Rank reversal aversion and fairness in hierarchies

Alexandre Foncelle, Elodie Barat, Jean-Claude Dreher, Jean-Baptiste van Der Henst

Adaptive Human Behavior and Physiology

Rank reversal aversion and fairness in hierarchies

--Manuscript Draft--

Manuscript Number:	AHBP-D-22-00022R2
Full Title:	Rank reversal aversion and fairness in hierarchies
Article Type:	Original Article
Order of Authors:	Alexandre Foncelle
	Elodie Barat
	Jean-Claude Dreher
	Jean-Baptiste Van der Henst
Abstract:	Abstract Objectives Despite the aversion to inequality in humans, social hierarchies are a fundamental feature of their social life. Several mechanisms help explain the prevalence of hierarchies over egalitarianism. Recent work has suggested that while people tend to reduce resource inequalities when given the opportunity, they are reluctant to do so when it results in a reversal of social ranks (Xie et al., 2017). In this study, we explore how the way in which hierarchies are established influences this mechanism. We propose that aversion to rank reversal depends on whether rank asymmetry is fair or unfair. Methods In an online study, participants read 12 vignettes depicting six hypothetical hierarchies that varied in fairness. In each vignette, one individual was endowed with more resources than another individual, and participants could reduce that inequality by transferring resources from the higher-ranked individual to the lower-ranked one. In half of the vignettes, reducing the inequality led to a reversal of ranks, while in the other half it did not. Results We observed that participants were more likely to reverse ranks and reduce inequality when the hierarchy was perceived as unfair. Conclusion Overall, our results suggest that considerations of fairness guide participants' in their decision to reverse ranks.
Corresponding Author:	Jean-Baptiste Van der Henst UMR 5292: Centre de Recherche en Neurosciences de Lyon FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	UMR 5292: Centre de Recherche en Neurosciences de Lyon
Corresponding Author's Secondary Institution:	
First Author:	Alexandre Foncelle
First Author Secondary Information:	
Funding Information:	Fondation de France (89590) — Dr Jean-Claude Dreher
Order of Authors Secondary Information:	
Author Comments:	This manuscript is submitted as part of a special issue edited by Dr Khandis Blake and Dr Pranjal Mehta on "Status, Power, and Social Hierarchy"

Rank reversal aversion and fairness in hierarchies

Alexandre Foncelle[*1], Elodie Barat[*1,2], Jean-Claude Dreher[3], Jean-Baptiste Van der Henst[1]

* equally contributed to the work

Affiliations:

[1]Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028, Université Claude Bernard de Lyon, Université Jean Monnet de Saint-Etienne

[2]École de Psychologues Praticiens – Institut Catholique de Paris

[3]Institut des Sciences Cognitives Marc Jeannerod – CNRS UMR 5229 – Université Claude Bernard de Lyon

Corresponding authors:

Alexandre Foncelle: alexandre.foncelle@gmail.com

Jean-Baptiste Van der Henst: jean-baptiste.vanderhenst@cnrs.fr

Keywords: Hierarchy, rank reversal, fairness, redistribution

Abstract

Objectives

Despite the aversion to inequality in humans, social hierarchies are a fundamental feature of their social life. Several mechanisms help explain the prevalence of hierarchies over egalitarianism. Recent work has suggested that while people tend to reduce resource inequalities when given the opportunity, they are reluctant to do so when it results in a reversal of social ranks (Xie et al., 2017). In this study, we explore how the way in which hierarchies are established influences this mechanism. We propose that aversion to rank reversal depends on whether rank asymmetry is fair or unfair.

Methods

In an online study, participants read 12 vignettes depicting six hypothetical hierarchies that varied in fairness. In each vignette, one individual was endowed with more resources than another individual, and participants could reduce that inequality by transferring resources from the higher-ranked individual to the lower-ranked one. In half of the vignettes, reducing the inequality led to a reversal of ranks, while in the other half it did not.

Results

We observed that participants were more likely to reverse ranks and reduce inequality when the hierarchy was perceived as unfair.

Conclusion

Overall, our results suggest that considerations of fairness guide participants' in their decision to reverse ranks.

Rank reversal aversion and fairness in hierarchies

Hierarchy is a central structuring feature of nearly all human societies and non-human primate groups, and is thus considered as an elementary form of sociality (Fiske, 1992). Among humans, hierarchies manifest themselves in different forms such as power, dominance, occupational status or prestige, and generally involve inequality in control of desired resources, whether material or symbolic. Several factors are likely to create and perpetuate hierarchies at the interpersonal and societal levels. Scholars have argued that social hierarchy has an adaptive functional value that consists of increasing social coordination, cooperation, and peacemaking (Halevy et al., 2011; Magee & Galinsky, 2008). Moreover, from a psychological standpoint, humans seem to be endowed with cognitive mechanisms biased toward asymmetrical relationships as hierarchies tend to be more easily detected, understood and remembered than other types of relationships (Zitek & Tiedens, 2012), and also require less deliberative thought than egalitarian relationships (Van Berkel et al., 2015). Hierarchy is also a dimension of the social world that is understood the earliest by children, especially when it manifests itself through dominance relationships (Mascaro & Csibra, 2012; Thomsen et al., 2011). For instance, before the age of one, infants use body size and the number of allies to predict who will prevail in a right-of-way conflict (Pun et al., 2016; Thomsen et al., 2011).

Moreover, attitudes towards hierarchy are likely to contribute to its legitimization. For instance, Social Dominance Orientation (SDO) is described as a personality trait characterizing the degree to which individuals desire and promote hierarchies between social categories (Pratto et al., 1994). Social Dominance Orientation leads to the development of attitudes that legitimize the mechanisms that produce institutional and behavioral inequalities, which in turn reinforce social hierarchies (Pratto et al., 1994). Other researchers have

proposed that social hierarchies are not only maintained through in-group favoritism held by dominant groups, but also through system-justification mechanisms in subordinate groups that result in outgroup favoritism and in the perpetuation of inequalities at their own expense (Jost & Banaji, 1994). These mechanisms rely on stereotypes and the belief in a just world, but they also "capture social and psychological needs to imbue the status quo with legitimacy and to see it as good, fair, natural, desirable, and even inevitable" (Jost et al., 2004), p. 887).

However, in tension with factors favoring hierarchies, there is much evidence that people have a strong concern for equality. The notion is at the heart of several state constitutions and national mottos (e.g. *Liberty, Equality, Fraternity* in France), and occupies a central place in the Universal Declaration of Human Rights. Equality largely shaped social exchanges and the sharing of resources in hunting and gathering societies (Boehm, 1993; von Rueden, 2020). Moreover, multi-nation studies indicate that more people prefer equality between groups (Fischer et al., 2012), more equal distribution of wealth (Norton et al., 2014) and greater pay equality between CEOs and unskilled workers (Kiatpongsan & Norton, 2014). In most societies, humans show remarkable adherence to principles of fairness and equality, underpinned by different mechanisms including social comparison, loss aversion or emotion (Dawes et al., 2007; Fehr & Schmidt, 1999). People not only show an aversion to distributions that disfavor them (i.e. disadvantageous inequity), but they are also averse to unequal situations that benefit them (i.e. advantageous inequity; Bechtel et al., 2018; Fehr & Schmidt, 1999; Loewenstein et al., 1989). Moreover, a majority of individuals are willing to pay to enforce equality norms between third parties even if these norms do not benefit them (Fehr & Fischbacher, 2004; Henrich et al., 2006). This manifestation of costly punishment has been reported in a large variety of societies (Henrich et al., 2006), and is hypothesized as a means to promote cooperation (Fehr & Fischbacher, 2004).

Also noteworthy is the fact that egalitarian concerns emerge in infancy and are observed in both predictions and preferences. Indeed, infants from 9 months of age expect resources to be divided equally between two recipients (Buyukozer Dawkins et al., 2019; Meristo et al., 2016) and from 13 months of age, they prefer to interact with an individual who distributed resources equally rather than unequally (Burns & Sommerville, 2014; Lucca et al., 2018). Moreover, in first-party tasks, 3-year-olds who received more resources share them equally with another child with whom they collaborated, or sacrifice their resources to reach equality (Hamann et al., 2011; Ulber et al., 2017). However, sensitivity to equality is conceived within the framework of moral motivations that have emerged through evolutionary pressures (Tomasello & Vaish, 213). Egalitarian preference is not the default mechanism of human social attitudes but is guided by interests in fairness (Starmans et al., 2017). People distribute resources equally or enforce norms of equality to ensure that no individual is unduly advantaged or disadvantaged relative to another, and at a more general level to promote cooperation and reciprocity.

In summary, it appears that antagonistic tendencies can influence how people deal with social hierarchies. On the one hand, cognitive processes biased towards hierarchical asymmetries as well as legitimation processes lead to maintaining hierarchies. On the other hand, egalitarian and fairness concerns can lead to the reduction or even elimination of inequalities induced by hierarchies. Given this antagonism, a key question is therefore to determine the factors that favor certain mechanisms over others. In the present work, we explore how the way hierarchies are established influences people's choices for and against equality. To better understand the articulation of egalitarian and hierarchical tendencies, one strategy is to pit them against each other to determine which prevails. A particularly interesting case is when the reduction of equality leads to the reversal of the hierarchy. Using a redistribution game, a recent study specifically examined this issue (Xie et al., 2017).

In Xie et al.'s (2017) study, participants saw pictures of two players who had received random computer-generated allocations, resulting in resource asymmetries that determined their respective social ranks. Then, participants were given the possibility to redistribute the allocated resources by transferring a certain amount from the richer (higher rank) player to the poorer (lower rank) player. In the four types of situations participants dealt with, the transfer always resulted in a reduction of inequality, but in one of these situations, the transfer also resulted in hierarchy reversal (i.e., the richer player became the poorer player, and vice versa). For instance, in one situation player A was endowed with ￥4, and player B was endowed with ￥1. Participants then had to decide whether to transfer ￥2 from the richer to the poorer. This transfer reduces the inequality (￥1 difference instead of ￥3) but also reverses the hierarchy as the richer becomes the poorer (A has ￥2 and B has ￥3). Results indicated that overall, 55% of participants (including Huan Chinese, Indians, and Caucasians) rejected the transfer when it led to a reversal of hierarchy, while only 23% did so when the transfer did not lead to a reversal of hierarchy. Hence, egalitarian tendencies prevail as long as initial hierarchies are preserved, but if reducing inequality leads to a reversal of the hierarchy, then people are reluctant to reduce inequality, and keep social ranks unchanged.

The purpose of the current study is to extend these results by manipulating how the hierarchy is implemented, and to examine whether participants take this information into account when deciding to reverse or to preserve social ranks. In Xie et al.'s (2017) study, hierarchy emerged independently of players' controls because their resources, and therefore their ranks, were determined by chance. Although arbitrary, chance is sometimes considered as a relatively fair way to make decisions (Bolton et al., 2005; Kimbrough, Sheremeta, & Shields, 2014). We often rely on chance when we have no good reason to choose one option over another or when we think our personal opinion might be biased. Indeed, flipping a coin, or using other random procedures, are quite common ways of deciding who will have an

advantage in an unequal situation, such as attributing the first-move advantage in chess (by playing white) or the crown in the galette des rois. Even children consider flipping a coin to be a fair procedure to decide who should be advantaged in third-party and first party tasks (Shaw et al., 2014; Shaw & Olson, 2014).

Consistent with this argument, some participants in the Xie et al.'s study may have viewed social ranks as being fairly assigned. Aversion to rank reversal may thus result in part from fairness considerations about how the hierarchy was established rather than from the desire to maintain the hierarchy whatever it may be. Of course, some hierarchies may seem even more fair or unfair than the one implemented in Xie et al's study. In fact, hierarchies rarely occur in a purely random fashion. They are most often the result of actions produced voluntarily by social agents according to their motivation and their ability to achieve a higher status. Importantly, these actions can be evaluated based on the legitimacy of the costs and benefits to individuals in the hierarchy. A social rank acquired through coercion or spoliation is likely to be perceived as more unfair than a rank acquired through merit.

Egalitarian and hierarchical tendencies are intuitively conceived as two radically different kinds of attitudes about how people view social order. However, preferring fair treatment between individuals can lead to fostering not only equality but also hierarchy. People may thus prefer hierarchy to equality if the former is perceived as fairer than the latter, and vice versa. In a recent proposal, Starmans et al. (2017) precisely argued that in much experimental work defending the existence of inequality aversion, situations of resource inequality presented to participants were confounded with economic unfairness. For these authors, it is therefore more accurate to interpret the conclusions following from these studies as an aversion to unfairness than as an aversion to inequality resulting from egalitarian tendencies. As Starmans et al. point out, preferences for merit-based inequalities, or

inequalities that favor helping behaviors, emerge as early as preschool age and prevail over allegedly egalitarian tendencies.

In the same vein, we consider here the idea that if people perceive a hierarchy as unfair, they may decide to reverse social ranks. Recent research in developmental psychology has shown evidence of rank reversal in the decision-making power. For example, in a study by Charafeddine et al. (2016), 8 years-old children who watched a higher-ranked puppet who repeatedly imposed their choice on a lower-ranked puppet decided to allocate more resources to the latter than to the former and referred to fairness considerations to justify their distribution (see also Cheng et al., 2021). Faced with a hierarchy of power, they established a reverse hierarchy based on material resources. However, these hierarchies were different in nature since one involved the power to decide and the other involved material resources. It thus cannot be said that the initial power hierarchy was reversed *per se*.

To provide a better understanding of people's willingness to reverse rank, we presented participants with 12 vignettes depicting six hypothetical hierarchies that varied in fairness. All hierarchies involved an inequality of resources, but like Xie et al., participants could redistribute resources to reduce inequality. In half of the vignettes, redistribution resulted in rank reversal, such that the higher-ranked individual became the lower-ranked one (and vice versa). We tested the prediction that participants would be more likely to reverse ranks between the two individuals when the hierarchy was established unfairly and we examined whether the unfairness of hierarchies neutralized rank reversal aversion.

Method

Participants

We used the online platform Prolific Academic (Palan & Schitter, 2018) to recruit 309 adults (155 females, $M_{age} = 39.5$, $SD = 11.7$). A description of the study was posted to

Prolific to identify eligible participants. Eligibility criteria were as follows: (1) age 18 to 60 years, (2) residence in England, and (3) native English speaker. An a priori power analysis was performed using G*Power version 3.1.9.7 (Faul et al., 2009) to estimate the minimum sample size needed to find a statistically significant effect in the model. The results indicated that the sample size required to obtain 80% power for detection of a small effect (Cohen's f = 0.05), at a significance criterion of $\alpha = 0.05$, was N = 282 for an ANOVA with repeated measures, between and within interactions. By adding about 10% for participants for possible missing data, we recruited 309 participants. Actually, all participants rated the 12 trials they received.

Material and design

Fairness Manipulation

The stimuli consisted of six different vignettes involving two individuals, A and B, who received unequal endowments of pounds. In one of the vignettes, ranks were randomly assigned as in Xie et al.'s study, but in all other vignettes, the ranks resulted from A and B's actions or resulted from a previously established hierarchy between them.

Thus, the six vignettes, which are presented in the Online Supplementary Materials, were based on:

i) *Merit*, where A worked harder than B,

ii) *Chance*, as in Xie et al.'s study,

iii) *Competition*, where A and B competed for more money,

iv) *Pre-existing hierarchy*, where A was described as B's superior,

v) *Despotic power*, where A arrogated to themselves the power to decide on the distribution of resources.

vi) *Theft*, where A stole resources from B.

To obtain a measure of fairness for each vignette, we conducted a rating study with a sample of 125 participants (64 females, M_{age} = 35.1, SD = 10.9) who were recruited online in the same manner as participants in the transfer experiment. An a priori power analysis was performed using G*Power version 3.1.9.7 (Faul et al., 2009) to estimate the minimum sample size needed to find a statistically significant effect in the model. The results indicated that the sample size required to achieve 80% power for detection of a small effect (Cohen's f = 0.11) at a significance criterion of $\alpha = 0.05$, was N = 113 for an ANOVA with repeated measures within factors. Thus, adding approximately 10% to account for possible missing data, the resulting sample size of N = 125 was adequate to test the model. Participants were asked to rate the extent to which each of the 6 vignettes in Online Supplementary Materials seemed fair or unfair on a Likert scale (from 1 “Totally unfair” to 7 “Totally fair”). The endowments were those used in Reversal situations (£4 vs. £1).

The results indicated that Merit, Competition, and Chance were the vignettes perceived as the fairest, and that Theft, Despotic Power, and Pre-Existing Hierarchy were the vignettes perceived as the most unfair, and with Merit and Theft at opposite ends. Figure 1 shows the boxplots of the fairness measure for each vignette. The intervals between positions on the scale of fairness are monotonous but never so defined as to be numerically uniform increments. Hence, to estimate the extent to which the fairness measure increases or decreases between vignettes and whether the difference is significant, the fairness measure was considered as an ordered variable between 1 and 7. Then, a cumulative link mixed model with a random effect on participants (to control for the repeated measures design) was implemented (Table ESM2), and a post-hoc analysis with pairwise comparisons was conducted (Table 1). Fairness scores were significantly different in all pairwise comparisons, except for Theft vs. Despotic Power and for Chance vs. Competition and allowed the

vignettes to be used as an ordered factor in the statistical model of the analysis of hierarchy reversal (Table 1).

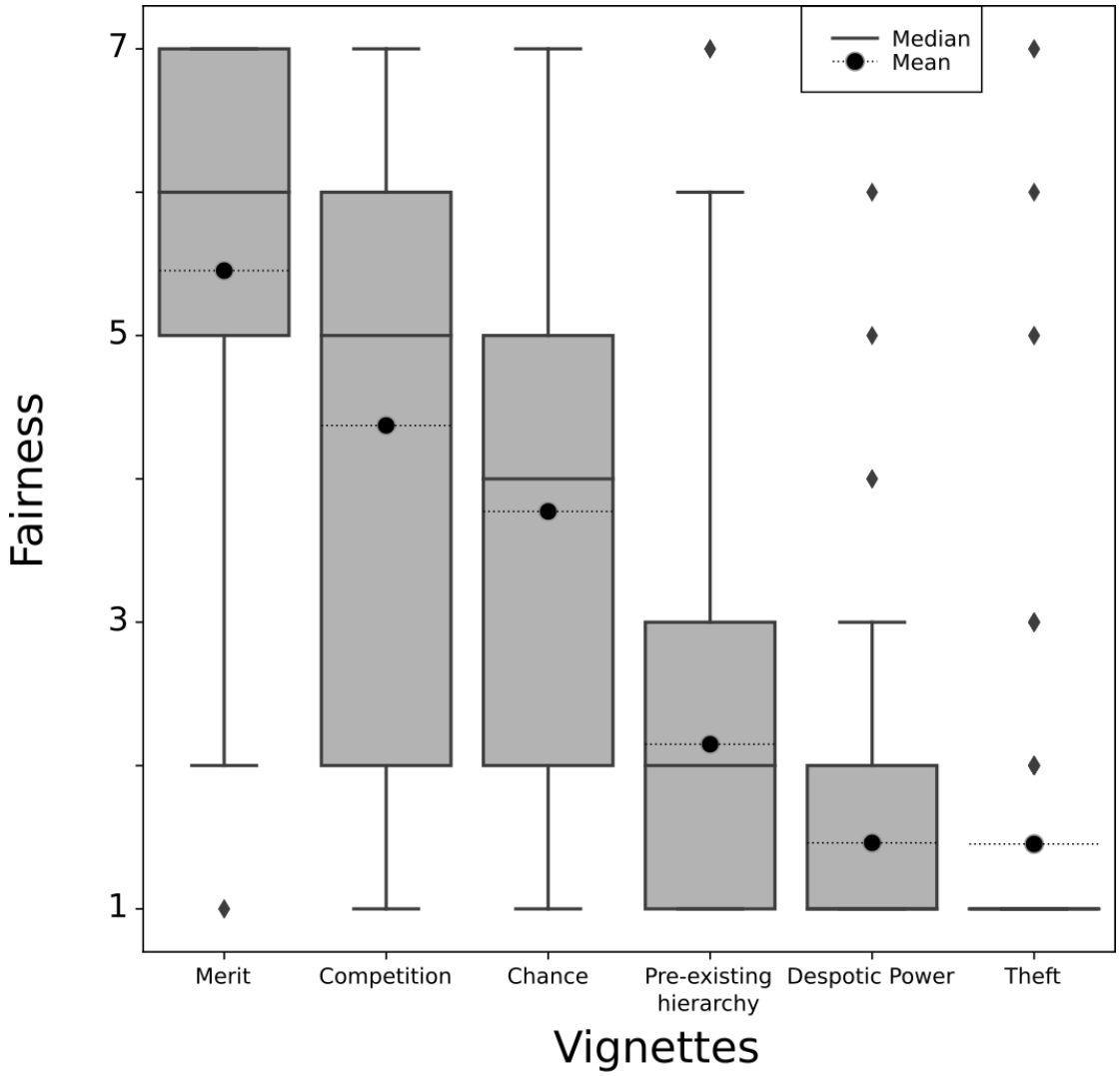


Fig. 1 Boxplots on the Likert scale (1-7) of fairness for each vignette. Mean ± SD: Merit: 5.5 ± 1.5, Competition: 4.4 ± 2.1, Chance: 3.8 ± 1.9, Pre-existing hierarchy: 2.1 ± 1.4, Despotic Power: 1.5 ± 0.8, Theft: 1.5 ± 1.1

Table 1. Pairwise comparisons of fairness between vignettes.

Vignette	**Merit** fairness = 5.5	**Competition** fairness = 4.4	**Chance** fairness = 3.8	**Pre-existing hierarchy** fairness = 2.1	**Despotic power** fairness = 1.5
Competition fairness = 4.4	**2.9 [1.8, 4.5]** ***p* < .001**				
Chance fairness = 3.8	**5.6 [3.4, 8.3]** ***p* < .001**	1.9 [1.2, 2.9] *p* = .07			
Pre-existing hierarchy fairness = 2.1	**33.3 [20.0, 50.0]** ***p* < .001**	**11.6 [7.0, 19.1],** ***p* < .001**	**6.2 [3.8, 9.9]** ***p* < .001**		
Despotic power fairness = 1.5	**107.8 [59.6, 194.8]** ***p* < .001**	**37.5 [21.4, 65.8]** ***p* < .001**	**20.0 [11.7, 34.1]** ***p* < .001**	**3.2 [2.0, 5.3]** ***p* < .001**	
Theft fairness = 1.5	**172.4 [90.5, 328.6]** ***p* < .001**	**60.0 [32.4, 111.0]** ***p* < .001**	**31.9 [17.6, 57.6]** ***p* < .001**	**5.2 [3.0, 9.0]** ***p* < .001**	1.6 [0.9, 2.9] *p* = .61

Note. Post-hoc comparisons on the cumulative link mixed model. For example, the Merit vignette is, on average, 33.3 times more likely to be evaluated as fairer than the Pre-existing hierarchy vignette. Results are reported with 95% CI and *p*-value. The mean fairness is reported under each vignette.

Rank reversal manipulation

The transfer situations proposed to participants varied according to the initial endowments and the amount of the transfer. In situations in which the proposed transfer led to a reversal of rank (Reversal condition), the higher-ranking individual had £4, the lower ranking individual had £1, and the proposed transfer was £2. For situations where the transfer led to the preservation of rank (Preservation condition), there were two possibilities. Either the initial endowments were identical to Reversal situations but with a lower transfer (i.e. initial endowments: £4 vs. £1, transfer: £1 - Preservation-£1 condition), or the transfer was identical to Reversal situations but with a higher endowment to the highest-ranking individual (i.e. initial endowments: £6 vs. £1, transfer: £2 - Preservation-£2 condition). We used these two preservation conditions to control for possible effects due to the difference in transfer (2£

vs. 1£) or the difference in initial endowments to the highest-ranking individual (6£ vs 4£). In all situations, the transfer led thus to the same level of inequality (i.e. £1).

Participants dealt with the six vignettes in Reversal situations as well as in Preservation situations and thus received 12 trials. About half of the participants received Reversal and Preservation-£1 situations and half received Reversal and Preservation-£2 situations. Hence, in the first variant of the binomial mixed-effects regression model of our analysis, the rank reversal condition variable had three levels (Reversal, Preservation-£1, and Preservation-£2). Pairwise comparisons in the post-hoc analysis showed that for each vignette there was no significant difference between Preservation-£1 and Preservation-£2. Therefore, we pooled the two preservation situations for further analysis (Table ESM1).

Gender manipulation

The gender of the individuals in the presented pictures was counterbalanced across vignettes, but the gender remained constant between the two rank reversal conditions of each vignette (Reversal and Preservation). Hence, if participants saw two women in the Reversal condition of the Merit vignette, they also saw two women in the Preservation condition of that vignette. Pictures were obtained from the KDEF database (https://www.kdef.se/index.html). The pictures from this database are standardized for posture, look and facial expression. In addition, we have matched the pairs by hairstyle and hair color.

Relevant covariates

The order of presentation of the vignettes was controlled by four counterbalanced runs of the 12 trials. Moreover, at the end of the survey, participants were asked to indicate their political position and their socioeconomic status (SES) on 9-point Likert scales (political position: 1 = conservative, 9 = liberal; SES: 1 = low, 9 = high).

Procedure

Participants did the experiment online on the SurveyMonkey platform. After they gave their consent, they received the following instructions: "You are going to take part in a short experiment in which two people - Person A and Person B - will be given different amounts of money, in 12 situations. For each situation you will have the option to make a transfer of money from one person to another." Participants were asked to make a yes or no decision on each of the 12 trials. The 12 trials were presented in random order.

Data analysis

First, we examined how rank reversal aversion was related to fairness. To do so, we reported participants' decision to reject the transfer in each trial. To analyze these data, we adopted a model-fitting approach on a binomial mixed-effects regression model. We started with a simple model (Tables ESM3 and ESM5), including the estimated fairness of the hierarchy vignette and the rank reversal condition as main fixed effects, and participants as a random effect (to control for repeated measures), and then added terms to the model to see if they were significant or improved model fit (e.g. the order of presentation of vignettes, participant gender, and picture gender). We performed Chi-squared tests to determine the best-fitting model (Table ESM6). Next, to characterize the effect size, we performed pairwise comparisons by running F-tests in post-hoc analysis. We reported the odds ratio of rejecting the transfer by taking the exponent of the estimate.

In the results section, we report the final reduced model of our fitting approach, its main effects and the interactions between these effects. The extraction of data and plots were performed in Python3 (Van Rossum & Drake, 2009) and statistical analysis were conducted withR (version 4.2.1) using the lme4, multcomp and emmeans packages (R. Core Team, 2018).

Results

To interpret the results, we evaluated the best model of our model-fitting approach (Tables ESM7 - ESM11) whose equation is as follows:

*Transfer ~ HierarchyFairness*RankReversalCondition*PictureGender +HierarchyFairness*Gender +RankReversalCondition*Gender +HierarchyFairness*HierarchyFairnessOrder +HierarchyFairness*SES +HierarchyFairness*Politics +(1|Participant)*

where *Transfer* is a binary factor variable, *HierarchyFairness* is an ordered factor based on fairness measure ("Theft" < "Despotic Power" < "Pre-existing hierarchy" < "Chance" < "Competition" < "Merit"), *RankReversalCondition* is a binary factor variable, *PictureGender* is a binary factor variable, *Gender* is a binary factor variable *HierarchyFairnessOrder* is a four-level factor variable, *SES* is an ordered factor, *Politics* is an ordered factor and *(1|Participant)* is the random intercept of the participant (Table ESM4).

The analysis of variance (Table ESM7) revealed a significant main effect of *HierarchyFairness* (χ^2 (5) = 686.9, p < .001). Given that our main hypothesis was that participants would be more likely to reverse the hierarchy when it was fair rather than unfair, pairwise comparisons of transfer rejection rates between vignettes are reported only for the Reversal condition (Table 2). In particular, to examine the reluctance to reverse ranks across vignettes, a post-hoc analysis estimated how many times participants were more likely to reject the transfer between two hierarchy fairness vignettes in the Reversal condition (Table 2). This analysis confirmed the prediction that when a hierarchy was unfair, participants were more likely to reverse the hierarchy than when it was fair (Table 2). Merit was perceived as the fairest hierarchy and it was also the one for which participants were most reluctant to accept the transfer. In contrast, Despotic Power and Theft were perceived as the most unfair hierarchies and were also the ones for which participants were most willing to accept the

transfer. It should also be noted that for the Chance vignette of the Reversal condition, 53.4% of participants rejected the transfer (Figure 2), a rate that was highly similar to the one reported by Xie et al. (i.e. 55.2%).

The analysis of the model also revealed a significant main effect of *RankReversalCondition* (χ^2 (1) = 78.6, p < .001) indicating that participants were more likely to reject the transfer in the Reversal condition than in the Preservation condition. Indeed, participants were 2 times more likely to reject the transfer (2.0 [1.6, 2.6], p < .001) in the former compared to latter condition. In addition, there was a significant interaction between the conditions (Reversal vs. Preservation) and the hierarchy fairness vignettes (χ^2 (5) = 14.0, p < .02). A post-hoc analysis estimated the odds of rejecting the transfer between the reversal and preservation conditions of each hierarchy fairness vignette (Figure 2 and Table 3). For Merit, Competition, Chance, and Pre-existing Hierarchy, the odds of rejecting the transfer in the Reversal condition were highly multiplied relative to the odds of rejecting the transfer in the Preservation condition. Moreover, this multiplier effect was at its maximum for Competition, and decreased at the extremes of the fairness scale, namely for Merit and Theft.

The interactions between *HierarchyFairness* and *PictureGender, SES, Politics* and *HierarchyFairnessOrder* were globally significant (Table ESM7) and even if the study was not designed to explore all pairwise comparisons of this model, we could investigate some pairwise comparisons for SES and Politics (Table ESM10 - ESM11). The triple interaction between *HierarchyFairness, RankReversalCondition* and *PictureGender* was also significant (χ^2 (5) = 13.0, p < .02). Studying some pairwise comparisons (Table ESM9), participants are significantly more likely to reject transfer in the Reversal condition for both Competition and Chance vignettes, compared to the Preservation condition, for both males and females in the pictures, although for the male pictures the odds are higher. Furthermore, *Gender* was globally non-significant (χ^2 (1) = 2.9, p = .090)) and other non-significant variables such as

PictureGender, *SES*, *Politics* and *HierarchyFairnessOrder* are reported in Table ESM7.

Random intercepts for participants improved the model compared to a version without it, according to a Chi-square test (χ^2 (1) = 259.1, p < .001).

Table 2. Odds ratios of transfer rejection between hierarchy fairness vignettes in the Reversal conditions.

Hierarchy Fairness Vignette	**Merit** fairness = 5.5	**Competition** fairness = 4.4	**Chance** fairness = 3.8	**Pre-existing hierarchy** fairness = 2.1	**Despotic power** fairness = 1.5
Competition fairness = 4.4	**3.1 [1.8, 5.2]** ***p*** **< .001**				
Chance fairness = 3.8	**10.7 [6.4, 17.8]** ***p*** **< .001**	**3.5 [2.2, 5.3]** ***p*** **< .001**			
Pre-existing hierarchy fairness = 2.1	**76.2 [43, 135]** ***p*** **< .001**	**24.6 [15, 40.5]** ***p*** **< .001**	**7.1 [4.5, 11.3]** ***p*** **< .001**		
Despotic power fairness = 1.5	**170.9 [91.4, 319.3]** ***p*** **< .001**	**55.3 [31.7, 96.4]** ***p*** **< .001**	**16.0 [9.5, 26.9]** ***p*** **< .001**	**2.2 [1.3, 3.9]** ***p*** **= .004**	
Theft fairness = 1.5	**205.4 [109.1, 386.9]** ***p*** **< .001**	**66.4 [37.7, 116.7]** ***p*** **< .001**	**19.2 [11.3, 32.5]** ***p*** **< .001**	**2.7 [1.5, 4.7]** ***p*** **< .001**	1.2 [0.7, 2.2] *p* = .55

Note. The values indicate the number of times a vignette in the header row was greater than the vignette in the column header. For example, participants were 24.6 times more likely to reject transfer for Competition compared to Pre-existing hierarchy. Fairness refers to the vignette fairness obtained in the rating study. Results are reported with 95% CI and *p*-value.

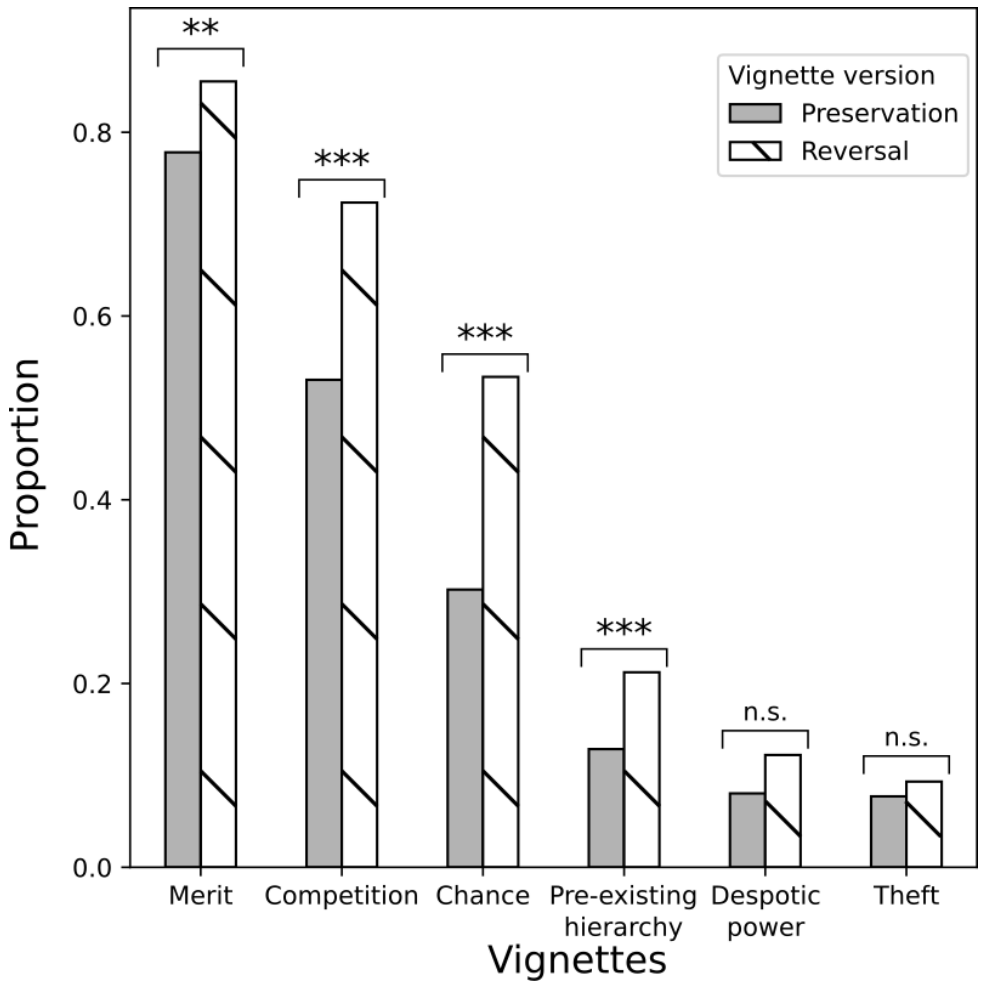


Fig. 2 Proportion of participants rejecting transfer for each hierarchy fairness vignette in Reversal and Preservation conditions. The significance of the comparisons is obtained from Table 3.

Table 3. Odds ratios of transfer rejection between hierarchy fairness vignettes in Reversal and Preservation conditions.

Hierarchy Fairness Vignette	**Merit** fairness = 5.5	**Competition** fairness = 4.4	**Chance** fairness = 3.8	**Pre-existing hierarchy** fairness = 2.1	**Despotic power** fairness = 1.5	**Theft** fairness = 1.5
Preservation to Reversal	**1.9 [1.2, 3.1]** ***p* = .009**	**3.3 [2.2, 4.9]** ***p* < .001**	**3.7 [2.5, 5.5]** ***p* < .001**	**1.9 [1.1, 3.1]** ***p* = .014**	1.7 [0.9, 3.1] *p* = .08	1.3 [0.7, 2.3] *p* = .47

Note. For example, for the Merit vignette, participants are 1.9 times more likely to reject the transfer in the Reversal condition compared to the Preservation condition. The mean fairness from our first experiment is reported under each vignette as an indicator. Results are reported with 95% CI and p-value.

General Discussion

Using a resource redistribution game, the current study investigated participants' decision to reverse social ranks in different types of hierarchies. Unlike non-human primate groups, where hierarchies rely mostly on dominance relationships, human societies give rise to a diversity of social orders that vary in fairness. Because fairness determines behavior in resource allocation tasks, we hypothesized that people may be inclined to reverse unfair hierarchies in order to reduce inequalities. Several main results emerged from the current study. First, the fairness of hierarchies did influence participants' decision to reverse ranks. The rating study confirmed that participants differentially judged the fairness of the vignettes and overall the study showed a clear parallel between the perceived fairness of vignettes and the transfer decisions. Participants were more likely to reverse ranks when the vignettes were more unfair. In particular, in the two most unfair situations, namely Theft and Despotic Power, about 90% of participants reversed social ranks.

Second, we did replicate the findings of Xie et al., who used a chance based-hierarchy. We found almost the same amount of rejection rate in the Reversal condition of the Chance vignette as they did (our study: 53.4%; their study: 55.2%). We also replicated the rank reversal aversion in the Chance vignette, since participants were more likely to reject the transfer in the Reversal condition than in the Preservation condition. Moreover, we also found evidence of rank reversal aversion in the four fairest situations. Indeed, the higher rate of rejections in Reversal than in Preservation situations was also observed in the Merit, Competition and Pre-existing Hierarchy vignettes. In contrast, aversion to rank reversal was absent for the most unfair vignettes, namely Theft and Despotic Power, which may be explained by the fact that these two situations were completely unfair, as revealed by the floor effect obtained for these situations in the rating study.

These results make sense in light of various findings in the literature that show critical postures towards certain forms of hierarchy. First, research on attitudes show that in third-party contexts, individuals who achieve a higher status through antisocial strategies, such as dominance and coercion, are judged as more negatively by adult and child observers than those who achieve it through prestige (Cheng et al., 2013; Kajanus et al., 2020).

Second, feelings people experience about hierarchy may lead them to take supportive actions toward low-status individuals and hindering actions toward high-status individuals. For instance, people tend to prefer the fall of a high achiever than the fall of an average achiever, and are more pleased by the fall of high-status individuals who do not fully deserved their position or who exhibit negative personality traits (Feather, 1994). In contrast, attitudes toward lower status individuals turn out to be more positive. Individuals of lower status who subvert hierarchies and triumph over those of higher status are often praised in popular cultures and mythologies, such as Rocky Balboa, Cinderella or David triumphing over Goliath. This positive stance has been experimentally evidenced by research on the underdog effect. Presented with situations of competition and international conflicts people tend to be more favorable to the underdog than to the top dog and attribute greater effort to the underdog's performance (Kim et al., 2008; Vandello et al., 2007). It has been argued that this positive view of the underdog operates as a means to restore a sense of fairness (Vandello et al., 2007, Quesque et al., 2021). It is therefore possible that the desire to support a lower rank individual is particularly active when a hierarchy is unfair.

Third, in unfair vignettes, the higher-status individual blatantly violates social norms such as the equality norm in distribution, the relational equality norm (e.g. Despotic Power), or the norms of respect and property (e.g. Theft), which may actually contribute to their being perceived as having higher status (VanKleef et al., 2011). In such a context, reducing inequality by reversing ranks is not only a way to restore equality, but also to punish the

higher-status individual who violates norms, even if the participant is not the victim of those violations. Sanctioning norm violation is precisely a way by which norms can be enforced and is willingly applied by third-parties even if it comes at a cost to them (Fehr & Fischbacher, 2004; Henrich et al., 2006).

A notable finding of the study is that while for the two most unfair vignettes (i.e. Despotic Power and Theft), participants reduced inequality in the same way when it led to rank reversal and when it led to rank preservation, this was not the case for Pre-Existing hierarchy. In this case, participants rejected the transfer more in the Reversal condition than in the Preservation condition. Hence, the mechanisms that lead to maintaining the hierarchical *status quo* can be triggered even when the hierarchy is unfair as for Pre-existing hierarchy. In other words, although unfairness in the hierarchy increases the likelihood of action against it, it does not guarantee such action. In the specific case of Pre-existing hierarchy, it is also possible that in addition to the aforementioned mechanisms that contribute to maintaining the status quo, face-saving values also influenced participants' choices. For people sensitive to such values, the reversal of ranks could constitute too great a punishment inflicted on the higher-status individual. Rank reversal aversion would thus be motivated by prosocial attitudes toward the higher-status individual.

Limitations and Future directions

In the rating study, the distribution of the most unfair vignettes did not mirror that of the fairest vignettes. Indeed, while Despotic Power and Theft, ratings were at the fairness floor (i.e. the means were at about 0.5 from the floor, see Figure 1), Merit was not at the fairness ceiling (i.e. the mean was at 1.5 from the ceiling, see Figure 1). In the event that a merit situation is seen as closer to the fairness ceiling, it might be useful to analyze whether the difference between the Reversal and Preservation conditions persists. Indeed, if

participants consider the inequality situation to be completely fair, they might feel that the resources obtained by each individual is perfectly justified, and decide to maintain the inequality as it is, and thus reject the transfer, not only in Reversal situations but also in Preservation situations. It might therefore be useful in future research to examine other situations where the inequality is considered fairer than that observed here for merit. At an even more general level, other more specific hierarchy situations could be studied to find out more precisely how participants evaluate the fairness of higher status individuals. For example, in the Competition vignette participants could not know exactly by which means the winner has won and in the Pre-existing hierarchy situation, they could not know to which type of hierarchy the two individuals belonged. Enriching these situations would better highlight the fairness or unfairness involved.

Regarding individual differences, we did observe a participant gender effect for Competition and Pre-existing Hierarchy, with males rejecting transfer more often than females. These results are in line with research showing that men endorse more positive beliefs about competition than women (Kesebir et al., 2019) and are more likely to endorse existing hierarchies as revealed by gender differences in SDO measures (Sidanius et al., 1994, 2006). The gender category of participants had thus an influence on their decision. However, a complementary line of research would be to explore gender further by examining how the gender of the stimuli affects participants' decisions. In the current study, a triple interaction indicates that the difference in transfer rejection between the Reversal and the Preservation condition was higher for male than for female pictures in some of the vignettes, namely Competition and Chance. This suggests that participants may be more conservative for some hierarchies when they involve men than when they involve women, in line with the stereotype that men are perceived as more hierarchical than women (Schmid-Mast, 2004). This could be further explored by examining more systematically transfer decisions in typical male

hierarchies (e.g. military hierarchies) when occupied by male or female characters. Moreover, given the existing hierarchies between men and women, one might wonder how participants would reverse ranks between a male and female individual, and whether equity considerations would influence transfer in the same way as in same-sex pairs.

In addition, it is also important to note that our results were obtained in the specific context of a population residing in England, and thus constitute a sample of a WEIRD population (Henrich, Heine, & Norenzayan, 2010). People from WEIRD countries share individualistic values that place more weight on meritocratic and competitive values and tend to challenge hierarchies relatively easily (Huppert et al., 2017; Shäfer, Haun, & Tomasello, 2015). Conversely, in collectivist cultures, which are more committed to group harmony, individuals place less weight on merit in resource distribution tasks but are also more sensitive to saving face, especially in the context of an established hierarchy (Oetzel et al., 2010). Thus, it may be useful to investigate how culture influences rank reversal in the context of our vignettes that may activate different fairness intuitions.

The current experiment and the one by Xie et al. placed participants in a third-person perspective. However, one question that arises and that would be worth examining is the attitude people may have toward their own rank when they are embedded in a hierarchy, and how likely they are to change rank to reduce inequality. There would obviously be two opposing situations for participants, one in which reducing inequality through rank reversal would result in a loosening of rank and one in which it would result in an increase in rank. Given individuals' vigilance about moving down the social hierarchy (Anderson et al., 2015), it seems clear that those in higher status positions should be much more likely to accept a reduction of inequality if it preserves their rank than if it reverses it. Things may be less clear for those in the lower status situation. In terms of their personal utility, it would be in their interest to accept a reduction in inequality, as this might be doubly beneficial to them: by

increasing their endowment and increasing their own rank. However, if rank-preserving mechanisms are at work, they may choose not to accept a change in rank in order to prevent the other player from suffering a social defeat due to rank reversal. Answering this question would allow us to know if the advantageous inequity aversion observed in the context of resource distribution also applies to rank reversal.

Another limitation concerns the implementation of status. In the current experiment, the status of the two individuals A and B is established by a difference in their monetary endowments. However, often status differences do not involve any financial reference, especially when it results from relational asymmetry such as for decision-making power, where one individual imposes their choices on another or exercises authority by giving orders to the subordinates. Even if in the current experiment the unequal endowment resulted from a diversity of contexts, participants could only act on the monetary part of status. To better understand attitudes toward social rank distinctions, it would be useful to compare participants' monetary actions on monetary hierarchies and their non-monetary actions on non-monetary hierarchies. The attitude of the participants via status reallocation actions would allow us to know if a relational hierarchy is perceived as a more rigid social order than a financial hierarchy or if, on the contrary, a relational hierarchy is perceived as being more unfair and more subject to being thwarted.

Conclusion

In conclusion, hierarchical structures are universal features of human societies and profoundly shape our dyadic interactions in a wide range of situations. When people are confronted with a hierarchical situation as a third-party actor, they may decide to endorse the social order resulting from the hierarchy or counter it to restore equality. This study shows that considerations of fairness guide participants to act for or against the preservation of social

ranks and provides insight into human political intuitions. Just as egalitarian tendencies often reflect a concern for fairness (Starmans et al., 2017), this same concern may also be at work in behaviors that favor rank order. Of course, this research does not establish that favoring hierarchy is exclusively motivated by considerations of fairness because, as some of our results have shown, preserving the status quo can occur even in the case of unfair hierarchies.

Data Availability: The data that support the findings of this study are available from the corresponding authors upon request.

Conflict of interest: On behalf of all authors, the corresponding authors state that there is no conflict of interest.

References

Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological bulletin, 141(3),* 574.

Bechtel, M. M., Liesch, R., & Scheve, K. F. (2018). Inequality and redistribution behavior in a give-or-take game. *Proceedings of the National Academy of Sciences*, *115*(14), 3611‐3616.

Boehm, C. (1993). Egalitarian behavior and reverse dominance hierarchy. *Current Anthropology*, *34*(3).

Bolton, G. E., Brandts, J., & Ockenfels, A. (2005). Fair procedures : Evidence from games involving lotteries. *The Economic Journal*, *115*(506), 1054‐1076.

Burns, M., & Sommerville, J. (2014). "I pick you" : The impact of fairness and race on infants' selection of social partners. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/article/10.3389/fpsyg.2014.00093

Buyukozer Dawkins, M., Sloane, S., & Baillargeon, R. (2019). Do Infants in the First Year of Life Expect Equal Resource Allocations? *Frontiers in Psychology*, *10*. https://www.frontiersin.org/article/10.3389/fpsyg.2019.00116

Charafeddine, R., Mercier, H., Clément, F., Kaufmann, L., Reboul, A., & Van der Henst, J.-B. (2016). Children's allocation of resources in social dominance situations. *Developmental Psychology*, *52*(11), 1843‑1857. https://doi.org/10.1037/dev0000164

Cheng, J. T., & Tracy, J. L. (2014). Toward a unified science of hierarchy : Dominance and prestige are two fundamental pathways to human social rank. In *The psychology of social status* (p. 3‑27). Springer.

Cheng, N., Wan, Y., An, J., Gummerum, M., & Zhu, L. (2021). Power grabbed or granted : Children's allocation of resources in social power situations. *Journal of Experimental Child Psychology*, *210*, 105192.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *nature*, *446*(7137), 794‑796.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Feather, N. T. (1994). Attitudes toward high achievers and reactions to their fall : Theory and research concerning tall poppies. In *Advances in experimental social psychology* (Vol. 26, p. 1‑73). Elsevier.

Feather, N. T., & Nairn, K. (2005). Resentment, envy, schadenfreude, and sympathy : Effects of own and other's deserved or undeserved status. *Australian journal of psychology*, *57*(2), 87‐102.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human behavior*, *25*(2), 63‐87.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817‐868.

Fischer, R., Hanke, K., & Sibley, C. G. (2012). Cultural and institutional determinants of social dominance orientation : A cross-cultural meta-analysis of 27 societies. *Political Psychology*, *33*(4), 437‐467.

Fiske, A. P. (1992). The four elementary forms of sociality : Framework for a unified theory of social relations. *Psychological review*, *99*(4), 689‐723.

Fragale, A. R., Overbeck, J. R., & Neale, M. A. (2011). Resources versus respect : Social judgments based on targets' power and status positions. *Journal of Experimental Social Psychology*, *47*(4), 767‐775.

Halevy, N., Y. Chou, E., & D. Galinsky, A. (2011). A functional model of hierarchy : Why, how, and when vertical differentiation enhances group performance. *Organizational Psychology Review*, *1*(1), 32‐52.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328‐331.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and brain sciences*, *33*(2-3), 61-83.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., & Henrich, N. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767‐1770.

Huppert, E., Cowell, J. M., Cheng, Y., Contreras-Ibáñez, C., Gomez-Sicard, N., Gonzalez-Gadea, M. L., Huepe, D., Ibanez, A., Lee, K., & Mahasneh, R. (2019). The development of children's preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental science*, *22*(2), e12729.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*(1), 1‐27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory : Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881‐919.

Kajanus, A., Afshordi, N., & Warneken, F. (2020). Children's understanding of dominance and prestige in China and the UK. *Evolution and Human Behavior*, *41*(1), 23‐34.

Kiatpongsan, S., & Norton, M. I. (2014). How much (more) should CEOs make? A universal desire for more equal pay. *Perspectives on Psychological Science*, *9*(6), 587‐593.

Kimbrough, E. O., Sheremeta, R. M., & Shields, T. W. (2014). When parity promotes peace : Resolving conflict between asymmetric agents. *Journal of Economic Behavior & Organization*, *99*, 96-108.

Kim, J., Allison, S. T., Eylon, D., Goethals, G. R., Markus, M. J., Hindle, S. M., & McGuire, H. A. (2008). Rooting for (and then abandoning) the underdog. *Journal of Applied Social Psychology*, *38*(10), 2550‐2573.

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*(3), 426‑441. https://doi.org/10.1037/0022-3514.57.3.426

Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. *PLOS ONE*, *13*(2), e0192848. https://doi.org/10.1371/journal.pone.0192848

Magee, J. C., & Galinsky, A. D. (2008). 8 social hierarchy : The self-reinforcing nature of power and status. *Academy of Management annals*, *2*(1), 351‑398.

Maner, J. K. (2017). Dominance and prestige : A tale of two hierarchies. *Current Directions in Psychological Science*, *26*(6), 526‑531.

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(18), 6862‑6867. https://doi.org/10.1073/pnas.1113194109

Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. *Infancy*, *21*(3), 353‑372.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annu. Rev. Econ.*, *3*(1), 601‑630.

Norton, M. I., Neal, D. T., Govan, C. L., Ariely, D., & Holland, E. (2014). The not-so-common-wealth of Australia : Evidence for a cross-cultural desire for a more equal distribution of wealth. *Analyses of Social Issues and Public Policy*.

Palan, S., & Schitter, C. (2018). Prolific. Ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22‑27.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation : A personality variable predicting social and political attitudes. *Journal of personality and social psychology*, *67*(4), 741‑763.

Pun, A., Birch, S. A., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences*, *113*(9), 2376‐2381.

R. Core Team. (2018). Package "Stats." *The R Stats Package, https://www.R-project.org.*

Schäfer, M., Haun, D. B., & Tomasello, M. (2015). Fair is not fair everywhere. *Psychological science*, *26*(8), 1252-1260.

Schmid Mast, M. (2004). Men are hierarchical, women are egalitarian: An implicit gender stereotype. *Swiss Journal of Psychology/Schweizerische Zeitschrift Für Psychologie/Revue Suisse de Psychologie*, *63*(2), 107.

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363.

Shaw, A., & Olson, K. (2014). Fairness as partiality aversion : The development of procedural justice. *Journal of experimental child psychology*, *119*, 40‐53.

Sidanius, J., Pratto, F., & Bobo, L. (1994). Social dominance orientation and the political psychology of gender : A case of invariance? *Journal of personality and social psychology*, *67*(6), 998.

Sidanius, J., Sinclair, S., & Pratto, F. (2006). Social Dominance Orientation, Gender, and Increasing Educational Exposure 1. *Journal of Applied Social Psychology*, *36*(7), 1640‐1653.

Smith, R. H., Turner, T. J., Garonzik, R., Leach, C. W., Urch-Druskat, V., & Weston, C. M. (1996). Envy and schadenfreude. *Personality and Social Psychology Bulletin*, *22*(2), 158‐168.

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(4), 1-7.

Thomas, A. J., Thomsen, L., Lukowski, A. F., Abramyan, M., & Sarnecka, B. W. (2018). Toddlers prefer those who win but not when they win by force. *Nature Human Behaviour*, *2*(9), 662‐669.

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M. C., & Carey, S. (2011). Big and mighty : Preverbal infants mentally represent social dominance. *Science*, *331*(6016), 477‐480.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual review of psychology*, *64*(1), 231-255.

Ulber, J., Hamann, K., & Tomasello, M. (2017). Young children, but not chimpanzees, are averse to disadvantageous and advantageous inequities. *Journal of experimental child psychology*, *155*, 48‐66.

Van Berkel, L., Crandall, C. S., Eidelman, S., & Blanchar, J. C. (2015). Hierarchy, dominance, and deliberation : Egalitarian values require mental effort. *Personality and Social Psychology Bulletin*, *41*(9), 1207‐1222.

Van Rossum, G., & Drake, F. L. (2009). *Manual-Python 3*. CreateSpace.

Vandello, J. A., Goldschmied, N. P., & Richards, D. A. (2007). The appeal of the underdog. *Personality and Social Psychology Bulletin*, *33*(12), 1603‐1616.

von Rueden, C. (2020). Making and unmaking egalitarianism in small-scale human societies. *Current opinion in psychology*, *33*, 167‐171.

Xie, W., Ho, B., Meier, S., & Zhou, X. (2017). Rank reversal aversion inhibits redistribution across societies. *Nature Human Behaviour*, *1*(8), 1‐5.

Zitek, E. M., & Tiedens, L. Z. (2012). The fluency of social hierarchy : The ease with which hierarchical relationships are seen, remembered, learned, and liked. *Journal of personality and social psychology*, *102*(1), 98‐115.

Rank reversal aversion and fairness in hierarchies

Dear Dr Khandis Blake,

Thank you for offering us the opportunity to revise and resubmit our manuscript entitled "Rank reversal aversion and fairness in hierarchies" (AHBP-D-22-00022) for possible publication in *Adaptive Human Behavior and Physiology*. We are very grateful for all the work you have done on our manuscript. The suggestions for deletions, reorganizations, and rewordings that you made were very helpful to us. We are well aware that this must have taken a lot of your time and we sincerely thank you for it. We hope that we have improved the manuscript and that it will meet your expectations.

Yours sincerely,

Alexandre Foncelle, Elodie Barat, Jean-Claude Dreher & Jean-Baptiste Van der Henst

Thank you for your revised submission of "Rank reversal aversion and fairness in hierarchies", AHBP-D-22-00022R1. I appreciate the time and attention paid to the comments of the two reviewers and myself.

Below I document areas for further modification, before the manuscript can be considered by AHBP for publication. Most of the improvements needed are stylistic - the manuscript is still rather difficult to follow in sections, especially the presentation of the results. Throughout each section, a better use of sub-headings would help focus the manuscript and its improve its concision.

Throughout the manuscript, it may be easier to follow if you explain that the scenarios are in fact vignettes, and refer to them as such. In the Methods section of the Abstract, I also suggest you explain the method as follows: "Participants read 12 vignettes depicting six hypothetical hierarchies that varied in fairness. In each vignette, one individual was endowed with more resources than another individual, and participants could reduce that inequality by transferring resources from the higher-ranked individual to the lower-ranked one. In half of the vignettes, reducing the inequality led to a reversal of rank, while in the other half it did not."

Response:

Throughout the manuscript, the scenarios are now referred to as vignettes. Moreover, in the Abstract, we describe the method as you suggested.

INTRODUCTION
p3L19 - replace 'in' with 'of'
p3L38 - replace 'the Social' with 'Social'
p3L53 - replace 'A social' with 'Social'
p4L2 - delete ideologies
p4L10 - replace 'in just world' with 'in a just world'
p4L17 - replace 'large' with 'much'
p4L24 - 'greater equality pay' --> 'greater pay equality'
p4L39 - 'different motive such as' --> 'different mechanisms including'
P5L54 - last two sentences should follow on from previous paragraph (ie remove line break)
P6L5 - Start paragraph 'In Xie...'

Response:
We have followed all these suggestions.

P6L20 - 'the transfer also resulted in hierarchy reversal'. Here, you could end this sentence as follows: 'the transfer also resulted in hierarchy reversal (i.e., the richer player became the poorer player, and vice versa).' and delete the next two sentences. Likewise, you could then delete the paratheses on L38, (Richer: 4;...).

Response:
We have followed all this suggestion, but we have deleted only the first of the two sentences you suggested in order to make the example easier to understand.

p8L34-44 - I suggest a re-write as follows: "To provide a better understanding of people's willingness to reverse rank, we presented participants with 12 vignettes depicting six hypothetical hierarchies that varied in fairness. All hierarchies were inequitable, but like Xie et al., participants could redistribute resources to reduce inequality. In half of the vignettes, redistribution resulted in rank reversal, such that the higher-ranked individual became the lower-ranked one (and vice versa). We tested the prediction that participants would be more likely to reverse ranks between the two individuals when the hierarchy was unfairly established." Following - The rest of the paragraph - if all of these sentences are necessary, please consider re-writing for brevity and clarity. I'm not sure they are necessary, but if they are, I'm also not sure exactly what you are trying to say with them.

Response:
We have adopted this rewriting but we have replaced “All hierarchies were inequitable” with “All hierarchies involved an inequality of resources”.

In addition, we have deleted the rest of the paragraph that you found unnecessary. We have only expanded the last sentence of the paragraph as follows:
“We tested the prediction that participants would be more likely to reverse ranks between the two individuals when the hierarchy was established unfairly and we examined whether the unfairness of hierarchies neutralized rank reversal aversion.”

METHOD
P9L14 - should be 'Prolific Academic'
P9L19 - replace 'interested' with 'eligible'

Response:
We have made the changes.

Currently, you describe part of the stimuli, then refer readers to the Fairness Measure scenario section. Instead, you should discuss these scenarios and explain the fairness manipulation at the same time. Hence:
I suggest you move the 'Fairness measure scenario' section to the start of Materials and Design, under a sub-heading called "Fairness manipulation". Begin with sentences P9L44-13, then continue with paragraphs in Fairness Measure Scenario section starting 'To obtain a measure of fairness...'. The next two paragraphs (P10L14 to P11L5 - I'm not convinced these are necessary. They explain your predictions for the fairness measures, but ultimately this isn't really relevant to the paper, its only relevant to the fairness manipulation. If that was its own study, then it might be worthwhile spending two paragraphs explaining your predictions. But given it is just a pilot study that allowed you to determine fairness scores for your manipulation, two paragraphs isn't justifiable in this case, and your predictions for the manipulation in my mind isn't relevant. Hence I'd suggest cutting these two paragraphs.

After your 'Fairness manipulation' section in the method, it would be helpful to then have a 'Rank reversal manipulation' section, which starts from P11L7. Afterward, you should include a heading for 'Gender manipulation', where you explain the gender manipulation. Then, a sub-heading for 'Relevant covariates', where you list and explain these.

Response:
We have followed your suggestion. We now present the vignettes and explain the fairness measures in the same subsection entitled "Fairness manipulation". In addition, we have also introduced a section called "Rank reversal manipulation" and a section called "Relevant covariates".

In the data analysis and beyond, it would improve the readability of the manuscript if you refer to the scenarios for what they actually measure, i.e., hierarchy fairness, rather than 'scenario'. Likewise, rather than referring to 'condition' - when you actually have three different manipulations - you would be best to specify 'rank reversal condition'.

Response:
We have taken your advice and are now referring to the more meaningful terms you suggested.

RESULTS
P15L46-56 would be best included in the method where you explain the rank reversal condition. There you can explain that while you included P-1 and P-2, there was no difference between them, hence you pooled them for further analysis.

Response:
We have now moved this paragraph in the Material and Design section of the Method part (page 13):
"Participants dealt with the six vignettes in Reversal situations as well as in Preservation situations and thus received 12 trials. About half of the participants received Reversal and Preservation-£1 situations and half received Reversal and Preservation-£2 situations. Hence, in the first variant of the binomial mixed-effects regression model of our analysis, the rank reversal condition variable had three levels (Reversal, Preservation-£1, and Preservation-£2). Pairwise comparisons in the post-hoc analysis showed that, for each hierarchy fairness vignette, there was no significant difference between Preservation-£1 and Preservation-£2. Therefore, we pooled the two preservation situations for further analysis (Table ESM1)."

Further, the results in general are still very confusing and difficult to follow. I take no issue with your actual approach - but others will if they dont understand what you've done, and its important that the manuscript is presented in the best possible way for people to grasp your approach. It appears that what you've done is essentially follow a model fitting approach, whereby you start with a baseline model, then you

add additional terms to the model to see whether they are significant/improve the model fit. If they do, you retain them, if they don't, you don't. This is a great approach, but the way the manuscript is currently framed, it isn't clear that you've done this. Hence, in the data analysis section, I need you to explain that you took a model fitting approach, and ran F-tests/chi-sqr tests to determine the best fitting models. Name the Models Model 1-n. Put the results of the ANOVAs comparing each model all in one table, which shows the logical progression of model fitting tests (i.e., Model 1 v Model 2, Model 2 v Model 3), and which variables were important, and which weren't. This table could go in the Online Supplementary Materials, or the main text if you think it necessary. A good example of how to do this for AHBP is the following paper, which I advise you consult: https://link.springer.com/article/10.1007/s40750-020-00151-3#Sec7.

Response:

We have followed your suggestion. The Data analysis section now describes our model fitting approach as follows (page 14):

”First, we examined how rank reversal aversion was related to fairness. To do so, we reported participants' decision to reject the transfer in each trial. To analyze these data, we adopted a model-fitting approach on a binomial mixed-effects regression model. We started with a simple model (Tables ESM3 and ESM5), including the estimated fairness of the hierarchy vignette and the rank reversal condition as main fixed effects, and participants as a random effect (to control for repeated measures), and then added terms to the model to see if they were significant or improved model fit (e.g. the order of presentation of vignettes, participant gender, and picture gender). We performed Chi-squared tests to determine the best-fitting model (Table ESM6). Next, to characterize the effect size, we performed pairwise comparisons by running F-tests in post-hoc analysis. We reported the odds ratio of rejecting the transfer by taking the exponent of the estimate.

In the results section, we report the final reduced model of our fitting approach, its main effects and the interactions between these effects. The extraction of data and plots were performed in Python3 (Van Rossum & Drake, 2009) and statistical analysis were conducted withR (version 4.2.1) using the lme4, multcomp and emmeans packages (R. Core Team, 2018).
”

Inspired by Luberti et al. (2020), we also created three tables to present the results in a more synthetic way:

- **Table ESM3.** Equations for GLMER models.
- **Table ESM4.** Parameters for CLMM and GLMER models.
- **Table ESM6.** ANOVA comparisons for the model fitting approach.

For the actual significant result odd ratio results - a good rule of thumb is that anything that is significant from the model fitting procedure should be tabulated in the paper, whereas tables depicting non-significant results can go in the

supplementary materials (not an appendix). If you need to explain something, add it into the table note, rather than explaining it as a dialogue in an Appendix. Also note that if the approach is too unclear without you adding a dialogue to each table (as you currently do in the Appendix), this is a good indication that the paper is not yet written to a high enough standard for publication.

Response:

Following the example of Luberti 2020, we have put the tables for the fitting approach and for the post-hoc comparisons in Online Supplementary Materials and we refer to them in the Results section. We also removed the dialogue style and privileged notes under our tables.

The new results section reads as follows:

"To interpret the results, we evaluated the best model of our model-fitting approach (Tables ESM7 - ESM11) whose equation is as follows:

*Transfer ~ HierarchyFairness*RankReversalCondition*PictureGender +HierarchyFairness*Gender +RankReversalCondition*Gender +HierarchyFairness*HierarchyFairnessOrder +HierarchyFairness*SES +HierarchyFairness*Politics +(1|Participant)*

where *Transfer* is a binary factor variable, *HierarchyFairness* is an ordered factor based on fairness measure ("Theft" < "Despotic Power" < "Pre-existing hierarchy" < "Chance" < "Competition" < "Merit"), *RankReversalCondition* is a binary factor variable, *PictureGender* is a binary factor variable, *Gender* is a binary factor variable *HierarchyFairnessOrder* is a four-level factor variable, *SES* is an ordered factor, *Politics* is an ordered factor and *(1|Participant)* is the random intercept of the participant (Table ESM4).

The analysis of variance (Table ESM7) revealed a significant main effect of *HierarchyFairness* (χ^2 (5) = 686.9, p < .001). Given that our main hypothesis was that participants would be more likely to reverse the hierarchy when it was fair rather than unfair, pairwise comparisons of transfer rejection rates between vignettes are reported only for the Reversal condition (Table 2). In particular, to examine the reluctance to reverse ranks across vignettes, a post-hoc analysis estimated how many times participants were more likely to reject the transfer between two hierarchy fairness vignettes in the Reversal condition (Table 2). This analysis confirmed the prediction that when a hierarchy was unfair, participants were more likely to reverse the hierarchy than when it was fair (Table 2). Merit was perceived as the fairest hierarchy and it was also the one for which participants were most reluctant to accept the transfer. In contrast, Despotic Power and Theft were perceived as the most unfair hierarchies and were also the ones for which participants were most willing to accept the transfer. It should also be noted that for the Chance vignette of the Reversal condition, 53.4% of participants rejected the transfer (Figure 2), a rate that was highly similar to the one reported by Xie et al. (i.e. 55.2%).

The analysis of the model also revealed a significant main effect of *RankReversalCondition* (χ^2 (1) = 78.6, p < .001) indicating that participants were more likely to reject the transfer in the Reversal condition than in the Preservation condition. Indeed, participants were 2 times more likely to reject the transfer (2.0 [1.6, 2.6], p < .001) in the former compared to latter condition. In addition, there was a significant interaction between the conditions (Reversal vs. Preservation) and the hierarchy fairness vignettes (χ^2 (5) = 14.0,

$p < .02$). A post-hoc analysis estimated the odds of rejecting the transfer between the reversal and preservation conditions of each hierarchy fairness vignette (Figure 2 and Table 3). For Merit, Competition, Chance, and Pre-existing Hierarchy, the odds of rejecting the transfer in the Reversal condition were highly multiplied relative to the odds of rejecting the transfer in the Preservation condition. Moreover, this multiplier effect was at its maximum for Competition, and decreased at the extremes of the fairness scale, namely for Merit and Theft.

The interactions between *HierarchyFairness* and *PictureGender, SES, Politics* and *HierarchyFairnessOrder* were globally significant (Table ESM7) and even if the study was not designed to explore all pairwise comparisons of this model, we could investigate some pairwise comparisons for SES and Politics (Table ESM10 - ESM11). The triple interaction between *HierarchyFairness, RankReversalCondition* and *PictureGender* was also significant (χ^2 (5) = 13.0, $p < .02$). Studying some pairwise comparisons (Table ESM9), participants are significantly more likely to reject transfer in the Reversal condition for both Competition and Chance vignettes, compared to the Preservation condition, for both males and females in the pictures, although for the male pictures the odds are higher. Furthermore, *Gender* was globally non-significant (χ^2 (1) = 2.9, $p = .090$)) and other non-significant variables such as *PictureGender*, *SES*, *Politics* and *HierarchyFairnessOrder* are reported in Table ESM7. Random intercepts for participants improved the model compared to a version without it, according to a Chi-square test (χ^2 (1) = 259.1, $p < .001$)".

In the tables - Please change 'f =' to 'fairness =". It will simply be too misleading if you leave it as 'f'.

Response:
We have made these changes.

The tables should also have clear headings, with any table notes written below the table as a 'Note'. Please see APA formatting for guidelines.

Response:
We have followed APA formatting for tables and figures.

There is a significant scenario:imageGender effect that is not discussed. Perhaps I mis-read. But if a finding such as this is significant, it should be discussed in the results.

Response :

Indeed, there is a global significant effect in the ANOVA for *HierarchyFairness*:PictureGender, however our study was not designed to explore these specific interactions with proper pairwise comparisons. Hence, we specified this limit in the last paragraph of the Results section, and we say a word on this interaction, also nested in a triple one (*HierarchyFairness, RankReversalCondition* and *PictureGender) :*

"The interactions between *HierarchyFairness* and *PictureGender, SES, Politics* and *HierarchyFairnessOrder* were globally significant (Table ESM7) and even if the study was not designed to explore all pairwise comparisons of this model, we could investigate some pairwise comparisons for SES and Politics (Table ESM10 - ESM11). The triple interaction between *HierarchyFairness, RankReversalCondition* and *PictureGender* was also significant (χ^2 (5) = 13.0, p < .02). Studying some pairwise comparisons (Table ESM9), participants are significantly more likely to reject transfer in the Reversal condition for both Competition and Chance vignettes, compared to the Preservation condition, for both males and females in the pictures, although for the male pictures the odds are higher."

DISCUSSION
P24L2 - 'does not completely unlock' --> 'does not guarantee'
P26L41 - 'digging into this question' --> 'Answering this question'

Response:
We have made these changes.

The Discussion needs headings - something like implications, future directions, limitations, conclusion. This will help you ensure that you are remaining on point and not meandering too much too and fro, and will help the reader follow the trajectory of the paper.

Response:
We have now added headings to the discussion (i.e., "Limitations and Future Directions"; "Conclusion"). In addition, after re-reading the manuscript with your concerns in mind, we realized that our statements were sometimes too convoluted, lacked conciseness, and were not always well organized. In addition the discussion was too long for a manuscript describing a single experiment. We have now shortened the discussion by 617 words (from 2676 to 2059 words) by improving conciseness and eliminating unnecessary details rather than deleting paragraphs. We have also put all the points about future directions in the same section, whereas in the previous version they appeared in places too far apart in the Discussion.

OVERALL
Please go through the whole manuscript for repetition, inconsistency of expression (calling something Imagegender and then Photogender, etc), and brevity. This is a rather long manuscript and often longer paragraphs could be shortened and made more succinct. Also ensure that hypotheses are consistent between each section in terms of wording and implications - sometimes it seems as if one person wrote the introduction but

another wrote the discussion, and they conceive of the hypotheses in slightly different ways.

Response:

We have paid more attention to the consistency of the terms and we have systematized the use of expressions that you suggested to us. We have also reduced the manuscript by almost 1000 words (from 7512 to 6599 words). In this respect, the suggestions of cuts that you made to us were very useful. In addition, the simplification of the presentation of the hypotheses allowed us to better harmonize the discussion and the introduction.

Finally, note that we have changed the order of the first two authors, which is now more faithful to their contribution in the manuscript.

Multidisciplinary Digital Publishing Institute

Journal — Change of Authorship Form

(Must be Completed and Signed by ALL Authors)

Manuscript ID: AHBP-D-22-00022R1
Manuscript Title: Rank reversal aversion and fairness in hierarchies

Description of the change: The order of the first two authors have been reverted

Original Authorship

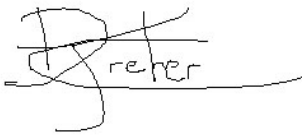
LIST ALL AUTHORS in the same order as the original (first) submission. For more than 10 use an extra sheet.

	Name	Affiliation
author (1)	Elodie Barat	Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028 École de Psychologues Praticiens – Institut Catholique de Paris
author (2)	Alexandre Foncelle	Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028
author (3)	Jean-Claude Dreher	Institut des Sciences Cognitives Marc Jeannerod – CNRS UMR 5229
author (4)	Jean-Baptiste Van der Henst	Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028

New Authorship

All authors must sign below agreeing to the new changes in authorship. The authorship order must match the new title page of the manuscript. Signatures below certify compliance with the author responsibilities on the next page. List ALL AUTHORS in the same order as the new version.

	Name	Email	Affiliation	Signature&Date
author (1)	Alexandre Foncelle	alexandre.foncelle@gmail.com	Centre de Recherche en Neurosciences	12/09/2022 

			de Lyon - CNRS UMR5292, Inserm U1028	
author (2)	Elodie Barat	ebarat@psycho-prat.fr	Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028 École de Psychologues Praticiens – Institut Catholique de Paris	12/09/2022 
author (3)	Jean-Claude Dreher	dreherjeanclaude@gmail.com	Institut de sciences cognitives, CNRS	12/09/2022 
author (4)	Jean-Baptiste Van der Henst	jean-baptiste.vanderhenst@cnrs.fr	Centre de Recherche en Neurosciences de Lyon - CNRS UMR5292, Inserm U1028	12/09/2022