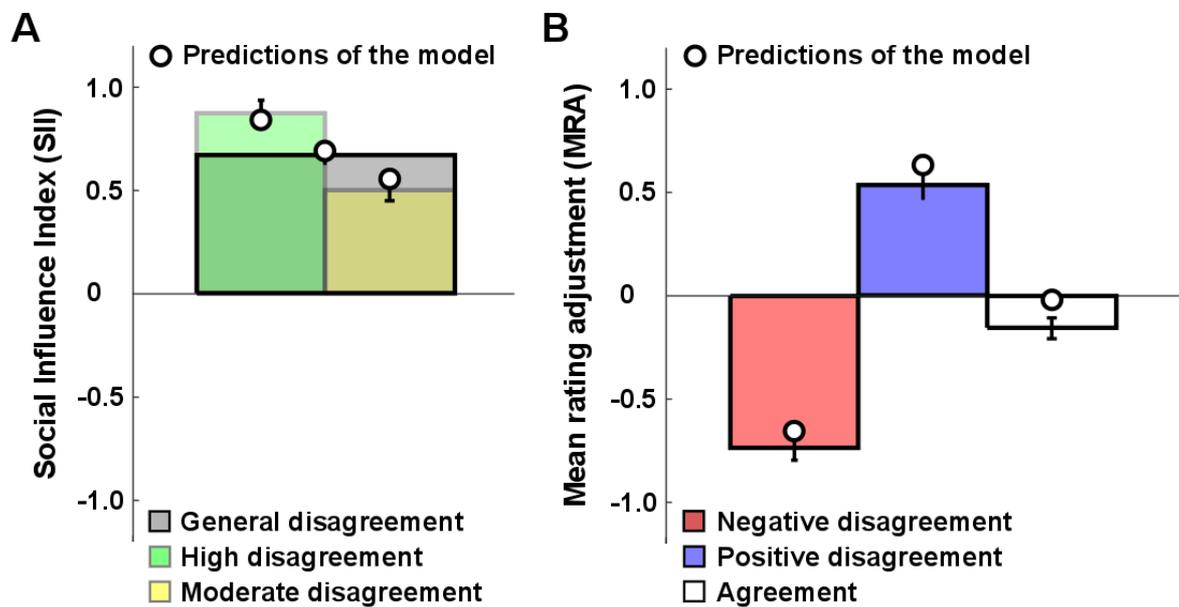


Figure 2. Influence of social information on observed and modelled social influence index (SII) and mean rating adjustment (MRA). A. Effect of social disagreement and of its magnitude. The SII (observed or modelled) represents how much the participants matched their ratings to the group (y axis). A positive SII therefore indicates that participants were influenced by social information when adjusting their ratings, a negative SII indicates that they were not influenced by social information, and a near-0 SII means that, in post-test trials, participants stuck on average on their initial ratings. The large grey column of the x axis represented the observed SII (\pm SEM) obtained for the general disagreement condition, i.e., the SII averaged on the high (green column) and moderate (yellow column) disagreement trials. The discs are the predictions made by the computational model for each type of social information (see Material and Methods for details). **B. Effect**

of the social disagreement valence. A positive MRA (y axis) indicates that participants rated the faces in post-test as more approachable than they initially did in test-trials (a negative MRA indicates that they rated the faces as less approachable). A near-0 MRA means that, in post-test trials, participants stuck on average on their initial ratings. The red, blue, and grey columns represent the observed MRA (\pm SEM) calculated for the negative, positive, and agreement sub-dimensions, respectively (x axis). The discs are the predictions made by the computational model for each social information type.

Susceptibility to social influence as a function of perceived childhood adversity and its dimensions of harshness and unpredictability

We ran independent linear mixed-effects models²⁴ fitted with the software Matlab version R2014b to investigate the contribution of childhood adversity and of its two dimensions – harshness and unpredictability – on the participants' social influence (SII) index and mean rating adjustment (MRA). For either SES, unpredictability, or adversity, two sets of models were conducted. The first set aimed to fit SII and therefore involved the magnitude dimension of social information (moderate disagreement vs high disagreement with the group) as within-subject factor. The second set aimed to fit MRA and involved the valence dimension of social information (positive disagreement vs negative disagreement vs agreement) as within-subject factor.

In the first set of models, we started with a ‘reduced’ model which included *Magnitude* as the within-subject fixed-effect predictor, and *Subject* as random effect factor. We then ran an ‘intermediate’ model in which the ecological index of interest (*SES*, *Unpredictability*, or *Adversity*) was added as an independent fixed-effect predictor, and a ‘full’ model in which the interaction term of the two fixed-effect predictors was added. The fixed-effect factors *SES*, *Unpredictability*, and *Adversity* were then tested independently by investigating whether the ‘intermediate’ and the ‘full’ models improved or penalized the fit of SII data relative to the ‘reduced’ model. To determine which of the SES, the unpredictability, or the adversity index most affected social influence, we compared the goodness of fit of the models that best predicted the SII.

The comparison procedure applied in the second set of models was roughly similar. Three differences should be noticed however. First, the within-subject fixed-effect factor *Magnitude* was replaced by *Valence* (i.e., positive disagreement, negative disagreement, and agreement). Second, models worked on the basis of mean rating adjustment (MRA) instead of social influence index (SII). The consequence is that we only compared a full model (which included an interaction term between the fixed-effect factors) to a reduced model (which only included *Valence* as fixed-effect predictor). The very reason why we did not design an intermediate model is that, taking MRA into consideration, a putative effect of any ecological index on the participants’ susceptibility to social influence can only be

indirectly inferred from its interaction with the *Valence* dimension of social information, i.e., a positive MRA for positive disagreement trials and a negative MRA for negative disagreement trials. Indeed, positive and negative MRAs do not indicate that participants conformed to or deviated from the group rating, but, instead, that they increased or decreased their approachability ratings along the task. Thus, a main effect of one of these indices would show that, on average and independently of the valence of the disagreement, participants increased or decreased their approachability ratings. In the full model, Interaction between the two fixed-effect factors – *Valence* and *SES*, *Unpredictability* or *Adversity* – was investigated by comparing the slope obtained in the agreement trials (in which each individual MRA represent the subjective baseline of rating changes) with slopes obtained in the positive and negative disagreement trials. We also compared the goodness of fit of the models that best predict the data in order to determine which of the SES, the unpredictability, or the adversity index most affected the MRA.

Models of the first and second sets all had random intercepts and random slopes for the within-subject factors. All model comparisons were performed using simulated likelihood ratio tests (1000 simulations per test). Simulated likelihood ratio test consists in generating the reference distribution of the likelihood ratio test (LR) statistic under the null hypothesis (e.g., the reduced model) from random data, and then assessing the statistical significance of the alternate model (e.g., the intermediate or the full model) by comparing the observed LR statistic to the

simulated reference distribution. The reported p -value is an exact value, which corresponds to the ratio of the number of times the simulated LR statistic is equal to or exceeds the observed value plus one, to the number of simulations plus one ($n = 1000$).

In the first set of models where *Magnitude* (moderate / high disagreement with the group of peers) was the within-subject fixed-effect factor we found that, relative to the reduced model, the best fitting model was the intermediate one in which the ecological index was added as an independent rather than an interacting fixed-effect predictor. The improvement of data fitting was marginal for the model including *SES* ($D(1) = 3.77$, $p = .067$), turned significant for the model including *Unpredictability* ($D(1) = 6.32$, $p = .014$), and reached its maximum with the model including *Adversity* ($D(1) = 7.36$, $p = .01$). More specifically, participants who grew up in an environment characterized by a low SES were more susceptible to social influence than participants who grew up in a high SES environment ($\beta = 0.09$, $t = 1.96$, $p = .05$) (Figure 3.A). The same pattern was observed with unpredictability: participants who grew up in an unpredictable environment being more influenced by social information than participants who grew up in a predictable environment ($\beta = 0.12$, $t = 2.55$, $p = .01$) (Figure 3.B). Finally, those of the participants who experienced a greater adversity – that is, a low SES paired with an unpredictable environment – were more susceptible to social influence than those for whom adversity was lower ($\beta = 0.08$, $t = 2.77$, $p = .006$) (Figure 3.C). Note that adding

Gender (female vs male) as a third fixed-effect predictor did not significantly improve the goodness of fit of intermediate models (D(1) range: 3.85 – 2.34; *p* range: .06 – .13).

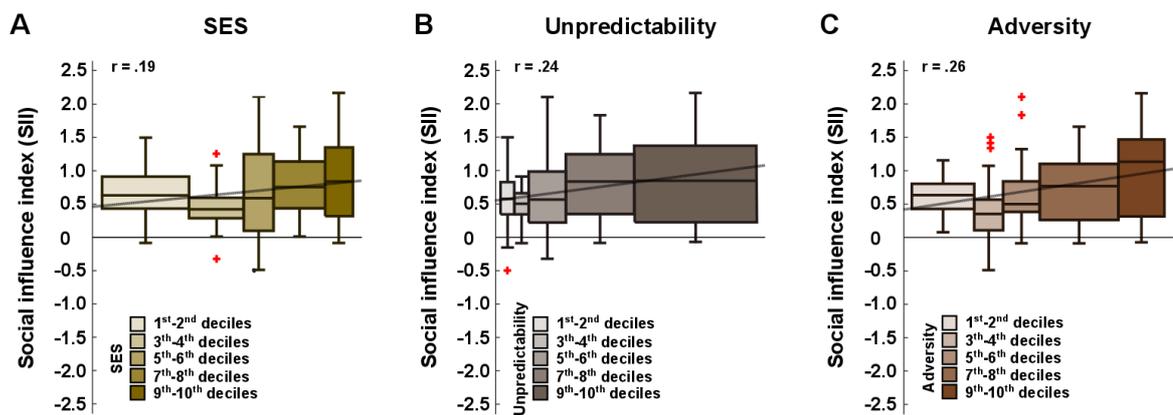


Figure 3. Social influence index (SII) as a function of A. childhood harshness, B. unpredictability, and C. adversity. The boxes represent the deciles of the distribution of each ecological index of interest, and their width are sized to their range. The horizontal lines inside the boxes are the median SII, and whiskers represent the SII's 25th and 75th percentiles. The red crosses depict outliers. The slopes are the linear correlations of the two variables, based on individual data points.

In the second set of models, we found that the full model better fit MRAs than the reduced one. This improvement was marginally significant for the full model which included *SES* ($D(3) = 7.19$, $p = .07$), and turned significant for the models including *Unpredictability* ($D(3) = 9.33$, $p = .031$) and *Adversity* ($D(3) =$

9.39, $p = .029$). Interestingly, the effect each ecological indices had on MRA depended on whether the disagreement with the group was of a positive or a negative valence. When compared to the agreement condition indeed, participants who experienced greater levels of harshness increased their approachability ratings more than others after they have been exposed to a positive disagreement (SES: $\beta = 0.21$, $t = 2.59$, $p = .01$). This was not the case when exposed to a negative disagreement (SES: $\beta = -0.005$, $t = -0.07$, $p = .95$). In that situation, there was no difference between the two slopes. When *Unpredictability* was considered, a reversed pattern was found. When compared to the agreement condition, participants who experienced less predictable environments decreased their approachability ratings more than others after they have been exposed to a negative disagreement with the group ($\beta = -0.16$, $t = -2.11$, $p = .036$). When exposed to positive disagreement, no difference was found between the slopes ($\beta = 0.11$, $t = 1.38$, $p = .017$). Finally, the greater sensitivity to positive disagreements was also observed with participants who grew up in more adverse environments, i.e. environments characterized by higher levels of both harshness and unpredictability (Adversity: $\beta = 0.12$, $t = 2.41$, $p = .016$); negative disagreement: $\beta = -0.06$, $t = -1.31$, $p = .19$) (Figure 3).

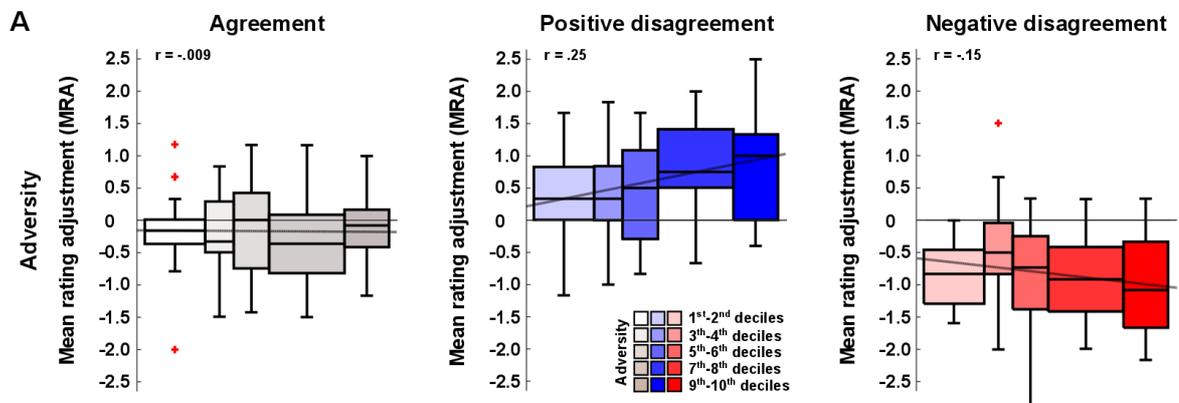


Figure 3. Mean rating adjustment (MRA) as a function of adversity. A. Agreement with the group of peers. **B.** Positive disagreement with the group of peers. **C.** Negative disagreement with the group of peers. The boxes represent the deciles of the distribution of each ecological index of interest, and their width are sized to their range. The horizontal lines inside the boxes are the median MRA, and whiskers represent the MRA's 25th (lower limit) and 75th (upper limit) percentiles. The red crosses depict outliers. The slopes are the linear correlations of the two variables, based on individual data points.

To better understand these last results we ran a set of complementary models in which each ecological indices was regressed against MRA obtained in the positive disagreement sub-dimension, and another set in which it was regressed against MRA obtained in the negative disagreement sub-dimension. All of them included *Subject* as random effect factor, had random intercepts and random slopes. Results showed that MRA obtained in the positive disagreement sub-dimension was

positively linked to SES ($\beta = 0.13$, $t = 1.89$, $p = .06$) and, in a more significant extent, to Unpredictability ($\beta = 0.18$, $t = 2.69$, $p = .008$) and Adversity ($\beta = 0.11$, $t = 2.79$, $p = .006$). By comparison, no link was found between MRA obtained in the negative disagreement sub-dimension and any of the three ecological indices (SES: $\beta = -0.08$, $t = -1.32$, $p = .19$; Unpredictability: $\beta = -0.09$, $t = -1.40$, $p = .16$; Adversity: $\beta = -0.06$, $t = -1.66$, $p = .10$). Note however that the sign of the relation between the two variables conformed to our expectations (i.e., negative correlation for the negative disagreement sub-dimension). In sum, we cannot strictly conclude, on the basis of these results, whether harshness and unpredictability independently and differently impacts MRAs in the positive and negative disagreements, or whether the absence of relation between the two variables in one or the other condition is rather due to a lack of statistical power. We thus hope that the bigger sample that we plan to recruit for the target study will help us disentangle between these two possibilities.

In a final step, we aimed to determine which of the SES, the unpredictability, or the adversity index most affected SII on one hand, and MRA and the other hand. To do so we used simulated likelihood ratio tests to find which of the intermediate models of set #1 and which of the full models of set #2 had the best predictive power (number of simulations = 1000). Between the three intermediate models involving *Magnitude* as within-subject fixed-effect factor, we found that *Unpredictability* ($D(0) = 2.55$, $p = .034$) and *Adversity* ($D(0) = 3.59$, $p = .023$)

provided a significant advantage over *SES*. Crucially, we also found that *Adversity* tended to better fit the data than *Unpredictability* ($D(0) = 1.04$, $p = .12$). Very similar results were found when we compared intermediate models involving *Valence* as within-subject fixed-effect factor, with *Unpredictability* ($D(0) = 2.14$, $p = .06$) and *Adversity* ($D(0) = 2.19$, $p = .095$) both marginally improving goodness of fit relative to *SES*. Evidence for a greater explanatory power of *Adversity* over *Unpredictability* was more meager ($D(0) = 0.06$, $p = 0.21$). Note however that the last three results turned significant (*Unpredictability* vs *SES*: $D(0) = 3.71$, $p = .034$; *Adversity* vs *SES*: $D(0) = 5.02$, $p = .019$) or marginally significant (*Adversity* vs *Unpredictability*: $D(0) = 1.31$, $p = .09$) once the agreement condition was withdrawn from the *Valence* within-subject fixed-effect factor. This is not surprising and it can be easily explained by the fact that ecological indices were expected to predict MRA only in the positive and negative disagreement sub-dimensions (i.e., where the participants and the group ratings conflicted), while they were expected to let the MRA obtained in the agreement condition unaffected (i.e., where the participants and the group ratings strictly matched).

Materials and Methods

Pilot study

Participants

In order to assess sample size, we capitalized on an existing data set that was initially collected to test the effect of disgust on social influence. The pilot experiment compared the influence exerted by social information on face approachability ratings in a condition in which participants were primed with disgusting images and in another condition in which they were primed with neutral images²⁵. For our current purposes we restricted our analysis to the neutral condition (a condition fairly similar to the procedure that we aim to use in the target study) that was conducted on 122 participants (59 females, 63 males). Our goal here is to compare ratings of computerized human faces on the approachability dimension *before* (test rating trials) and *after* (post-test rating trials) participants have been confronted with the group's modal rating. All participants were at least 18 years old (mean age = 35, SD = 9), reported being naïve to the purpose of the experiment, gave their written informed consent and received payment for their participation in accordance with the standards of Mechanical Turk. The experimental protocol that was performed was approved by the local Ethical Committee (Conseil d'évaluation éthique pour les recherches en santé – CERES n°201659) and is in accordance with the Declaration of Helsinki (World Medical Association, 2008).

Stimuli and procedure

Stimuli consisted in 24 emotionally neutral faces generated using FaceGen Modeller 3.5 (Singular Inversions, 2007) according to the methods developed by Oosterhoof and Todorov²⁶. The experimental procedure was adapted from⁵, and was divided into three phases.

In a first phase, participants watched a series of 24 faces each presented for 500 msec, a duration that has been shown to minimize participants' ability to encode the identity of the faces but that provides enough time to consistently estimate social traits such like trustworthiness²⁶⁻²⁸. After the presentation of each face, a numerical scale appears to assess approachability. The scale remained on the screen until participants selected a value between 1 and 8 (Figure 1).

Immediately after the completion of the test phase, participants perform a distractive task where they were asked to watch neutral images²⁵ and judge how disgusting they were on an 8-point scale. Each image was shown for 3 sec. A numerical scale then appears on the screen until participants select a value comprised between 1 and 8, indicating how disgusting they judged the stimulus (1 = not disgusting at all, 8 = very disgusting). The disgust task lasted approximately 2 minutes.

In a third phase, we asked participants to rate the very same faces (presented in a random order) for a second time, but informed them that before watching each of them, they will see other M-Turker's modal rating. This social group rating was

represented in a numerical format and centered on a card featuring a social group symbol in its angles (Figure 1). A non-informative feedback was also introduced by means of a card featuring a question mark in its center. In these trials participants were expected to stick to their initial rating. This control condition enabled us to ensure that participants perform the task appropriately and do not answer randomly.

Unbeknownst to the participants, the group rating was fictive and generated on-line by means of a simple algorithm. Note that previous studies using a similar procedure have shown that participants were convinced that the group ratings were provided by real individuals^{5,6,29-31}. The fictive group rating could deviate up or down the participant's initial rating to a moderate (+2/-2 points deviation) or high (+3/-3 points deviation) extent, or could not deviate at all (0 point deviation). The delivery of these various feedbacks was thus conditioned on each participant's response during the test phase, and the feedback algorithm was built such that a trial could be assigned a social disagreement that deviates up or down the participants' rating only when this rating was equal to or greater than 2 (i.e., the minimum value of the scale is 1) or equal to or smaller than 7 (i.e., the maximum value of the scale is 8). Overall, there were two different types of trials: those displaying an agreement with the group (0 point deviation: same ratings as the participant) and those displaying a disagreement with the group (-3/-2/+2/+3 points deviation: different ratings). Disagreement trials could then be split in two sub-dimensions depending on the valence of the group deviation (positive vs. negative disagreement), or in two

sub-dimensions depending on the magnitude of the group deviation (moderate vs. high disagreement). An additional condition without social information was also introduced: instead of presenting the group rating, a question mark appeared up to the fixation square. These non-informative trials were used to ensure that participants did not performed the task in a random fashion but, instead, adjusted their responses according to the type of social information they were exposed to. A t-test against 0 indeed showed that, on average, the participants' initial ratings remained unchanged after the exposure to the non-informative feedbacks ($t = 0.18$, $p = .86$). On average, the feedback conditions were equally distributed, i.e., 6 trials \times 4 feedback conditions (non-informative feedback, agreement, 2 sub-dimensions of disagreements based on either valence or magnitude). However, some subjects could rate approachability more extremely than expected, which would have led to a skewed distribution of feedback types. For instance, for a participant who rated the approachability of the 24 faces between 1, 2 and 3, the probability of being exposed to agreement feedbacks and positive disagreement feedbacks (of both moderate and high magnitude) was very high, while the probability of being exposed to negative disagreement feedbacks was very low, and null for negative disagreement feedbacks of high magnitude (i.e., the group feedback could deviate up or down the participant's rating of 2 or 3 points). Such a strategy could result in a rating in test trials that was very low or very high on average, while we expected subjects to rate faces around the intermediate ranges of the approachability scale. Our pilot data

acquired from a sample of 125 participants are in line with this prediction: the median approachability rating was 4.66 ± 0.7 (95% CI, lower = 0.63, higher = 0.81), indicating that our face stimuli were indeed perceived as emotionally neutral. Subjects showing an average test-rating exceeding the mean of the group by $\pm 2SD$ were discarded from the analysis ($n = 3$).

Social influence index (SII) and Mean rating adjustment (MRA)

Participants' propensity to adjust her/his approachability ratings as a function of the group ratings was assessed by computing the average rating difference between initial and post-test ratings. On one hand, the mean rating adjustment (MRA) could be signed by the valence of the rating adjustment (in which case a positive MRA indicated a more positive rating in post-test trials and a negative MRA a more negative rating). On the other hand, the social influence index (SII) could be signed by the magnitude of the disagreement relative to the initial test rating (in which case a positive SII indicated that participants aligned their ratings on social information, whereas a negative SII indicated that their ratings went away from social information, and so independently of the valence dimension).

Model description and fitting

The fitted computational model hypothesizes that the decision to adjust a rating after the integration of social information is formed on the basis of a comparison between the faces presented in post-test trials and the type of social information it is paired with (group rating). The model consists of two free parameters, fitted to each participant's behaviour: 1) a social influence index δ corresponding to the adjustment of an initial rating in post-test trials (superior to zero for adjustments using social information, equal to or inferior to zero for adjustments not using social information), measured as the signed fraction of disagreement between the initial rating and the subsequent group rating, and 2) an internal noise magnitude σ corresponding to the standard deviation of the post-test rating. The mean rating in post-test trials μ thus corresponds to a linear combination between the initial rating x_{ini} and the group rating x_{group} following:

$$\mu = x_{\text{ini}} \cdot (1 - \delta) + x_{\text{group}} \cdot \delta$$

The probability of choosing the discrete rating x in post-test trials can be computed using the following equation:

$$p(x) = \Phi\left(x + \frac{1}{2}, \mu, \sigma\right) - \Phi\left(x - \frac{1}{2}, \mu, \sigma\right)$$

where $\Phi(\cdot)$ is the cumulative normal density function.

We obtained maximum-likelihood estimates of the two parameters δ and σ separately for each participant's behaviour using gradient descent of the negative model likelihood using the 'interior-point' algorithm of the *fmincon* routine implemented in MATLAB (Mathworks, Natick, MA). We derived model

predictions in terms of SII and MRA for all measures that were made directly from participants' behaviour, as means to test the adequacy of the model.

Assessing childhood adversity and its unpredictability and harshness dimensions

Exposure to adverse environments, including its unpredictability and harshness dimensions was computed from scores obtained in established questionnaires³²⁻³⁶.

Unpredictability was assessed following the methods developed by Mittal and colleagues (2015)³⁵. Participants first read the following instructions: "Think back to your life when you were younger than 10. This time includes preschool, kindergarten, and the first few years of elementary school." They were then asked to say how much they agreed with the following three statements whose aim is to estimate the extent to which participants grew-up in an unpredictable environment: "When I was younger than 10... : (a) things were often chaotic in my house, (b) people often moved in and out of my house on a pretty random basis, and (c) I had a hard time knowing what my parent(s) or other people in my house were going to say or do from day-to-day." Responses to these three items were made on a 7-point scale, from 1: strongly disagree, to 7: strongly agree. Scores on the 3 items were

averaged and z-scored to provide a childhood unpredictability index. Thus, the higher this index is, the less predictable the childhood environment is.

Consistent with prior research³⁷⁻³⁹, we assessed participants' childhood socioeconomic status (SES) as a proxy measure for early-life environmental harshness. We used established items³²⁻³⁶, asking participants to say how much they agreed with the following three statements: (a) "My family usually had enough money for things when I was growing up," (b) "I grew up in a relatively wealthy neighborhood," (c) "I felt relatively wealthy compared to the other kids in my school". Responses to SES items were made on a 7-point scale, from 1: strongly disagree, to 7: strongly agree. Scores on each item were reversed, averaged and z-scored to provide an SES index that matched with the unpredictability index. Thus, the higher the SES index was, the harsher the childhood environment was.

Finally, the z-scored unpredictability and SES indices were summed together in order to obtain a single index of childhood adversity in order to test the joint effect of both environmental dimensions on social influence.

Childhood harshness, unpredictability and their combination into an adversity index will be used as continuous variables in independent linear mixed-effects models²⁴ fitted with the software Matlab version R2014b to predict inter-individual variation in the susceptibility to social influence.

Target study

Based on the results that were obtained from the 122 participants who performed the pilot study, we estimated that the sample size necessary to obtain a main, positive effect of childhood SES on social influence (independently of the magnitude and valence dimension of social information) is $N = 223$, with a power of .90. A minimum of 223 healthy human adults will thus be recruited for the target study.

The stimuli, procedure, computational model of behaviour and statistical analyses that we aim to use in the target study will be similar to the pilot study. Few differences should be highlighted however.

First, we will modify the distractor task that separated the test and the post-test phase of the pilot experiment. Our aim here is to better control for a potential effect of aversion on the subsequent participants' behaviour. Indeed, even though the images we used in the pilot studies were judged low on the disgust dimension, they could nonetheless convey a very small aversive emotional response. Therefore, instead of judging neutral images on the disgust dimension, we will ask participants to compare the surfaces of two squares, i.e., a grey square and a black square (Figure 5). Eighteen trials will be performed. The image featuring the two squares will be presented for 3 seconds. The side on which the grey and black squares appear will be counterbalanced across trials. A numerical scale will then appear on the screen until participants select a value comprised between 1 and 8, representing how much

bigger the surface of the black square is compared to the grey square (1 = same size, 8 = 8 times bigger). The square task will last approximately 70 secs.

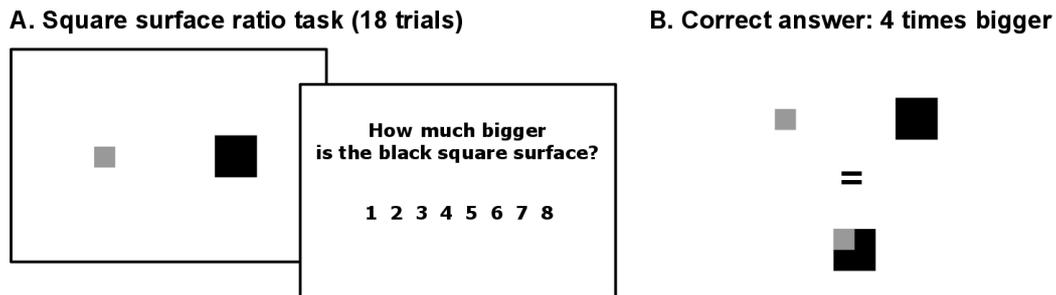


Figure 5. Distracting task of the target study: square surface ratio task. A. Participants will be asked to estimate the ratio between the surface of a black square varying across trials, and the surface of a grey square held constant across trials. Sixteen trials will be performed and the ratio between the two square surfaces could vary from 1 to 8. Each image featuring the two squares will be presented for 3 seconds. After each image, participants will select a value on an 8-point scale, representing how much bigger the surface of the black square is compared to the grey square (1 = same surface; 8 = 8 times bigger). **B.** In the example featured on the left panel of the figure, the black square surface is 4 times bigger than the grey square surface.

Second, we will assess the participants' current SES as a proxy measure for current environmental harshness. We will use items adapted from standardized questionnaires³²⁻³⁶ asking participants to say how much they agree with the

following three statements: (a) “My family and I have enough money for things” (b) “I currently live in a relatively wealthy neighborhood,” (c) “I feel relatively wealthy compared to other people of my neighborhood”. Responses to SES items were made on a 7-point scale, from 1: strongly disagree, to 7: strongly agree. Scores on each item will be reversed, averaged and z-scored to provide a current SES index that matched with the childhood indices described above. Thus, the higher the current SES will be, the harsher the current environment is.

The same approach will be used for unpredictability. We will adapt standard items used in the pilot study, and will ask participants to estimate the extent to which their current environment is unpredictable: “In the past few years: (a) things are often chaotic in my life, (b) people often move in and out of my neighborhood on a pretty random basis, and (c) I have a hard time knowing what people in my family or in my neighborhood are going to say or do from day-to-day.” Responses to these three items will be made on a 7-point scale, from 1: strongly disagree, to 7: strongly agree. Scores on the 3 items will be averaged and z-scored to provide a current unpredictability index. Thus, the higher this index is, the less predictable the current environment is.

Aknowledgement

Pierre O. Jacquet was supported by the LabEX IEC. We are very grateful to Alex Todorov and Ron Dotsch for having given us access to their computerized faces database.

Competing interests

The authors declare no competing financial interests.

References

1. Schaller M. Evolutionary bases of first impressions. In: Ambady N & Skowronski JJ, editors. *First impressions*. New York: Guilford Press; 2008. p. 15-34.
2. Sechrist GB, Stangor C. Perceived consensus influences intergroup behavior and stereotype accessibility. *J Pers Soc Psychol* 2001; 80:645-54.
3. Jones BC, DeBruine LM, Little AC, Burriss RP, Feinberg DR. Social transmission of face preferences among humans. *P Roy Soc B Bio* 2007; 274:899-903.
4. Little AC, Burriss RP, Jones BC, DeBruine LM, Caldwell CA. Social influence in human face preference: men and women are influenced more for long-term than short-term attractiveness decisions. *Evol Hum Behav* 2008; 29:140-46.

5. Klucharev V, Hytönen K, Rijpkema M, Smidts A, Fernández G. Reinforcement learning signal predicts social conformity. *Neuron* 2009; 61:140-51.
6. Shestakova A, Rieskamp J, Tugin S, Ossadtchi A, Krutitskaya J, Klucharev V. Electrophysiological precursors of social conformity. *Soc Cogn Affect Neurosci* 2013; 8:756-63.
7. Molleman L, van den Berg P, Weissing FJ. Consistent individual differences in human social learning strategies. *Nat Commun* 2014; 5:3570.
8. Richerson PJ, Boyd R. Not by genes alone: how culture transformed human evolution. The University of Chicago Press; 2005.
9. van Schaik CP, Burkart J, Damerius L, Forss SIF, Koops K, van Noordwijk MA, Schuppli C. The reluctant innovator: orangutans and the phylogeny of creativity. *Phil Trans R Soc B* 2016; 371:20150183.
10. Pepper GV, Nettle D. Perceived extrinsic mortality risk and reported effort in looking after health. *Human Nature* 2014; 25:378-92.
11. Belsky J. Attachment, mating and parenting: An evolutionary interpretation. *Human Nature* 1997; 8:361-81.
12. Nettle D, Dickins TE, Coall DA, de Mornay Davies P. Patterns of physical and psychological development in future teenage mothers. *Evolution, Medicine, & Public Health*. 2013:187-96.
13. Griskevicius V, Ackerman JM, Cantù SM, Delton AW, Robertson TE, Simpson J, Thompson ME, Tybur JM. When the economy falters, do people spend or

save? Responses to resource scarcity depend on childhood environments.

Psychol Science 2013; 24:197-205.

14. Baumard N, Chevallier C. The nature and dynamics of world religions: A life history approach. *P Roy Soc B Bio* 2015; 282:20151593.
15. Ellis BJ, Figueredo AJ, Brumbach B, Schlomer GL. Fundamental dimensions of environmental risk: the impact of harsh versus unpredictable environments on the evolution and development of Life History strategies. *Hum Nat* 2009; 20:204-268.
16. Adler NE, Boyce WT, Chesney MA, Folkman S, Syme SL. Socioeconomic inequalities in health: no easy solution. *Journal of the American Medical Association* 1993; 269:3140-3145.
17. Chen E, Matthews KA, Boyce WT. Socioeconomic differences in children's health: how and why do these relationships change with age? *Psychol Bull* 2002; 128:295-329.
18. Albrecht, C., & Teachman, J. D. (2003). Childhood living arrangements and the risk of premaritalintercourse. *Journal of Family Issues*, 24, 867–894.
19. Capaldi DM, Crosby L, Stoolmiller M. Predicting the timing of first sexual intercourse for at-risk adolescent males. *Child Dev* 1996; 67:344–359.
20. Woodward L, Fergusson DM, Horwood LJ. Risk factors and life processes associated with teenage pregnancy: Results of a prospective study form birth to 20 years. *Journal of Marriage and the Family* 2001; 63:1170–1184.

21. Wu LL. Effects of family instability, income, and income instability on the risk of premarital birth. *American Sociological Review* 1996; 61:386–406.
22. Wu LL, Martinson BC. Family structure and the risk of premarital birth. *American Sociological Review* 1993; 58:210–232.
23. Mittal C, Griskevicius V, Simpson JA, Sung S, Young ES. Cognitive adaptations to stressful environments: when childhood adversity enhances adult executive function. *J Pers Soc Psychol* 2015; 4:604-621.
24. Pinheiro JC, Bates DM. *Mixed-effects models in S and S-PLUS*. 2000. New York: Springer-Verlag.
25. Curtis V, Aunger R, Rabie T. Evidence that disgust evolved to protect from risk disease. *Proc R Soc Lond B (Suppl.)* 2004; 271:S131–S133.
26. Oosterhof NN, Todorov A. The functional basis of face evaluation. *Proc Natl Acad Sci USA* 2008; 105:11087-92.
27. Willis J, Todorov A. First impressions: Making up your mind after 100 ms exposure to a face. *Psychol Sci* 2006; 17:592-598.
28. Todorov A, Duchaine B. Reading trustworthiness in faces without recognizing faces. *Cogn Neuropsychol* 2008; 25:395-410.
29. Zaki J, Schirmer J, Mitchell JP. Social influence modulates the neural computation of value. *Psychol Sci* 2011; 22:894-900.
30. Huang Y, Kendrick KM, Yu R. Conformity to the opinions of other people lasts for no more than 3 days. *Psychol Sci* 2014; 27:1388-93.

31. Simonsen A, Scheel-Krüger J, Jensen M, Roepstorff A, Møller A, Frith CD, Campbell-Meiklejohn D. Serotonergic effects on judgments and social learning of trustworthiness. *Psychopharmacology* 2014; 231:2759-69.
32. Griskevicius V, Delton AW, Robertson TE, Tybur JM. Environmental contingency in life history strategies: the influence of mortality and socioeconomic status on reproductive timing. *J Pers Soc Psychol* 2011; 100:241–254.
33. Griskevicius V, Ackerman JM, Cantú SM, Delton AW, Robertson TE, Simpson JA, Tybur JM. When the economy falters, do people spend or save? Responses to resource scarcity depend on childhood environments. *Psychol Sci* 2013; 24:197–205.
34. Mittal C, Griskevicius V. Sense of control under uncertainty depends on people’s childhood environment: A life history theory approach. *J Pers Soc Psychol* 2014; 107:621-637.
35. Mittal C, Griskevicius V, Simpson JA, Sung S, Young ES. Cognitive Adaptations to Stressful Environments: When Childhood Adversity Enhances Adult Executive Function. *J Pers Soc Psychol* 2015; 109:604–621.
36. Hill SE, Prokosch ML, DelPriore DJ, Griskevicius V, Kramer A. Low Childhood Socioeconomic Status Promotes Eating in the Absence of Energy Need. *Psychol Sci* 2016; 27:354-364

37. Nettle D. Dying young and living fast: Variation in life history across English neighborhoods. *Behav Ecol* 2010; 21:387-395.
38. Belsky J, Schlomer GL, Ellis BJ. Beyond cumulative risk: Distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Dev Psychol*, 2012; 48:662-673.
39. Simpson JA, Griskevicius V, Kuo SIC, Sung S, Collins WA. Evolution, stress, and sensitive periods: The influence of unpredictability in early versus late childhood on sex and risky behavior. *Dev Psychol* 2012; 48:674-686.

Formations suivies

2015 Séminaire communication écrite et orale (IFD – 9h)

2015 Séminaire pratiques managériales – principes généraux (IFD – 4h)

2015 Evaluating social programs (MITx – 20h)

2016 Séminaire entrepreneuriat innovant (IFD – 8h)

2017 Colloque de Roscoff (ED3C – 13h)

2017 Présentation d'entreprise – Métiers de l'édition (IFD – 4h30)

2017 Aegina summer school – the social brain (University of London – 33h)

Nombre total d'heures de formation : 91h30