THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse - Jean Jaurès

Présentée et soutenue par :
Andrei Ivănescu

le lundi 16 octobre 2017

Titre :

Third-party expectations of nepotism and mating preferences from facial similarity

---

École doctorale et discipline ou spécialité :

ED CLESCO : Psychologie

Unité de recherche :

Cognition, Langues, Langage, Ergonomie (CLLE), UMR 5263

Directeur/trice(s) de Thèse :

Jean-François Bonnefon, Astrid Hopfensitz

Jury :

Jean-François Bonnefon, directeur de recherche, CNRS
Coralie Chevallier, chargée de recherche, ENS
Sigrid Suetens, professeur associé, Université de Tilburg
Hugo Mercier, chargé de recherche, CNRS

THIRD-PARTY EXPECTATIONS OF NEPOTISM AND MATING PREFERENCES FROM FACIAL SIMILARITY

ANDREI IVĂNESCU

CLLE
l'ED CLESCO
Université de Toulouse

2017

Do not go gentle into that good night,
Old age should burn and rave at close of day;
Rage, rage against the dying of the light.

— Dylan Thomas (1952)

Dedicated to the loving memory of Ion Szabo.

1939 – 2009

ABSTRACT

Our relation to our kin shapes much of our social world. It's no surprise then, that how we recognize and react to our own kin has been a widely investigated topic. In particular, when tackling direct kin recognition, facial similarity has emerged as a putative cue of relatedness. In this thesis, I investigate whether or not the same can be said for third party kin recognition. Split between two lines of research, we explore individuals' predictions of nepotistic and mating behavior in third party scenarios using facial stimuli. These two domains provide the backbone of our research. *Categorization must serve action*. So, what would strengthen the notion of a presence of third-party kin recognition in humans? Facial similarity *must have* a context-dependent effect on participants predictions, susceptible to valence changes in scenarios and switches from the prosocial and mate choice domains. This is precisely what we set out to do with our two lines of research. Though our literature review revealed that when context is starved participants seem to be able to detect similarity and seemingly connect it to relatedness. Our nepotism and mating series of experiments, by re-inserting context, offers us a different conclusion altogether. Within scenarios in which valence is modified and our participants analysis is bounded by predictions made by kin selection, their choices do no reflect a connection between similarity and relatedness.

Keywords: *kin recogntion, third parties, inclusive fitness, phenotype matching, facial similarity.*

RÉSUMÉ

Notre relation avec nos apparentés forme une grande partie de notre monde social; et la façon dont nous reconnaissons et traitons nos apparentés a donné lieu à une importante somme de recherche. Lorsqu'il s'agit de reconnaître un apparenté direct, la similarité faciale est considérée comme un indice d'apparentement. Dans cette thèse, j'étudie si elle joue un rôle comparable lorsqu'il s'agit de reconnaître un apparentement entre des tiers, en menant deux lignes de recherche: les prédictions de comportement népotistiques et les prédictions de préférences de couple, par des tiers, en présence de stimuli faciaux. *La catégorisation devant servir l'action*, la similarité faciale doit avoir un effet dépendant du contexte sur ces prédictions, susceptible à des changements de valence et de domaine. En l'absence de contexte, les individus semblent pouvoir détecter la similarité faciale et la mettre en relation avec l'apparentement. Nos deux séries d'expériences offrent une confusion différente. Quand la valence du contexte change et que nous analysons les prédictions des participants en terme de kin selection, leurs choix ne semblent pas mettre en relation similarité faciale et apparentement.

Mots-Clés: *reconnaissance des apparentés, tierces parties, fitness inclusive, appariement de phénotype, similarité faciale*

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor Jean-François Bonnefon. Your guidance and patience has made all of this possible. I couldn't have asked for a better teacher and mentor. From the bottom of my heart, thank you!

I am particularly grateful to my co-supervisors, Astrid Hopfensitz and Gwenaël Kaminski, for their continued support and helpful feedback.

I am indebted to the wonderful members of our lab group André Tricot, Florence Lespiau and Marina Miranda Lery Santos . Our meetings helped shape and grow this research!

In particular, I would like to express my sincere thanks to Florence Lespiau. You are a brilliant colleague and friend. As I have said many times, a true life saver. Thank you for always being there to help!

I would also like to thank the following institutions for their assistance and support:

CLESCO, CLLE and TSE

Finally, I would like to offer my very special thanks to my family and friends. I might have become a hermit, but I was never alone.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Corb la corb nu scoate ochii.

— Romanian Proverb

We search for meaning and we search for patterns. We read a strange proverb, in a strange language, and we wonder. *What does it mean? Why is it here? How will this serve the purpose of this thesis?* We are dependent on context and without it these question remained unanswered and the proverb remains meaningless.

These questions echo a greater search for an understanding of our own nature. In that, our evolutionary history offers us a wealth of topics to explore and ask these fundamental questions. Our seemingly unique status in the animal kingdom only further entices us. It should come as no surprise to the reader that this very thesis is a symptom of our extraordinary human ailment. Our need to understand.

One topic that has been a fixture of our search for understanding is our uniquely cooperative nature. A great deal of research has been dedicated to elucidating the evolutionary reasons behind our unique propensity to cooperate. This feature of the human species has allowed us to rise from humble beginnings and reach a point where humanity has been able to break the bonds of earth and touch the heavens. None of the great works of engineering or even art could have been achieved if we lacked this one simple ability: to restrain our more aggressive instincts in favor of joining hands and cooperating towards a common good. Essential to our understanding of cooperation is kinship. Which brings us back to our strange proverb.

One might still be wondering, what does the proverb mean, what is the context in which it is relevant? Ad literam, it refers to a raven not pecking out another raven's eyes, akin to birds of a feather flock together, only in a more gruesome package. Given context both this proverb and its countless variations found around the world offer the same insight. People will not harm and will protect their own. More often than not, in our history, "our own" has meant our close kin and our more distant relatives. Hamilton (1964) realized that these proverbs hold a fundamental truth at heart. Through his theory of kin selection we find that these cooperative and altruistic acts towards kin in fact provide an evolutionary advantage. An advantage that in more ways than one, provides the springboard for what makes us truly unique in the animal kingdom, our cooperative nature. Thus, a quintessential question in this field has been, *how do we spot our own? How do we recognize kin?*

To answer this question, the first part of this thesis will be dedicated to an exploration of the literature connected to kin recognition. Our initial foray will begin with human cooperation and cooperation within the animal kingdom in chapter 2, while slowly progressing towards a key issue, altruism. Kin selection provides the much needed lens through which cooperation and more importantly altruism can be examined and understood. We will present our in-depth exploration as to the implications of kin selection in section 2.1 before presenting, in chapter 3 the mechanism that is fundamentally required for it to function, kin recognition. Finally, our expedition will bring us to a major focus of current research on human kin recognition. Specifically direct kin recognition.

The first part of chapter 4 will be dedicated to the research that has been done on the topic of how we recognize our own kin and how this information affects behavior (i.e., direct kin recognition). As one might expect this line of research and the use of manipulated facial stimuli has proven to be extremely important in expanding our understanding of kin related behavior and the decision making processes involved.

The same depth of research has not, unfortunately, been dedicated to how and if we recognize and are affected by perceptions of relatedness in third-parties. This is an issue we hope our work will partially remedy. We will begin by discussing the state of the field in regards to third party kin recognition in section 4.1, delving into the issues and pitfalls this research entails and what we can do to improve our understanding of the underlying processes involved. As our research involves utilizing some of the same methodological and experimental procedures of manipulating facial similarity used in experiments investigating direct kin recognition, section 4.2 will be wholly dedicated to understanding how these procedure work and how they can expanded to third party scenarios.

The final chapter in this venture in the literature of kin recognition will serve one important role. To drive home one overarching theme of this thesis. Context matters. Investigations in third-party kin recognitions must a make this theme a central component of their experimental environment for any progress to be made in discerning the factors at play in directing predictions of the others. When context is starved and valence is deprived any results will be hard pressed to reveal the nature of the mechanism to be investigated. If third party kin recognition is present we must be able to investigate its effects. Categorization must be in service of actions, if participants categorize individuals as kin their predictions of these individuals behavior must follow discernible patterns of kin mediated interactions.

To achieve our goals, the experimental breath of this research has been two fold which allows us to precisely discern the nature of third party recognition effects. The first line of research has been directed towards investigating the relationship between third party kin recognition and third party expectations of nepotistic behaviour (i.e., Chapter 6). While the second involves third party kin recognition and its effects on decision making related to mating behavior(i.e., Chapter 7).

Concerning our first line of research, our premise following the context-dependent nature of kin recognition mediate behavior is that generally people cooperate more with people who look like them. This has been interpreted as an effect of kin recognition coupled with kin cooperation. Additionally, people also are thought to expect agents that look like each other to cooperate more and this, again, can be interpreted as an effect of third party kin recognition coupled with the same expectation that kin cooperate. Here we investigate whether people expect agents that look like each other to provide benefits to each other in a work context, in which this preference would amount to nepotism. Will people expect bosses to promote employees that look like them, or conversely fire employees who do not look like them?

Our second line of research follows the general principles set in our initial experiments. Using the same procedure to create our stimuli we investigate the relationship between third party cues of kinship and the predictions our participants would make based on them. As opposed to our first line of investigation, participants are presented with a mate choice scenario in which they are tasked to predict the choices of an agent in search of new romantic partner. Will people expect others to date potential mates that do not look like them, or conversely reject mate that look like them?

This juxtaposition of our two lines of research allows us to fully investigate if participants predictions are in fact connected to a third-party kin recognition mechanism or the result of other factors.

Part I

LITERATURE REVIEW

2 A NEED FOR COOPERATION

Kaplan et al. (2009), argued that in the case of humans, the adaptation towards the niche of skill-based hunting entailed increasing the time necessary to obtain nutritional independence, along with a net period of negative production in the learning phase compensated by a higher productivity period during adulthood. The increased ontogeny allowed the necessary time required to obtain the skill level necessary to exploit high value patches of resources. This change further reinforced cooperation through the need of parental investment by both the mother and the father. The theory provided by Kaplan et al. (2009), provides a strong basis for human social organization and the development of high level cognitive skills. What is lacked, however, is an explanation of how humans reached a position where high quality patches were accessible and exploitable. This theory presupposes the existence of a predisposition for cooperation that was pre-existing and the transition towards a skill-based niche further reinforced and stabilized it.

As we can see, what we are faced with, is that, fundamentally, human cooperation is based upon a "stag hunt" scenario, where mutualism is the most important aspect. Cooperation in the vein of mutual benefit provides the required framework in which a fully "human" prosociality can develop. As far as the steps required for this development to occur, Tomasello (2009), offers an important set of requirements for this propensity for cooperation to develop:

1. Evolve a shared intentionality with "serious social-cognitive skills and motivations for coordinating and communicating with others in complex ways involving joint goals and coordinated division of labor among the various roles".

2. For collaboration to even start, "early humans had first to become more tolerant and trusting of one another than are modern apes, perhaps especially in the context of food".

3. Tolerant humans "had to develop some group-level, institutional practices involving public social norms and the assignment of deontic status to institutional roles".

While these steps proposed by Tomasello provide a compelling timeline of our evolution, the social-cognitive skills needed for serious coordination and communication with conspecifics, fundamentally require a more tolerant and trusting "animal". Thus, the first step proposed by Tomasello is a development that first requires a fundamental shift in the aggression levels of early humans. For example, chimpanzees are able and do altruistically help a conspecific with novel tasks that present no reward (Warneken et al., 2007). They do understand when a conspecific requires help. What they do not understand is when a conspecific is trying to help them in the absence of a request. "Even when chimpanzees communicate with a human (e.g., by pointing), they are virtually always attempting to get him to do something for them", (Warneken et al., 2009). Theoretically, chimpanzees have most of the prerequisites to solve this task and in some cases their abilities far surpass these requirements. However, they are missing two crucial parts of this jigsaw puzzle: decreased aggression and tolerance towards conspecifics. These two factors impact on the propensity to cooperate. Without a predisposition towards cooperation any kind of unsolicited help would not be understood.

As with everything in nature, attending to the social cues of conspecific and sharing information is a cost benefit analysis. The benefits from our standpoint (i.e., humans) are immense. However, the large majority of possible advantages these abilities present are not readily apparent. The "fog" of strong competition for resources, abundantly present in the animal kingdom, shifts the balance in the detriment of using or obtaining such abilities. Hunting for example is most often shown as a sign of the high level of cooperation in chimpanzees. Teleki (1973), reported active cooperation as a major aspect of chimpanzees predatory behavior, with several instances of male chimpanzees pursuing infant baboons while other chimpanzees position themselves in "strategic" positions to cut off possible escapes routes. Active cooperation, however, is not the only way in which this strategy might emerge. As Busse (1978), points out, while it might seem that chimpanzees possess a great degree of coordination and cooperation, they might in fact be trying to maximize their own chances of obtaining the prey. This stems from the competitive nature that their environment imposes on them. Without some other factor modulating this competitiveness there are slim chances of evolution selecting traits in favor of true cooperation such as tolerance and decreased aggression towards conspecifics. Without these traits a slew of human behaviors would be impossible. A climate of constant competition presupposes a concrete lack of propensity to cooperate at the human level. Additionally, there is little incentive, to chimpanzees, to attend the social cues of conspecifics in non-competitive situations as with the example of helping behavior. This is mostly because the situations in which there is a lack of competition for food or other resources are few and far between. In addition, as seen in the case of sharing information, some types of abilities are simply not appropriate if the goal is individualistic, selfish competition for resources.

Thus, we see that chimpanzees like modern humans can both act in selfish ways to maximize their benefits and more importantly they seem to be able to altruistically help both unrelated conspecifics and humans. The issue of attending the social cues of conspecifics and decreased overall aggression remains, however we can plainly see that the scaffolding for human-like cooperative behavior is present in some of our closest relatives in the animal kingdom.

Selfish behaviour has been easily explained through natural selection, conversely altruistic behaviour has been a long subject of debate and research starting from Darwin himself, when he realized that this type of behaviour posed a serious problem and could not be accounted for in a straightforward analysis of natural selection, (Richards, 1987). His proposed solution relied on group level analysis in which one group would be more likely to survive as a function of the number of "well-endowed" individuals within it when compared to a rival group.

" A community including a number of well-endowed individuals increases in number and is victorious over others and less well-endowed communities; although each separate member may gain no advantage over the other members of the same community. " (Darwin, 1871).

While at the time this explanation would prove to be sufficient for Darwin, current understanding of the mechanisms behind natural selection, results in this explanation not aging well. The issue is made evident when analyzing costly behaviors for the individual, that greatly benefit his group. A classic example is signal calling, the caller at great personal cost, helps her group by announcing the presence of a predator, while this type of behaviour will greatly benefit her group the signal itself would put the caller in mortal danger, in the end making the likelihood of individuals surviving to pass on this trait greatly diminished. In this type of scenario, group level selection pressures would play no role in the evolution of altruistic behaviors. In practice the first true explanation that would resolve this long standing issue of altruistic behaviour would come from Hamilton in 1964. He recognized that the solution was not to be found at the group level but in a mechanism that would cause individuals to act selectively towards others with whom they would be likely to share the same genes, individuals would selectively interact in specific positive ways towards kin.

2.1 THE KEY IS KINSHIP

Hamilton realized that our ability to act, according to the degree of relatedness, is predicated upon our social nature as a species. Continuing the classic example of alarm callers, we see that the cost/benefit ratio is fundamentally changed when we take into account the group composition in terms of kinship. While at the proximal level the caller would remain in mortal danger by providing her group with the need time to escape from the signaled predator, the fact that her call would provide a fitness benefit for her kin, sharing the same genes, would ultimately facilitate the selection for this trait. This process can be expanded to any type of behavior that incurs a high cost on the actor while providing significant benefits towards a recipient that descents from the same common ancestor (i.e., kin).

The definition of altruism can seem at times extremely unreliable when going through these types of scenarios. This unreliability stems from common usage of the word altruism. In the common language, an altruistic act should not give any type of benefit towards the actor. When analyzing, however, altruistic acts through an evolutionary lens, it is very important to distinguish between the proximal and ultimate levels of the analysis. At the proximal level a caller announcing to her tribe the presence of a predator receives no immediate or in many case even long term benefit. By all intents and purposes the actor is altruistic in the dictionary definition of the term. However, when analyzed at the ultimate level, the indirect fitness benefits received are made evident through the continued survival of her kin that might in future pass on their genes which are shared with the actor as a result of their descent from a common ancestor.

Within this framework we can see that selection pressures can and do provide the necessary scaffolding for proximate altruistic behaviors to occur as long as certain factors are present. At the basic level this framework provides a straightforward and eloquent rule, Hamilton's rule. If the costs of performing an action(c), are less than indirect/direct fitness benefits (b), where these benefits are a function of the degree of relatedness(r) between actor and the recipient, then this behavior be it alarm calling, or any other type of behavior, it will be favored by natural selection. More succinctly put $rb-c>0$. Simple as it may be, the mere addition of the coefficient of relatedness(r) provides the needed insight to finally resolve this long standing issue of altruistic behavior, at least at the level of interactions between kin.

Predictions made on the basis of Hamilton's Rule offer us testable framework in which we can analyze what kind of altruistic behavior we should expect given the type of kin relationship present between actor and recipient. We would expect that altruism should be limited to kin and furthermore truly costly acts should be relegated to close kin. For example $r=0.5$, or more precisely, when we are dealing with interactions between parents and their offspring the amount of benefit need can be only twice the cost which allows altruistic behaviors by parents towards their children to be easily selected for. The interactions between kin play a fundamental role in shaping the social environment in many if not all of the animal taxa. Kin selection plays a significant role in understanding complex behavioral patterns that we would be otherwise hard pressed to explain. From the the very example of alarm calling, we discussed previously, which can be readily found in nature in species such as the ground squirrels, (Sherman, 1977), to allomaternal behavior in a number of species from elephants, (Lee, 1987), to vervet monkeys, (Fairbanks, 1990). Kinship is the first foundation in developing any kind of cooperative behavior be that mutualistic or altruistic.

Conversely, kin selection is not a context-independent process. While we would expect increased cooperation and altruism towards kin, interactions between kin also involve sexual attraction and more specifically the issue of inbreeding avoidance. Generally speaking inbreeding depression is the primary reason behind a need for bare minimum ability to recognize and avoid sexual relations with close kin. Inbreeding depression has been widely investigated and researched. Mate choice in this context is dependent on three types of benefits: direct benefits, genetic benefits and indirect fitness benefits. In terms of direct benefits issues such as opportunity costs and the ability of the mate to offer protection, shelter, resources to potential offspring have a large role in directing potential choices of mates. Genetic benefits are dictated by the fitness advantage gained from potential heritage of advantageous genes (Puurtinen et al., 2009). Indirect fitness benefits present the more interesting dilemma when compared to cooperative and altruistic behavior. Incestuous sexual relations at the same time raise the likelihood of inbreeding depression in potential offspring while also increasing the representation of the parents allele in the offspring. So, inbreeding avoidance should be function of the relation of inbreeding avoidance and increased relatedness where a maximal fitness benefit of inbreeding can be ascertained. While a large number of studies have investigated inbreeding avoidance and inbreeding depressions, fewer studies have focused on this nuanced issue of the maximal fitness benefit of inbreeding. One such study is that of Puurtinen (2011) modeling mate choice for optimal inbreeding. While the predictions one makes on the basis of specific context discussed might differ (i.e., altruistic and cooperative interactions vs. mate choice), the mechanism used to ascertain the viability and desirability of a potential partner does not face the same level of modification.

This brings up the important issue of the different interactions between individuals in social environments. While in principle, cooperation is forced upon our species as a result of the environmental pressures governing the development of characteristics beneficial to cooperation. Maintaining cooperation raised some of the most significant issues to our continued understanding of our cooperative nature.

While this will not be a focus of our research here, it is important to note that there is a great deal of debate and research still on-going as to how humans as a species have evolved to reach the level of cooperation we see in today's "modern" world. Kin selection is only one of many models and systems that allow us insight into this topic. Our social world is greater than just the sum of our kin. In 1964, Hamilton, cataloged the different types of interactions between "actors" and "recipients" in social situations. These classifications revolve, as discused above, around the analysis of measurements of the cost/benefit ratio of the interactions as it relates to the actor and the recipient. Interactions that are beneficial only for the actor are cataloged as being selfish, interactions that benefit only the recipient are defined as being altruistic and finally interactions that benefit neither the recipient nor the actor are spiteful. Furthermore, in light of the papers of a number of authors (Baumard et al., 2013; Clutton-Brock, 2002; Emlen, 1997; Gardner et al., 2004; Krebs et al., 1993; Ratnieks, 2006), we would define actions that provide benefits for both the actors and the recipients as being mutualistic.

This framework allows us to see that altruism via kin selection is not the only way in which this type of behavioral pattern can come to be in nature. Its evolution can also result from reciprocal altruism. Since the famous paper of Trivers (1971), the issue of cooperation and subsequently mutualism has been presented and analyzed through the lenses of partner control. Generally associated with the Iterated Prisoner Dilemma (IPD) type setting, the main question in this framework revolves around the issue of how an individual should treat his partner in a given series of encounters so as to maximize the payoff. In this scenario, what should an individual do? Is cooperation a profitable choice? Interestingly, the main finding of this approach is that, given certain circumstances, the cost paid by the individual towards his partner can be recouped if his partner, in turn, reciprocates. The most common example of a strategy that utilizes this insight is the Tit-for-Tat strategy in the IPD tournament presented in Axelrod (1984). Two conspecifics or non-conspecifics must have an extended series of interaction with each other varying their behavior as a response to the feedback received from their partner (i.e., tit-for-tat strategy).

Cooperation on the whole can be separated into a number of different models and systems that govern what type of behavior we would expect from participants in any cooperative endeavor. These systems are complementary and each of them offer, in turn, their own addition to create in part the great complexity we see in cooperative behavior in humans and in the animal kingdom as a whole. Sachs et al. (2004), separates the different models of cooperation into kin selection, direct reciprocation and byproduct. I will briefly discuss the last two models before returning to kin selection.

Sachs et al. (2004), underline two models that relate to direct reciprocation, Partner Fidelity Feedback (PFF) and Partner Choice. An important note here is that PFF can be seen as a modification of the traditional Partner Control model based on the IPD and the tit-for-tat strategy we discussed earlier. These two models differ in two basic ways; first of all, in PFF (see figure 1) fitness feedback is automatic, determining the decline of a cheater's fitness if he does not maintain the fitness of his partner (i.e., if they cheats); secondly, the maintenance of the cooperative interactions in PFF is not grounded in a conditional reciprocation strategy that bases ones behavior on the behavior of his partner (i.e., tit-for-tat). Nevertheless, it is important to keep in mind that these seemingly small changes do provide important benefits to the traditional IDP framework. An important typical trait of partner control models is the fact that individuals have no control over their partners. Forming the partnership is fixed and the individuals exert no ability to choose their partner. This typical characteristic of partner control models is curiously also the most unrealistic considering how group/partnership formation occurs in nature. Here individuals often have a high degree of choice on who they choose as their partner and when this partnership is terminated in favor of a more profitable one.

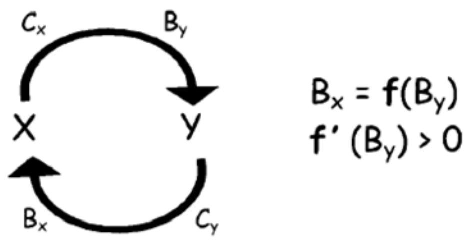

Figure 1: Partner Fidelity Feedback (Sachs et al., 2004)

"Benefits transferred from X to Y feed back through an extended series of exchanges."(Bx and By represent the benefits to X and Y, while Cx and Cy represent the costs for X and Y)."

This brings into focus the Partner Choice model of cooperation. Kaplan et al. (2005) comment on this issue: individuals in hunter-gatherer societies "vote with their feet". We have to take into consideration not just the individuals who want to be chosen as partners in cooperative enterprises but also the way individuals choose their partners. In the IDP framework the only choice is to either cooperate or not (i.e., tit-for-tat) forcing the "cheater" to either resume a productive relationship or degrade the interactions into a vicious circle diminishing the fitness benefit of both. Allowing the ability to choose the partner opens the possibility of dealing with the uncooperative partner by simply discontinuing the relationship and choosing a different partner. This focus leads to competition among members of a group enhancing a selection pressure on traits related to reliability as a partner, ability to recognize worthy partners and also traits related to their ability to complete and produce better results in cooperative tasks. Sachs et al. (2004) present, in their review of the different models of cooperation, a deceptively simple and elegant description: "An individual Y interacts with and rewards a specific cooperative partner X and avoids rewarding less cooperative partners. By choosing a cooperative partner X, individual Y not only enhances its own fitness but it promotes the evolution of cooperation in species X. This latter effect occurs because Y selectively benefits cooperative individuals of X through its cooperation." (see figure 2).

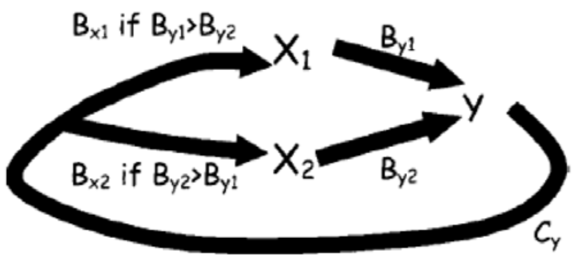

Figure 2: Partner Choice (Sachs et al., 2004)

"Either individual X1 or X2 receives a benefit from Y, depending on Y's choice. Y chooses to interact with the more cooperative X individual." (Bx1 and Bx2 refer to the benefits to X1 and X2 receive, while By1 and By2 represent the benefits to Y from X1 and X2, Cy representing the cost incurred by Y)"

On the other hand, the byproduct model (see figure 3), is separated into three categories: One-way Byproduct; Two-way Byproduct and Byproduct Reciprocity. While in the case of one-way byproduct (e.g., carrion feeders) and two-way byproduct (e.g., predator dilution in bugs), true cooperative behavior is not in earnest being selected. These models provide an important foundation for more complex behaviors to develop. Finally, byproduct reciprocity can provide a selection pressure for a type of pseudo reciprocation, where an individual's greater cooperation towards a given partner results in more benefits being received from that respective partner as a result of the cooperative behavior.

As mentioned, all of these models provide important insight into cooperative behavior. However, here we will focus on kin selection and the mechanisms and abilities associated with it. While on the surface the model proposed by Hamilton, the shared genes model or kin selection (see figure 4 from Sachs et al., 2004), seems simple enough only involving an actor engaging in behavior that would benefit a recipient with whom he shares alleles, this is not the whole picture. The model theoretically operates strictly within species and necessi-

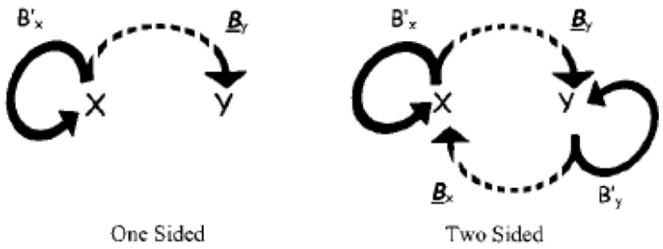

Figure 3: One-way and Two-way Byproduct Benefits (Sachs et al., 2004)

"Left: An act of X benefits Y as an automatic consequence (byproduct) of X's self-interested action (one sided).Right: Likewise, individual Y may, when performing an act that benefits itself, also benefit X" (B'x, and B'y are benefits of self-interest to X and Y, while Bx and By are byproduct benefits to X and Y. Dashed lines refer to byproduct benefits)"

tates a mechanism that limits the altruistic actions towards recipients that the actors shares a common ancestor. This ability results in different approaches as to the explanation of the process through which an individual achieves this feat: spatial association versus kin recognition, (Grosberg et al., 1986). Our ability to act according to the degree of relatedness is predicated upon our social nature as a species. This notion is predicted by the indirect and inclusive fitness benefits of investing in kin and the costs and benefits associated with inbreeding. However, for an individual of a species to gain the benefits of prosocial actions towards kin, one does not necessitate an ability to directly recognize cues of relatedness, the only thing practically needed is a high probability of altruistic behaviors to be directed towards kin. This can either be achieved by a mechanism that allows for kin discrimination or, on the other hand, indiscriminate altruism if high kin density is present through limited dispersal or budding dispersal.

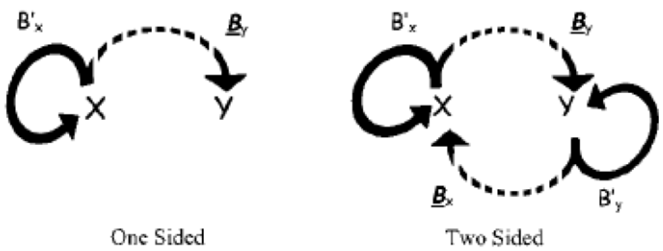

Figure 4: Kin Selection (Sachs et al., 2004; apud Hamilton, 1964)

"X evolves to benefit Y if r(By)= Cx>0. (r= coefficient of relatedness between X and Y; Cx= Cost of the act to X; By= Benefit of the act to Y.)"

3

SPOTTING KIN AND THE BLIND ALTRUIST

The necessity of a mechanism directing kin discrimination has been a hotly debated topic ever since the development of kin selection by Hamilton. And, Hamilton's rule has been largely accepted and verified by a large number of studies both in humans and in the animal kingdom as a whole, especially true in primates. Silk (2002) provides a fascinating review of kin selection and its influences in the animal kingdom from the strong coalitions of the macaques to maternal kinship the apparent driving force behind the social organization in female baboons and bonnet macaques; all this factors supporting a high likelihood of these primates and others being discriminating altruists acting accordingly to predictions one might make based on Hamilton's rule. Even so, the debate over the validity of Hamilton's rule and whether or not it presents a testable framework has continued. Reviews of this debate by Birch et al. (2014) show us the diverging opinions of a number of researchers on this topic. From remarks of Nowak et al. (2010), that this rule "almost never holds" to counters from Gardner et al. (2011) defending the applicability of this rule, be it in expanded form. This debate is highly contested and offers us a wealth of promising intellectual discussions on the nature of the mechanism, our conceptualization of it and the application of Hamilton's rule in whatever form it may come. However, our focus here will take us in a slightly different direction. While the conceptualization of formulation of rules related to kin selection might be debated, we would argue that any notion of kin selection needs to take into account and further explore the question of how. How do actors selectively interact with their kin? Any interpretation of kin selection must account for this simple notion, that for any evolutionary pressure of altruistic behavior to exist, some force must increase the probability of directing these actions towards kin be it in an ability of the actor allowing kin discrimination or the very makeup of the social structure providing the necessary enviroment.

For some species for whom their habitat consists mainly of burrows or nests localized in discreet clusters, the likelihood of an interaction partner being kin is increased making the need for a dedicated mechanism responsible for kin recognition unlikely (Silk, 2002; Blaustein et al., 1987). The social complexity and localization of conspecifics is not always so easily restrained, so as to produce the needed kin density in which indiscriminate altruism might operate. Behaviors that take advantage of indiscriminate altruism are all to prevalent in the animal kingdom, such as the infanticide in primates (Van Schaik et al., 2000). We discussed at length the issue of cooperation in chimpanzees, with the competitive nature of their environment being a major factor in determining, at least in part, the reason why they can, in some cases, reach levels of cooperation, altruism and other complex behaviors rivaling human ability while being unable to overcome others. With the increased size of the social groups that actors navigate comes greater need for a specific mechanism that maintains a high degree of interactions with kin fostering cooperation and reducing the competitiveness that stifles their ability to overcome environmental obstacles.

Here we see the continued debate over the two possible solutions offered by Hamilton himself: kin discrimination or limited dispersal. Kümmerli et al. (2009) present an in depth look at the current landscape of research into limited dispersal. They acknowledges a great deal of interest from a great many authors on this very mechanism (Pollock, 1983; Grafen, 1984; Murray et al., 1984; Goodnight, 1992; Kelly, 1992; Nowak et al., 1992; Queller, 1992; Taylor, 1992a; Taylor, 1992b; Wilson et al., 1992; Kelly, 1994; Nowak et al., 1994; Queller et al., 1989; Frank, 1998; Van Baalen et al., 1998; Mitteldorf et al., 2000; Griffin et al., 2002; West et al., 2002; Hauert et al., 2004; Rousset, 2013; Gardner et al., 2006; Killingback et al., 2006; Lehmann et al., 2006; Ohtsuki et al., 2006; Traulsen et al., 2006)

Dispersal rates and other characteristics vary greatly in nature with some organisms being motile throughout their lives, while others are adapted to move or be moved at precise, limited phases of their life cycles. Whatever the case, dispersing is a remarkable if costly behavior that does provides a number of benefits. From benefits such as the avoidance of competition with kin in the natal areas, to the sometimes high costs in terms of energy expenditure and inherent risk associated with relocation. In terms of kin selection, limited dispersal could in part create the necessary kin density in which the probability of interacting with kin is high enough for altruism to be selected in the absence of a kin recognition mechanism, thus fostering an environment in which indiscriminate altruism provides a fitness benefit. Though, as shown by Taylor (1992a) and Taylor (1992b), with increased kin density as result of limited dispersal, comes increased competition with kin, over resources. This competition modifies the relative value of any benefits that might be offered. Helping one relative to survive by providing them with food for the winter invariably leaves another relative with less food, thus if all kin are competing over the same patch of resources this quickly negates possible benefits. One possible solution to this issue is explored by Kümmerli et al. (2009) which show that "budding" dispersal, dispersing in groups of relatives, might alleviate the issue of increased competitiveness while maintaining the high kin density needed to produce altruistic behavior.

Though as shown, there have been solutions provided to issues related to dispersal as a catalyst for kin selection, it is not the only option available. Both dispersion and mechanisms allowing for kin discrimination can play an important role in kin selection. The propositions of either/or do not necessary have a place in this discussion. Even with the benefits and scaffolding provided by limited dispersal or budding dispersal, a kin recognition mechanism could help optimize cooperative actions on the actors' part increasing their overall fitness benefit. To optimize benefits, organisms have to be able to recognize and discriminatingly act according to different levels of relatedness when selecting their partners for cooperative or altruistic behavior. The important distinction to note here is recognition and discrimination. By recognition, we refer to the mental mechanism that facilitates the expression of discriminate behavior towards kin. Importantly when investigating the existence of a kin recognition mechanism what is most of the time salient experimentally is the discrimination that results from this recognition.

The duality and plurality of possible explanations and theoretical frameworks for investigating cooperation cannot be avoided even in our next step. Even if granted the need for kin discrimination requires the existence of a kin recognition ability that raises additional questions. How is this achieved? What constitute a reliable indicator of kinship? Should we explore direct or indirect mechanisms of kin recognition?

3.1 THE CUES OF KINSHIP

Krupp et al. (2011) discuss the different kin recognition mechanisms and the functional implication of their existence. One of the issues discussed at length is the notion of direct and indirect kin recognition. The notion that, the way in which these systems might be implemented remains quite basic is striking. The fact that it can be "positively brainless" as exemplified in the case of the annual plant Cakile edentula's ability to recognize kin and in reaction to this recognition retarding their root growth provides a strong starting point from which we can discuss kin recognition.

As noted, all that is practically needed is a mechanism that increases the probability of altruistic interactions to be directed towards kin. On that basis kin recognition systems have been found in many species the world over (Krupp et al., 2011). The environmental and social context has a great deal of influence in what kind of kin recognition system is best suited for the job at hand. From the positively brainless to the more "advanced" social interactions of great apes and humans, the amount of complexity required of this mechanism varies. The most straightforward way in which to analyze the different ways that kin recognition might vary is the distinction between direct and indirect kin recognition.

3.1.1 *Indirect Kin Recognition*

The distinction between indirect and direct kin recognition can become less clear when these mechanisms become responsible for giving an organism the ability to navigate complex social-environmental contexts. At the basic level this context closely correlates with the kinship of its conspecifics. This means that indirectly individuals can extrapolate kin relationships from environmental cues such as cohabitation during rearing. This strategy is a relatively straightforward endeavour when the mobility of offspring is limited making the likelihood of costly errors in discrimination unlikely. On the other hand, this type of indirect mechanism can take into account natal dispersion and developmental milestones as factors that impact the output of kin recognition systems. This in turn allows greater flexibility in determining kin from non-kin in circumstances where the offspring, as they mature become more mobile. One such case is that of Belding's ground squirrels where Holmes et al. (1982) showed that rejection rates of juveniles at their burrow are correlated with the age at which offspring start weaning. As the likelihood of interaction with non-kin during the initial stages of rearing are unlikely, this makes the the temporal/developmental cue of weaning a highly salient cue in distinguishing between kin and non-kin. So, we can see that a rather simple mechanism of location based kin recognition is further enriched by accounting for other types of cues such as developmental stages. We see the same type of social-environmental context mediated behavior when it comes to sexual attractions among closely related kin, inbreeding avoidance. One such example is the Westermarck effect which postulates that siblings that mature together undergo a type of reverse imprinting in which they become desensitized to sexual attraction of siblings. The classic case study of this type of effect usually offered is the kibbutzim communal living. Here children are raised together not based on kinship but on age. Marriages between individuals raised this way were found to be extremely rare. Only occurring in cases where they did not cohabitate during the initial years of rearing (i.e., before the age of 6). Indirect kin recognition, when compared to its direct counterpart, is an easier and more efficient way in which kin discrimination can occur. Even so we see that context sometimes determines a need for additional factors to be accounted for if kin discrimination can continue to function in cases where false positives might entail a high cost on the individuals fitness. The higher the chance of false positives the greater the need for more complex implementation of kin recognition. In that regard indirect kin recognition might not just stop being useful.

3.1.2 *Direct Kin Recognition*

When entertaining ideas of direct kin recognition, a simple way in which one can conceptualize it is that individuals build a mental template from previous experiences or start out with an innate base template for kin. This kin template is contrasted with templates representing the average local population. A new interaction partner, be them non-kin or kin, can be analyzed on this dimension matching their phenotype to one of these two points. The closer an individual is matched to a kin template the higher the likelihood that they are related to the observer. In contrast to indirect forms of kin recognition, phenotype matching not only allows asserting the kinship relationship with novel interaction partners, but also degrees of relatedness of kin. This allows a greater degree of flexibility in the types of behaviors that can be selected. Costly actions are made even more dangerous when coupled with higher chances of false positives in kin recognition. False positives are not the only issue here. Say, an organism falsely classifies a new interaction partner as kin, when in truth this partner is wholly unrelated to it. The fitness cost of any altruistic actions would evidently make such actions highly undesirable. Biologist Haldane's famous question is fitting: "Would I lay down my life to save my brother? No, but I would to save two brothers or eight cousins."(Gould, 1976). Attributing the same level of relatedness to an offspring as compared to a third cousin is also damaging. Granted, the cost/benefit ratio is less damaging, though it nevertheless does not favor the same type of altruistic actions being undertaken for both the offspring and the cousin. Phenotype matching allows this type of fine tunning when the social-environmental context necessitates it. Experience can also allow individuals to encode the phenotype of their kin. As early association is highly correlated with kinship this allows building a robust template that can help determine if future interaction partners are indeed kin or not by contrasting and comparing observations of that partner with this kin template.

On the other hand, an innate template would require, at the very least, some type of template to be present from the start. Definitions vary as to what consists an innate mechanism of kin recognition, but at the very least a predisposition to attend cues correlated with kinship would be a good start. Rosa-Salva et al. (2010) investigation if chicks have an innate preference for faces finds that "2-day-old domestic chicks, visually naïve for the arrangement of inner facial features, spontaneously prefer face-like, schematic, stimuli." And recent studies have found the same face preferences in humans, with researchers Reid et al. (2017) using lasers to project a series of three red dots onto the abdomen of the mothers. With the lasers being bright enough for the fetus to see from inside the womb. The idea is that if the fetus have an innate (or at least developed in the womb) preference for facial-like stimuli they will turn their heads and attend to these stimuli at a higher rate above chance from the inverted stimuli. This is precisely what they find and it provides a striking result. While these results are not the holly grail of innate template for kin recognition, they do show that the two diverging opinions here require more nuance than nature vs. nurture thinking. A preference for facial stimuli, even if it would be the only thing that we ever discover to be innate, would be a very important mechanism that allows future development of any ability to directly recognize kin. By directing attention towards faces, this preference might be the catalyst necessary to build and develop a base template of kin. In either case the starting point for this recognition is the ability to distinguish conspecifics relatedness on an individual basis. As we discussed, these templates can be based on the information gleamed from observing others (i.e., other-referent phenotype matching). The validity of others as referents for kin in this case would be based on indirect kinship cues . Conversely, another possible mechanism of creating kin templates could be based on traits of the individuals themselves, be that through self-inspection or other observations such as reflections (i.e., self-referent phenotype matching).

Research within the animal kingdom of the mechanisms behind the "assembly" of kin templates shows that self-referent phenotype matching seems to be less common. The issue with this type of phenotype matching is the difficulty in divorcing its predictions from those made of other-referent matching. In most cases, one would except that both matching techniques produce functionally identical results. One of the exceptions that has been tentatively examined is that in which parentage is highly-uncertain. This context allows a clear distinctions between the different predictions one might make using the two phenotype matching mechanisms. Krupp et al. (2011) provides interesting review of research in this field. The prominent example they give is that of the peacocks. What stands out is that even when they are not raised together, peacocks lekking is still assorted as a function of relatedness. They make clear that this result needs further investigation. Nevertheless, this example and others given from the animal kingdom are fascinating. These range from the golden hamsters (Heth et al., 1998; Mateo et al., 2000), brown-headed cowbirds (Hauber et al., 2000), white-bearded manakins (Shorey et al., 2000), to the chacma baboons (Alberts, 1999).

This distinction between self vs. other referent matching harkens back to our earlier discussion of the duality and plurality that seems to propagate throughout this field. While it could be argued that we only need one type of referent, it seems far more likely that, again, we are faced with shades of gray type of scenario in which we might need both, depending on the social-environmental context. In this case, other-referent phenotype matching is by far the more well explored mechanism. One famous example is the green-beards of Dawkins (1976) which he first used to exemplify the notion that an allele can present a marker that can be readily identified through direct observation. In this case individuals grow beards colored green which present a strong indication that an other is carrying the same allele as the observer thus allowing them to discriminatively cooperate with kin. While initially designed as a simple thought experiment, green-beard type effects have been found in a number of species. Examples range from odor as a marker for the "revolts" of red imported fire ants against their homozygous (BB) queens when they do not share the same genotypes (Keller et al., 1998); to the spermatozoa of the common wood mouse creating "trains" that help propulsion over and above what could be achieved individually, based on a unique morphological transformation (Moore et al., 2002). Some studies such as that of Bressan et al. (2009) have tried to investigate the issue of self vs. other-referent using facial stimuli in humans. Their results shows that there is a least a tentative indication that self-reference might be part of kin recognition in humans. And while the focus of our research will be on direct kin recognition, as far as the phenotype matching mechanism involved in producing kin templates is concerned, this research will remain agnostic as to its origin. Whether the kin template is based on self or other referencing, the putative cue of resemblance (i.e., facial stimuli) we choose to focus on in our research, provides functionally identical predictions for the purpose of our experimental design.

Stepping back from this distinction of self vs. other-referent, in the larger scope of kin recognition mechanism in humans, there have been a number studies investigating direct and indirect kin recognition. While in the case of indirect kin recognition we have examples such as the Westermarck effect; in the case of direct kin recognition much of the research has been made possible with the development of methodologies to produce and test putative cues of kinship in particular facial cues of kinship.

4 THE FACES OF KIN

In the rat race for understanding the mechanisms and foundations of human prosociality, a number of fields have had large contributions. From evolutionary biology and psychology to behavioral economics. In a previous segment we discussed the importance of the iterated prisoner dilemma in understanding social interactions and building a foundation for models intended to elucidate human nature such as partner control. Economic games play a large role in helping us conceptualize human behavior by creating controlled environments ripe for experimentation. A good example of the benefits and pitfalls of this approach is the ultimatum game.

As with any economic game there are a number of different variants that have been used throughout the years. In its basic form, two players engage in an economic decision involving a proposer and responder. The proposer is given a resource by a third-party and is tasked with offering a portion of that to the responder. In turn, the responder can either accept the offer in which case the resource are distributed as per the agreement or they can refuse which results in neither player receiving anything. From an economic rationality standpoint the success of the strategy chosen by the individuals depends on their ability to optimize payoffs. In this context the offer made needs to be the smallest acceptable option. While on the other hand the responder needs to optimize their response in such a way as to receive the maximal possible amount. This results in the Nash equilibrium for this type of one-shot ultimatum game is to accept any offer that is greater than zero while offering any amount higher than zero.

The surprising result from the point of view of a rational human actor driven solely by self-interest is that humans participants do seldom if ever follow the most optimal strategy vector (Güth et al., 1982). They seem to be more generous than one would expect (Fehr et al., 2006), offering more than economic rationality would predict. Taking fairness into account, they do not expect more than they would be willing in turn to offer (Nowak et al., 2000). A sense of fairness basically entails that one should "treat others with impartiality and to share the costs and benefits of cooperation equally" (Baumard et al., 2013). Thus, everything else being equal, an individual endowed with an innate sense of fairness will be more likely selected for cooperative tasks than an otherwise selfish individual that "acts" fair because it suits them (André et al., 2011). In this framework individuals have proximal moral sense, they are not selfish individuals who take into consideration the cost and benefit of their cooperative actions; at the proximal level they genuinely are moral, fair. They genuinely want to be moral, acting in such a way as to maintain fairness, which brings with it positive feelings; while acting unfair, does not. This is because, at the ultimate level this behavior is based on the fitness benefits fairness provides (Baumard et al., 2015). Kinship should also play a large role in modifying individuals responses to ultimatum games. Traditionally, economic games are produced following strict principles of anonymity. Which is one of the many explanations offered for the diverging strategy vectors chosen by human participants. Studies such as those of Hintze et al. (2016) show that by introducing ability to detect kin in computational evolutionary models deviations from the Nash equilibrium can be more easily explained.

Human behavior cannot be divorced from the evolutionary context in which it has developed. Kinship plays an integral role in the shaping of the social environment and removing this factor from any computational model cannot hope to capture results from actual human participants. We would expect humans to take into account kinship when dealing with situations akin to an ultimatum game. The proposer and responder divorced from cues of kinship would ultimately still have to take into account maximal payoff structure within a world in which kinship is actually present.

Kin selection and its effects on our evolution as a species, have been numerous, confounding our expectations of rationality with a new incentive to integrate other factors in our analysis. This need to integrate findings has historically resulted on focus on developing models that can account for our seemingly altruistic and cooperative nature that is over and above what we find in the the animal kingdom. As a section of this topic, research on kin recognition has been a widely investigated issue. Some research has focused on the effects and predictions made based on co-residence (Lieberman et al., 2007) and Westermarck effect in terms of indirect kin recognition. While others, have focused on direct kin recognition and the possible mechanism that would allow individuals to spot whether a conspecific is related. As we have seen throughout our discussion here, there is a great deal of proximal cues that can help "build" a kin template. Research in this field has been split between direct and indirect kin recognition. Many having focused on direct kin recognition. Techniques used thus far in the literature have been directed almost exclusively towards creating procedures that allow researchers to manipulate facial stimuli in such a way as to create an effect of self-resemblance. That is to say, researchers have used the participant's own face to create their stimuli. This has presented us with both strong literature resources to draw upon in developing our own procedures and a challenge in adapting these previous findings to the expanded scope of our experimental endeavours.

Why would facial similarity be a putative cue of resemblance? One clear connection can be seen when analysis predictions people make as to the relatedness of others through facial stimuli presented to them, where these judgments are associated with actual kinship (Kaminski et al., 2009) and in turn facial similarity predicts judgments of relatedness (Maloney et al., 2006; DeBruine et al., 2009). In addition, researchers like DeBruine (2002) and DeBruine (2005) and Krupp et al. (2008) have shown that in accordance with predictions one might make based kin selection, participants are likely to be more trusting and cooperative with self-resembling interaction partners. When mating choice comes into play participants seem to find opposite-sex faces that resemble themselves as being less attractive (DeBruine, 2005; DeBruine et al., 2011).

All of these findings provide a strong basis to ascertain that the manipulation of similarity in facial images is for all intents and purposes creating a putative cue of resemblance, of kinship.

4.1 FACED WITH THE KIN OF OTHERS

One issue with research of kin recognition has been the understandable focus on investigating the ability to recognize one's own kin. However, in the ever expanding social landscape of complex relationships and intertwined coalitions of individuals, the relevance of kinship seems self-evident. These far ranging interactions would inevitably lead to individuals coming into contact with groups that are comprised of related individuals not to themselves, but to each other. In this context, an ability to recognize third-party cues of kinship would be beneficial to determine the most efficient course of action be they selfish, cooperative or altruistic.

Parent-offspring interactions have been the main focus of research thus far, leaving little room to directly analyze whether this ability to recognize kinship extends to unrelated third parties. Researchers such as Lieberman et al. (2008) stress the fact that in the field of psychology, kinship has been "largely neglected". Using a memory confusion paradigm they show that kinship does seem to be in fact an important dimensions of social categorization. To bridge this gap in our understanding of human kinship recognition, there have been a some researchers posing new and interesting questions as to the scope of our kinship recognition ability. The first and most important question being whether or not humans are able in fact to detect third-party kinship. On this topic, the work of Maloney et al. (2006) is invaluable. Being one the first to investigate this topic, they find that in fact unrelated third parties are able account for facial similarity as reliable predictor of kinship by labeling children as siblings accurately. This finding is also expanded upon by Alvergne et.al.(2007) to unrelated individuals ability to the detect the kin relationship between parents and their offspring. In addition, the work of DeBruine et al. (2009) replicates the initial findings of Maloney et al. (2006) using adult stimuli. Strikingly this ability is unaffected by exposure, participants maintaining their accuracy cross-culturally as seen in the work of Alvergne et al. (2009a). This growing body of evidence seems to make clear that humans are in fact influenced by facial cues of similarity when dealing unrelated conspecifics. The issue becomes not their ability to detect similarity in others, but whether or not this ability is connected to kin recognition or is it a simple by-product of more general facial recognition processes. Raising this issue creates a clear need to disentangle similarity and kinship. Facial similarity seems to undeniably capture information related to kinship. However, this high correlation between kinship and facial similarity creates a type of feedback-loop confounding our analysis of the mechanism behind it. Individuals ability to spot similarity might be unrelated to their ability to detect kinship in third parties as both would provide the same result when faced with context-independent experimental settings. When it comes to first party kinship recognition (i.e., recognizing one's own kin) this issue has been partly resolved by the growing number of experiments that find its effect in a number of different experimental settings (i.e., cooperative endeavors vs. incest avoidance) offering a glimpse at this mechanism's context-dependent nature. This nuance has been unfortunately lacking in third party kin recognition studies.

Our earlier discussion of kin selection provides the perfect springboard for this very problem. From the family to the clan kinship has played an important role in shaping our cooperative behaviors, social relationships, group dynamics and mating choices. Individuals should selectively interact with kin. This allows a pathway towards collective actions that benefit groups of individuals. The kin relationships within those groups provide the spring board to increase the likelihood of these types of collective actions being undertaken in the first place. This interwoven nature of kinship, groups and collective actions form the basis of strong cohesive group structures where altruism is fostered (Lieberman et al., 2007; Lieberman et al., 2008) and aggression is lowered (Daly et al., 1988). At the same time, kinship structures behaviors withing the domain of mating, avoiding kin as sexual partners and shaping moral sentiments relating to third party incest (Fessler et al., 2004). The nature of kin relationship and their effects on group dynamics coupled with findings that support the notion of human ability to detect third party kinship leads us to a clear avenue through which to disentangle facial similarity detection from kinship recognition. Put simply in the words of Lieberman et al. (2008), categorization is for doing. Individuals possessing a third party kinship recognition ability should act in a context-dependent way in accordance with the indirect fitness impact of their actions.

To exemplify this call to action, economic games again play an important role in directing our exploration of this topic. During our discussion of the ultimatum game we noted that human behavior cannot be divorced from its evolutionary context and kinship plays a large role in it. In the ultimatum game we were discussion the simpler case of two person interactions, however the same remains true for circumstances in which individuals have to interact with groups. Fittingly for our purposes, in their investigation of the evolution of fairness through computational models, Santos et al. (2015) use a multi-player extension of the Ultimatum Game. In this variation proposers are given resources that should be divided with the members of the group. Responders, as in each member of the group, can either reject or accept this offer based on their individual expectation threshold (q). If the number of members accepting the offer meets a minimum quorum (M) set by experimenter then the offer is accepted by the group and the resources are distributed evenly among responders while the proposer keeps the rest. Conversely if the minimum quorum of yes votes is not met the payout is lost. By modeling a population of learning agents interacting within this context they show that, tentatively, stringent group criteria result in fairer offers. It seems that the evolution of fairness within multi-player ultimatum game is highly dependent on the acceptance threshold (M). A stricter more cohesive group, higher M, would result in fairer offers (p) by the proposer as this would mean the the offer has to accepted by more responders. At the same time responders are free to increase their expectation threshold (q) which makes it highly beneficial to be part of this group just through the direct fitness benefits gained. Coming back to kinship, research has shown that it increases group cohesiveness through cooperation and decreased aggression. Then, if kin relationships are a reliable predictor of group cohesiveness it follows that kinship is also a predictor of an increased M within the same group group. Thus, a kin recognition module might help direct individuals, looking to participate in cooperative actions, towards joining groups with high kin density which provide a productive and fair environment for cooperative activities. On the other hand, individuals looking to undertake selfish actions would likely seek out group in which kin density is low, signaling a potentially less cohesive group, where M would be lower, thus allowing for more unfair offers.It is important to note that experiments with human participants have not, to our knowledge, yet been implemented utilizing this multi-player ultimatum game framework. Nevertheless, these speculative behavioral characteristics, fit the context-dependent nature of actions resulting from processes taking into account outputs from kin recognition mechanisms. While not using this specific multi-player ultimatum games framework, studies have investigated some of these predictions. Researchers such as Dasgupta et al. (1999) have investigated how perceptual similarity among group members might influence impressions of unity of these groups. These studies have used physical characteristics such as body color (Dasgupta et al., 1999) and body movement (Ip et al., 2006) to show that increased similarity produces this effect. Other studies such as Wang et.al,(2016) have used facial stimuli to experimentally manipulate the perceived similarity of groups. The results of Wang et al. (2016) seem to indicate the same pattern of cooperative intent assigned to groups which appear to be perceptually more similar.

Not all predictions based on kin density provide positive reasons to join certain groups. High kin density can signal an environment in which nepotist behavior is highly prevalent. Nepotism, as in the preferential treatment of kin over non-kin, is only highly beneficial if one's on the receiving end of such preferential treatment. Participating in collective actions of groups with high kin density as an unrelated member (i.e., non-kin) runs the risk of being subject to unfair distribution of payoffs. In traditional economic theory nepotism is seen as a clear detriment to the profitability of certain "modern" collective endeavors such as corporate decisions making and profitability (Bennedsen et al., 2007). This balancing act between the benefits and costs of nepotist behavior have always been a factor in human development and is just another example of how context-dependent our analysis must be.

The same remains true for mate choice and the predictions kin selection allows us to make. When dealing with direct kin recognitions (i.e., recognizing one's own kin) the same type of balance must achieved between the opportunity costs of rejecting kin as a potential mate vs. avoiding the costs of inbreeding. Expanding this notion to the kinship ties of unrelated others, offers some of the same issues. Predicting the behavior of others is a clear benefit in determining the best course of actions. Be this in ascertaining the cooperative intent of others or whether they will be a potential sexual rival. Categorization in service of action. In this case recognizing the kin relationship between a potential mate and a potential sexual rival can help direct behavior. If they are recognized as being related this would modify the perception of conflict and costs associated with this potential mate by removing a potential rival from the list of concerns. Investigations into incest aversion and third party incest have been numerous, in regards to unrelated third parties, incest taboos and prohibitions have been some of the hallmarks of this research topic. Westermarck (2006) viewed social norms related to incest as a continuation of an individuals own aversion to engaging in incestuous sexual activity. At the ultimate level incest taboos and prohibitions can be viewed as investment by individuals to avoid the cost of inbreeding depression. While the proximal mating of two related individuals does not pose and fitness cost on a unrelated third observer, it does present future opportunity costs and potential inbreeding depression resulting in fewer viable mates. Again this is a balancing act, as too stringent third party incest prohibitions coupled with strong first party incest avoidance can lead to outbreeding depression. While there have been many studies investigating prevalence of third-party incest taboos (e.g., Antfolk et al., 2012; Fessler et al., 2004; Lieberman et al., 2003) and some connecting direct kin recognition and inbreeding avoidance both through possibly olfactory-based mechanisms (Weisfeld et al., 2003) and facial self-resemblance cues (DeBruine, 2005), the effects of third party kin recognition have not been as popular a topic.

Testing both predictions made based kin selection related to mate choice and prosocial tendencies in third-party scenarios within experimental sound environments is a very important requirement for any further developments to be made in this field. To achieve this, one must first investigate what has been one of the most fruitful avenue of research in recent years, facial stimuli.

4.2 THE GENESIS OF A FACIAL STIMULI

It is no surprise that the main drive in this field has been the development of software that allows the manipulation of facial stimuli, creating a new avenue through which one can investigate kin recognition and its effects on decision making and moral judgments. Psychomorph (DeBruine, 2002; Tiddeman et al., 2001) is one such piece of software. To better understand how participants would respond to our stimuli, we must first understand how these stimuli are created using this type of software. Thus, here we will discuss the steps involved in averaging, transforming and morphing two or more faces to create facial stimuli that, when presented to participants, will be able to elicit our desired effect.

One of the first examples of manipulation of facial stimuli was achieved by Francis Galton almost 100 years ago. The procedure he used is very similar to what software today allows researchers to do through a process called warping. His breakthrough was achieved by effectively layering a series of faces which had their eye position aligned, creating a multiple exposure, composite of those faces. Today warping is why modern researchers are able to create highly complex face composites to be used in experimental settings. By unchaining researchers from the need to recruit participants that are able to bring in or provide pictures of siblings and other family members, this digital ability to create the carefully controlled stimuli needed for their experiments has led to a growing interest in this field.

Software such as Psychomorph (see figure 5) is in a sense a direct successor of the multiple-exposure procedure used by Galton. By analzsing the shape texture and color of the facial images provided, the software compresses and stretches the shape of those inputs together into composite images in which the shape is a weighted average of those values. In addition, the pixel color and texture of the image can also be averaged, thus creating a "complete" composite face.

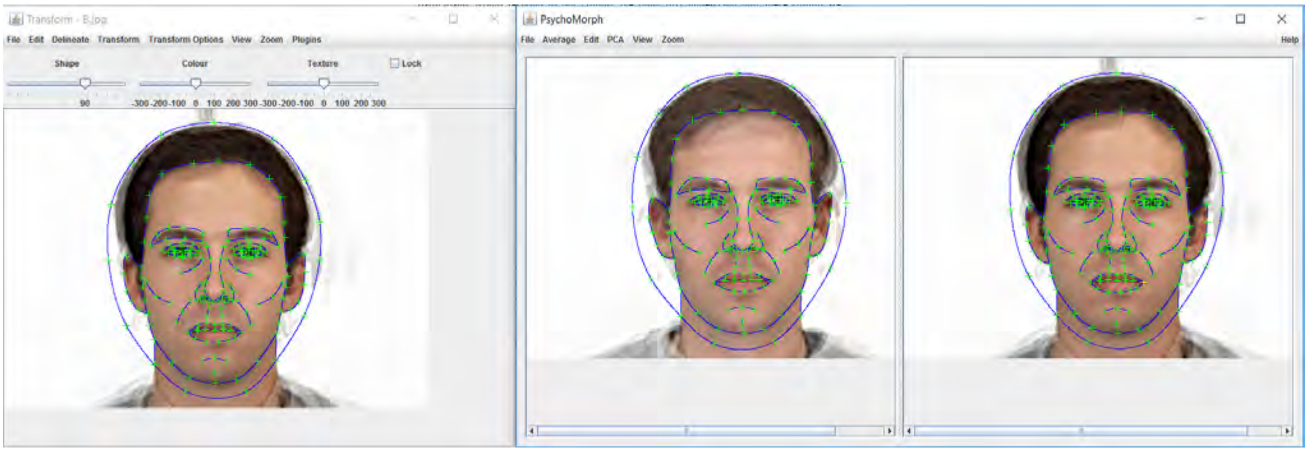

Figure 5: Psychomorph Interface.

The process itself begins from the moment the photograph is created. To achieve the best results when manipulating faces, one must take great care when producing the pictures themselves. As one might expect, quality pictures of higher resolution with the background and lighting standardized offer a more robust opportunity to create higher quality composites and morphs during the manipulation procedure. In addition, another important aspect is the position of the face itself in the frame, with wildly varying frame positions making aligning the faces during the manipulation procedure more problematic and reducing quality. These issues can be remedied digitally, however as a rule it is better if control is maintained when taking the pictures rather than applying corrections digitally. Another important aspect that must be carefully controlled are the emotional cues that participants present when their picture is taken. As one might expect, the ability to precisely control what goes into manipulation is highly important, as emotional cues such as those of happiness can affect how "drawn" our participants will be towards a certain picture. In the same way, clothing, jewelry or hairstyles that vary wildly among the faces that will be used to create stimuli using these techniques can create graphical artifacts and other unintended effects that might impact experiments. All of these issues are controlled in the Chicago Face Database (Ma et al., 2015) which we use to create our stimuli.

With the pictures in hand, be them the result of the researchers concerted effort to photograph their participants or from high quality face databases, the manipulation process can begin in earnest. The first step in any such endeavour is aligning the eye positions after which the process of delineating the fiducial point of the faces begins. Delineating the faces involves creating a template that will "tell" the software where the fiducial points (i.e., "feature points") of the image are located; in this case the mouth, nose, eyes and other important facial features. In general, 179 fiducial points (example using 189 fp, figure 6) are used when delineating facial images, however this figure might vary according to a number of aspects specific to each research objective. An increase in the number of fiducial points may allow greater quality in manipulating complex facial features such as the eyes. However, this is highly dependent of the quality of the base picture in terms of resolution. Higher resolution pictures provide greater robustness, but require greater attention to detail and an increase in the number of fiducial points to avoid graphical artifacts. That is not to say that lower resolution images do not present the same issues: distortions are just far more likely to be evident in higher resolution pictures if the delineating process is not carefully implemented. In our research the majority of studies use 189 fiducial points applied to high quality pictures with a resolution of 2,444 width by 1,718 height from the Chicago Face Database.

Figure 6: Delineation using 189 Fiducial Points.

The template of the delineation process allows us to use Psychomorph to continue the manipulation process by using either morphing or transforming to create our resemblance effect. In this research we used a combination of these two techniques to create our stimuli. However, in the literature, studies usually use either morphing or transforming to create their stimuli. Both of these techniques offer certain benefits and disadvantages which we will discuss now.

When discussing morphing (also known as averaging), the principal concept behind it is that, with the help of two images that can be conceptualized as endpoints of a continuum we can move, by a specified percentage, the target image's fiducial points towards the other image in terms of xy-coordinates. The target image (i.e., the image which is manipulated) is defined by the positive or negative value of the determined percentage. Negative values move the second image towards the first while positive values have the reverse effect. This distinction is important because the target image maintains its specific characteristics. Briefly, this is done by applying the specified percentage of the distance between the xy-coordinates of one of the endpoint images and the other to the target image. When only using the steps outlined thus far, the image that would result is catalogued as a shape only morph, as this image maintains the colour and texture characteristics of the target image. However, additional steps can be taken to create a full composite by also modifying the pixel color and texture in the target image according to the specified percentage of the difference between the two endpoint images. Some experiments require these additional steps by design, such as investigations where the perceived age is manipulated. As important cues of age (e.g., hair colour, wrinkles, etc.) are derived from colour and texture of the face presented, using shape-only morphs would be detrimental. However, even so the nature of morphing limits the scope of the manipulations that can be done in such experimental settings. One of the most frequent applications of the morphing technique is creating so called prototype images. Prototypes are composites of multiple images in which the average shape, colour and texture of the set of images is used to create the final composite image. These types of images are extremely useful in creating prototypes of the average man or woman or other definable groups, which can be used in conjunction with the transform technique to create even more complex manipulations.

As mentioned, morphing presents us with some distinct disadvantages, such as when we want to manipulate a target image in a set direction on a definable spectrum. While it would be possible to use our target image as one of the endpoints of that spectrum, we are faced with the issue that the target image intrinsic characteristics might diminish the quality of our manipulation. To continue the previous example, say our participant is middle aged and we wish to manipulate their apparent age. We first create a prototype image of the average elderly person and use this image as one of our endpoints. However, we still need the other endpoint of our spectrum to create our morph and that is where issues might arise. By using the participant's image as the other endpoint of our young, old spectrum we might be greatly limiting our ability to apply specific cues of age in our manipulation. Even though the participant's age is average they might already have facial characteristics that evoke a more advanced age. Thus, by using their image directly in our spectrum, the value applied to them as a percentage of the difference between their face and our average elderly person would be insignificant, failing in manipulating their perceived age. This case is a prime example of why modifying our technique to take advantage of the benefits of transforming is sometimes necessary. That is to say that the solution to our dilemma is to replace our target picture as the "young" endpoint of spectrum with a prototype face of the average young person and using the differences between young and old to manipulate our participant's face.

When transforming a target picture, we are modifying this picture by moving the fiducial points by the specified percentage of the difference between the two endpoints of our spectrum. By doing this we are ensuring that our manipulation of the target image is limited to our desired trait, in this case we are targeting general facial features that influence perceived age. There are of course a number of variations on both the transforming and morphing techniques, developed to overcome various shortcoming that might arise from this process. These issues vary from the attractiveness effect, that results from increased averageness of prototype images (Langlois et al., 1990); the difficulty in creating natural other-sex morphs; or graphical artifacts when creating full composites or transforms that manipulate the colour. All these techniques and caveats play an important role in our research of both expectations of third-party nepotism and mating preferences.

As previously mentioned, studies thus far have generally involved participants being presented with facial stimuli that have been manipulated to be more similar to the participants themselves (i.e., using so called self-morphs). As I have shown, this line of research has proven to be extremely interesting in expanding our understanding of kin related behaviour and the decision making process involved. Both lines of research followed in this research involve utilizing some of the same methodological and experimental procedures used in the experiments produced in this field (e.g., DeBruine, 2004; Tiddeman et al., 2001). However, as the scope of our research has been to expand these findings to third-party expectations, we have adjusted our procedures and techniques to account for these changes. Using the Chicago Face Database, we created novel stimuli using a combination of transforming and morphing images from this database in such a way as to elicit an effect of similarity between agents present in our scenarios while maintaining the participants ignorance of the manipulation effect and procedure, a manipulation procedure which we will explore in depth in our experimental section.

Indeed, our research focus has been twofold. The first line of research focused on investigating the relationship between third party kin recognition and expectations of nepotistic behaviour. The second line of research focused on the relation between third party kin recognition and expectations of decision making related to mating behavior. Though any meaningful investigation of kin recognition, no matter the complexity of its methodological framework or viability of presented stimuli must account for one important aspect. Categorization is for doing. Before getting into the experimental depth of our investigation here it is important to draw attention to our continued use of context-dependent as a description of the nature of kin recognition mediated behavior.

5 CONTEXT MATTERS

The construction of our experimental design must take into account the different predictions that kin selection requires for a perceptual recognition to be connected to kin recognition. We discussed at length the issue of disentangling similarity and kinship when presenting conclusions based on past research on third-party kin recognition. Imagine participants are asked to predict the actions of of a unrelated individual. That individual is first faced with a choice of cooperating with a similar partner and a dissimilar partner. Subsequently, the same individual is faced with a choice of punishing a similar partner and a dissimilar partner. The valence of the question being posed to the participants has changed. A context-dependent response dictates that participants will take notice of this and will predict the preference of similar partner in the first case and the dissimilar one in the second. As opposed to this, if participants ability to recognize similarity is not connected to a kin recognition then we would not expect this pattern of predictions. Possible results could range from choosing the similar partner in both cases, signaling a possible perceptual matching bias. Or the similarity of the partners might not affect the participants predictions either way, signaling no significance being assigned to similarity.Additionally, these predictions extent to third-parties choices of mates. Here predictions made based on output from kin recognition mechanisms are for all intents and purposes the opposite of our first example. With participants predicting the choice of the dissimilar partner as a mate.

In our use of the term, context-dependent responses encompass both the distinction between the domain of prosocial activity and mating choice, and the valence of encounters and actions within those domains (i.e., seeking vs. avoiding; accepting vs. rejecting; etc.). With our present research we hope to create an experimental framework in which this aspect of kin selection is accounted for in determining if humans indeed posses an ability for third-party kin recognition. In the following chapter we will begin by discussing the first line of in this research.

Part II

EMPIRICAL STUDIES

6 THIRD-PARTY EXPECTATIONS OF NEPOTISM

6.1 INTRODUCTION

While research on direct kin recognition has been instrumental in broadening our understanding of the dynamics of kin relationships, another important facet of kin relationships is, as mentioned, how third-parties recognize and predict the behaviours of other individuals which are related to each other. If human behaviour can be predicted in some cases based on the individual's kin relationship to others, it then follows that an ability to recognise the nature of those relationship and predicting behaviour based on them may be highly beneficial in navigating a social landscape.

When it comes to self-resemblance, Hauber et al. (2001) point towards the possibility that genetic relatedness might be extrapolated from how much others resemble ourselves. Self-resemblance of course would not be the only way to spot kin and, as we discussed previously, there are a number of other ways in which humans are able to determine their kin relationship with others. However, the notion that self-resemblance plays a role in determining relatedness partly motivates the current research. The two-pronged argument follows as such: if individuals are expected to recognize kinship, thus implicitly self-resemblance, and their behaviour towards kin is affected by the costs and benefits of such social actions from an inclusive fitness perspective, then, it follows that kin directed behaviour will be context-dependent, increasing cooperation in social circumstances and decreasing attraction when it comes to mating related interactions. So, if this argument is accepted we have established that individuals have the ability to recognize kin and subsequently react appropriately based on this information. Does this ability imply that humans are also expected to respond to cues of kinship between third-parties and thus make predictions in the same context-dependent way as they do when responding to self-resemblance?

People cooperate more with people who look like them, and this has been interpreted as an effect of kin recognition coupled with kin cooperation (Krupp et al., 2008). People also seem to expect agents that look like each other to cooperate more and this, again, can be again tentatively interpreted as an effect of third-party kin recognition coupled with the same expectation that kin cooperate (Wang et al., 2016). This inspires our goal to investigate whether people expect agents that look like each other to provide benefits to each other in a work context, in which this preference would amount to nepotism. Will people expect bosses to promote employees that look like them? Or conversely, to fire employees who do not look like them? This latter treatment will be especially important for the purpose of interpreting results. Indeed, suppose that we find that people expect bosses to promote employees that look like them (i.e., the boss). This could be due to the fact that people implicitly detect kinship and predict nepotism as a consequence. But this result could also be due to some perceptual matching bias; in other words, a tendency to select the employee that looks like the boss, whatever the question is. To tease out these two explanation, we need to ask other participants who the boss will fire. If they still select the employee who looks like the boss, we have evidence for a perceptual matching bias. If they select the employee that does not look like the boss, then we have evidence for a kinship-nepotism explanation.

Our experiments use a combination of facial stimuli, vignettes and performance cues. The flow of our experiments, briefly, begins with participants being presented with a vignette/scenario in which the boss agent needs to make a decision on firing or promoting one of the two candidate agents. The faces of these agents are presented in addition to performance review sheets for each candidate. At this point participants are asked who they expect the boss agent to either fire or promote.

6.2 FACIAL STIMULI PREPARATION

The procedures and techniques used to create our stimuli were the same across all our studies. The only difference in terms of stimuli was that our first pool of subjects was separated into two groups. Our female participants in that pool took part in a similar, but unrelated study and were shown a different set stimuli using the CAL/-PAL Face Database (Minear et al., 2004; Ebner, 2008) images to create composites and respectively the final stimuli presented to this set of participants. We discontinued its use in subsequent experiments as the higher resolution images from the Chicago Face Database (CFD) allowed us to create higher quality composites as compared to the 640x480 pixel resolution of the CAL/PAL Face Database. An additional difference resulting from the characteristics of the images from the CAL/PAL Face Database was that they were delineated using only 179 fiducial points, however this change does not present a distinct departure from other studies presented here, as the higher number of fiducial points was in part implemented to complement the increase in resolution of the new images from the CFD. The base images that were used in our experiments were selected at random, however we made the decision to overall limit this selection on the basis of two characteristics. As we wanted to first create a series of distinct composites that would represent base agents that would be later manipulated according to the requirements of the experiment, the first characteristic chosen was to maintain the composites as same-sex and same-race. The same-race characteristic was necessary as using images of multiple races when creating composite images limits the quality of these composites. Similarly, the same-sex requirement was maintained to avoid the side effect of increasing the androgyny of some of the composite images.

After the pictures selection, they were randomly distributed into 6 groups of 5 images each, which were combined using the averaging technique outlined in chapter 4.2. From these groups, 6 distinct composites were created. In addition, an average male prototype was created from the majority of white male faces in the database. This procedure was used to control for any distinctive facial feature that might be present in any of the individual faces. We created these distinct composites and assigned them a proto-identity that can be in turn manipulated in terms of similarity. Accordingly, these composites were each designated to a corresponding proto-identity, such as the proto-boss agent (pB), proto-candidate I (pC1), proto-candidate II (pC2) and finally as a control we designated another face as a control boss (CB). Using these base images and the other 2 composites images, that were not assigned to any identity, we constructed the final stimuli corresponding to 3 agents that will later be presented to our participants. The procedure used the following steps:

Table 1: Nepotism Manipulation Procedure Steps

1.	Boss = pB + 90%(Mcom1 - aMale)
2.	Control Boss = pCB + 90%(Mcom2 - aMale)
3.	Candidate II = pC2 + 90%(pCB - aMale)

Note Proto-Boss (pB); proto-candidate I (pC1); proto-candidate II (pC2); control-boss (pCB); average male prototype (aMale); Male Composite 1 (Mcom1); Male Composite 2 (Mcom2); 90% (the strength level of the similarity manipulation)

The important procedure here is related to Candidate I to which we applied the similarity manipulation. To achieve this, the image designated as proto-candidate 1 (pC1) had its similarity to the boss agent manipulated by applying 90% of the difference between the face of the boss agent, in terms of xy-coordinates of the fiducial points delineated on this face, and the prototype of an average male (aMale) which was constructed using all the male faces in the CFD. Both the second candidate and the boss were created using the same procedure. In the case of the boss we transformed proto-boss composite by applying 90% of the difference between one of the unassigned male composites (Mcom1) and the prototype of an average male. On the other hand, to maintain consistency and avoid any discrepancy in the way in which we created the two candidates, we used the designated control boss prototype which would take the place of the boss in the procedure used for this candidate.

6.2.1 *Pretest: Manipulation Strength*

We used the 90% level as the final strength of our manipulation, as opposed to 50% or 60% which are commonly used in the literature when investigating direct kin recognition. To determine the strength of our manipulation of similarity, we first ran a pretest in which we investigated the different threshold levels (see figure 7) at which our participants would be able to spot at an above chance level which of the face they were presented resembled the image of the designated Boss agent. This step was necessary as the literature was not as developed as direct kin recognition when it comes to determining the strength level needed for manipulations of third-party similarity.

In addition, we included basic demographic information in our models (age, sex) as well as information about family structure, namely the number of siblings and whether the participant was the youngest sibling. Studies involving third-party kin recognition have been limited and as such little information can be found on the possible influences of age and sex in relation to recognizing kin relationships in others. In terms of direct kin recognition, studies have found that sex and age differences do not influence judgments of resemblance of child sibling faces (Maloney et al., 2006). However, in terms of adult faces DeBruine et al. (2009) did find that resemblance perceptions were enhanced for same-sex vs. opposite-sex kin. If we conceptualize this ability for third-party kin recognition as an extension of direct kin recognition it follows that it might be still be subject to some of the same influences. The familial structure has been shown to influence a number of factors. For example the maternal perinatal association, that is the close association that older siblings observe between their mothers and a newborn. Individuals subject to the MPA, being the older sibling, have higher levels of altruism and incest avoidance toward the younger sibling. The nature of these developmental environments might impose different requirements to enhance kin recognition mechanisms. MPA provides strong, reliable cues of kinship making direct recognition mechanism less important in older siblings modulating this abilities performance. Conversely, this means that we would expect the younger siblings to have better performance when using facial kin recognition. This factor might be highly relevant for our pretest here in participants accuracy, but also for our participants predictions in our experiments proper. As such we used these control variables in our pretests here systematically in all our experiments.

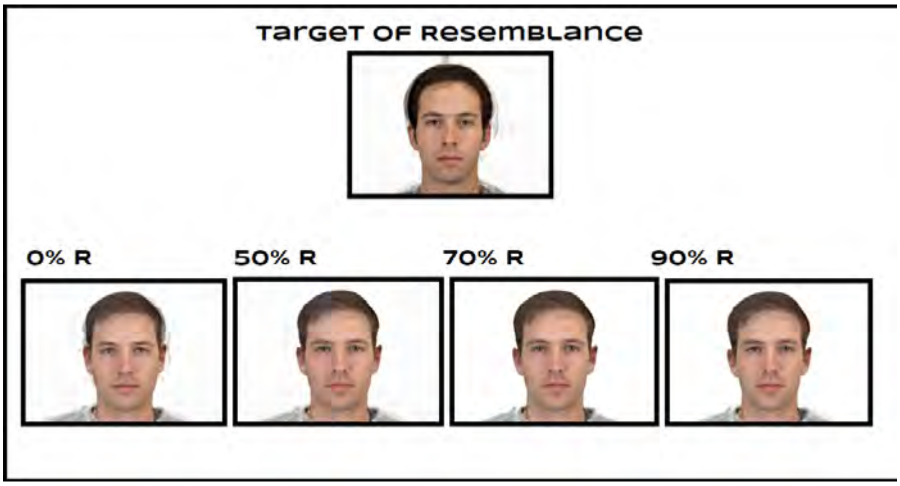

Figure 7: Nepotism Similarity (R) Manipulation Sequence

We recruited 150 participants, 115 male and 35 female from the Crowdflower platform between the ages of 19 and 65 (mean age 31.45 years). We randomly distributed them to 3 groups with varying strength levels of our manipulation of similarity:

- 50% increase in similarity (n=50),
- 70% increase (n=50),
- 90% increase (n=50)

Participants in each of the 3 groups were presented with the picture of the agent we previously catalogued as the boss agent. They were instructed that this individual is named John and in the following steps they will be asked a series of questions in related other people's likeness to him and that they should take the time to become familiarized with his appearance. Subsequently, they were presented with the images of candidate 1 and candidate 2 and were asked which of these individual looked like John. The participants responses were recorded on a standard 6-point response scale with 1-*Person 1 is very similar to John* and 6-*Person 2 is very similar to John*. Scores were then recoded so that higher scores would always denote successful detection of similarity.

Table 2: Nepotism Linear Regression and Logistic Regression Models

```
lm(NepotismQ~Condition+Sex+Age+Siblings+Youngest,data)
glm(NepotismBinary~Condition+Sex+Age+Siblings+Youngest,
family = binomial(link="logit"),data)
```

What we discovered was that participants, at all similarity manipulation strength levels, were able to select the correct response at an above chance level. By combining both the level of correct responses (see figure 8) and fitting our data to linear regression models and logistic regression, we determined that, to ensure our participants would be reliably influenced by the cues of relatedness resulting from our manipulation, we should use the 90% strength level of our manipulation. At this level, 66% of our participants choose the correct response (i.e., our manipulated candidate). In addition, both the linear and logistic models showed that resemblance strength level had an impact (see table 4 and 3 for full results of these models).

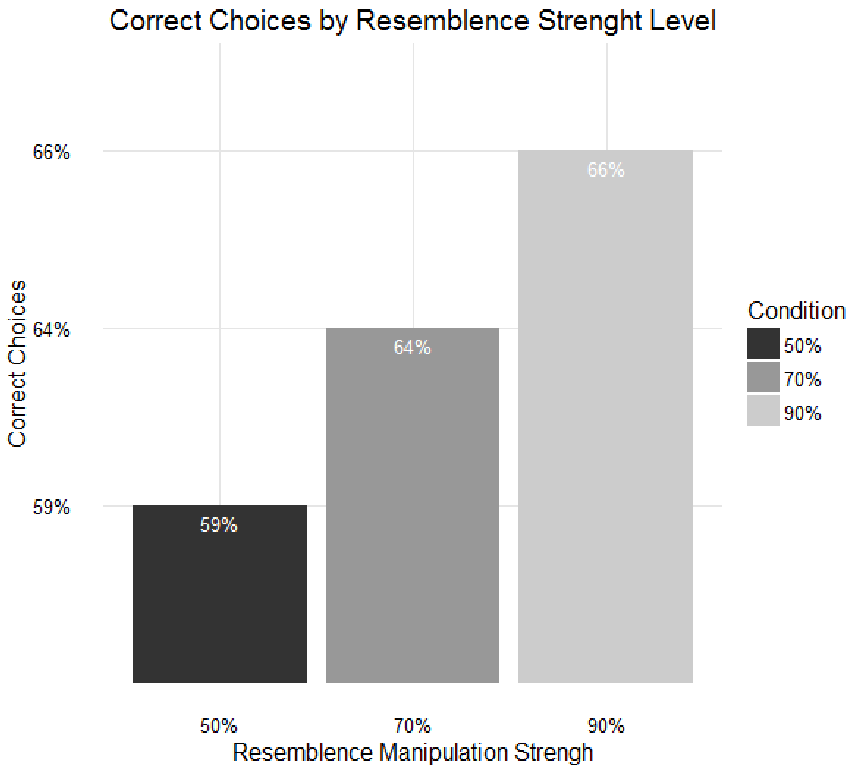

Figure 8: Nepotism Correct Choices by Similarity Threshold

Table 3: Pretest Nepotism Experiments - Linear Regression Models Results

	Dependent variable:		
	NepotismQ		
	(1)	(2)	(3)
Condition	0.011	0.013*	0.013*
	(0.007)	(0.007)	(0.007)
SexWomen		0.397	0.372
		(0.258)	(0.261)
Age		−0.014	−0.013
		(0.012)	(0.012)
Siblings			−0.001
			(0.076)
YoungestYes			0.289
			(0.401)
YoungestNo			0.059
			(0.406)
Constant	3.134***	3.348***	3.161***
	(0.470)	(0.588)	(0.672)
Observations	150	150	150
R^2	0.018	0.039	0.048
Adjusted R^2	0.011	0.020	0.008
Residual Std. Error	1.303 (df = 148)	1.297 (df = 146)	1.305 (df = 143)
F Statistic	2.657 (df = 1; 148)	1.993 (df = 3; 146)	1.193 (df = 6; 143)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 4: Pretest Nepotism Experiments - Logistic Regression Models Results

	Dependent variable:		
	NepotismBinary		
	(1)	(2)	(3)
Condition	0.008	0.012	0.013
	(0.010)	(0.011)	(0.011)
SexWomen		0.957**	0.907*
		(0.460)	(0.464)
Age		−0.008	−0.005
		(0.019)	(0.020)
Siblings			0.031
			(0.124)
YoungestYes			0.801
			(0.639)
YoungestNo			0.262
			(0.644)
Constant	−0.017	−0.266	−0.930
	(0.745)	(0.949)	(1.090)
Observations	150	150	150
Log Likelihood	−98.273	−95.892	−94.309
Akaike Inf. Crit.	200.546	199.784	202.617
Note:	*p<0.1; **p<0.05; ***p<0.01		

Additional Notes: (1), (2), (3) represent the different resemblance manipulation strength levels of 50%, 70% and 90%; Ordinary Least Squares (OLS) & logistic represent the two different types models used in our analysis the standard regression model and logistic regression model respectively.

6.3 vignette development

In the next couple of chapters, my objective will be to present all our studies individually. As the experimental framework in terms of stimuli presented has remained constant, this has allowed us to steadily improve upon the robustness of our datasets by testing our scenarios and stimuli under different constraints. These have involved creating and testing scenarios that vary by type and context while maintaining the overall narrative scope.

Most studies present participants with straightforward positive context in which the nature of the kin relationship between the presented agents remains unknown to the participants. This of course is a very worthwhile setup and in no way does our results undermine it. However, if one is to fully explore this topic, great care must be taken to verify the context-dependent nature of social action related to kinship. In terms of direct kinship recognition, this has long been investigated by taking into account the fact that kinship cues should decrease mating type behaviour directed towards close kin and, at the same time, increase cooperation towards kin.

Additionally, even cooperation itself as a social behaviour should be context-dependent, in the sense that individuals should not have an all encompassing preference towards kin irrespective of context. The valence of the action taken towards kin should always be paramount and actions should not be simple low-level perceptual matching bias (i.e., people choosing kin as the target of the action whether it be positive or negative). When it comes to third-party kinship recognition and the predictions and expectations that people have based on these cues, behaviours and decisions should remain context-dependent if the underlying mechanism is to be linked to an effect of kinship recognition and not low-level perceptual matching biases.

Table 5: Positive Scenario

John Smith is the manager of a small business and he needs to fill a position that has recently become vacant. He decides that he will promote a current member of his organization to this position. After extensive interviews, two possible candidates remain.

As such the scenarios we chose to present to our participants in our 11 studies varied according to their designation in terms of type (explicit kin information/ no kin information) and context (positive or negative actions directed towards candidates). I will not delve deeper into the specifics of these variations as I will present each individual study and their respective scenario in the following chapters.

Table 6: Negative Scenario

John Smith is the manager of a small business and he needs to let go of one his employees. Out of all his employees, two are currently occupying a position where it has recently become infeasible to keep both. After extensive reviews, he is ready to make a decision.

6.4 PERFORMANCE REVIEW SHEETS

Another major component of our experimental procedure involved presenting our participants two review sheets that would be associated with the candidates in our scenarios (see table 7).

Table 7: Performance Review Sheets

Higher Performance Sheet

Performance Review	
Job Knowledge	79
Productivity	84
Work Quality	82
Technical Skills	81
Work Consistency	86
Enthusiasm	77
Cooperation	75
Attitude	75
Initiative	72
Work Relations	85
Creativity	85
Punctuality	78
Attendance	85
Dependability	86
Communication Skills	85
Overall Rating	81

Lower Performance Sheet

Performance Review	
Job Knowledge	79
Productivity	86
Work Quality	71
Technical Skills	81
Work Consistency	79
Enthusiasm	71
Cooperation	70
Attitude	78
Initiative	71
Work Relations	74
Creativity	83
Punctuality	74
Attendance	87
Dependability	79
Communication Skills	87
Overall Rating	78

The main reason for their inclusion in our experimental framework was to allow us to capture the effect of our manipulation of similarity. By using a between-subject design where the only difference would be that one group of participants would have the higher performance sheet associated with the similar candidate we are able to clearly identify strictly the effect of our manipulation as the difference between the predictions of these two groups. The review sheets were created by randomly generating the individual values in such as way as for overall score (i.e., average score) would be equal to the predefined amount. In the case of higher performance sheet 81, while for other 78. The reasoning behind this choice of overall scores is the fact that we wanted these values to be as close as possible while remaining distinct. To make this distinction more prominent we choose to situate the 3 point difference in overall score at the junction between 70 and 80, we create the illusion that in fact the difference is greater when it was relatively insignificant. The categories chosen here are typical of job performance reviews.Our intention is for the participants to be drawn towards the more performant candidate with this preference being mediate by the effect our similarity manipulation.

6.5 EXPERIMENTAL STUDIES

I will first present the studies in which the participants are not informed as to the kin relationship of our agents. We begin with the four studies that investigate this issue through a positive context, in which participants are asked to predict which candidate the boss agent will select to be promoted. Continuing, we will tackle the issue of context-dependent responses in kin-recognition effects by presenting the four studies in which our participants are asked to predict which candidate will be fired by the boss agent. This contrast in scenarios will allow us to verify the context-dependent nature of our participants' responses. Finally, we will switch to our second type of experiments, that explicitly offer the participants the full information related to the kin relationship of the agents in with both our positive and negative context scenarios.

6.5.1 *Study 1*

6.5.1.1 *Introduction*

As people seem to cooperate more with people who look like them, does this inclination impact the predictions people make when observing third-parties interact? Even though limited in breadth, research on this topic has provided some evidence that indeed people seem to have these expectations (Wang et al., 2016). One natural interpretation would be that this is an effect of third-party kin recognition coupled with the expectation that kin cooperate. By moving this topic towards a work context and utilizing the techniques and procedures used in studies investigating the effect of self-resemblance, we wished to bring more clarity to this topic and provide a framework in which future studies might be produced. For this study, we recruited 140 participants age 18 to 59 years (M=29.69 years), having on average 2.4 siblings, 41% being the younger sibling, through the online Crowdflower platform. This platform provides some of the same features available through MTurk with the addition of being directly opened to researchers outside of the US.

6.5.1.2 *Methods*

As in other crowdsourcing environment such as MTurk, potential participants were presented with a "work order" for which they were remunerated after the study. Our study presented participants with the opportunity to participate in a short questionnaire that investigated decision making and moral judgments. They were informed that they would receive 0.30 euro for their participation. In terms of remuneration per time spent, this was above average, and the exit survey showed most of our participants were satisfied with the requirements and pay of the task (i.e., mean contributor satisfaction = 4.2/5 and pay satisfaction= 4.4/5; data provided by Crowdflower exit survey). After accepting the task participants were provided with a link to the Qualtrics survey. Steps were taken to prevent participants from responding multiple times to the survey (i.e., filtering through Qualtrics repeated IPs of participants), both on our end and through security measures of the Crowdflower platform itself.

In the first screen participants were presented with all 3 pictures of the agents that were to be part of the scenario. Without further context, we asked the participants whether "these people remind you of someone?". They were allowed to select any of the 3 pictures or none at all. This step ensured that participants would pay close attention to the faces, which is necessary for our protocol.

All participants saw the picture of John and the scenario in which they were told that he is the manager of a local business and he needs to promote one of his employees. We informed them that an extensive review of his employees had been done and resulted in only two possible candidates remaining. At which point we presented them with pictures of both of these candidates.

To reinforce the narrative, participants saw the performance review sheets assigned to each candidate. Participants were randomly distributed into one of two groups, as per our between subject design . Half of our participants (n=70) were randomly assigned to the condition in which the similar candidate (i.e., Candidate I) was assigned the performance review sheet with the higher overall score, 81 vs. 78 overall score of dissimilar candidate. The rest of our participants (n=70) were assigned to the other group, which had the higher overall score performance sheet assigned to the dissimilar candidate (i.e., Candidate 2).

All participants were asked the same experimental question: "Who do you think John Smith is going to choose?" Their responses were recorded on a standard 6-point response scale with 1-*Most certainly Candidate 1* and 6-*Most certainly Candidate 2*. In addition, they were asked to justify their choice "Why do you believe that John Smith will choose this candidate?" and as an attention check "Which candidate had the higher overall rating in his performance review sheet?".

6.5.1.3 *Results*

Data was recoded to facilitate model fitting. The 6-point scale capturing our dependent variable was recoded so that higher scores would always reflect the selection of the similar candidate. Then, two steps were taken which together allow to analyze the preference for the similar candidate while controlling for the response sheet assigned to this candidate.

First, responses were coded -3, -2, -1, +1, +2, and +3. That is, the value zero was omitted. Second, we created a variable named 81Sheet, which took the value +1 when the similar candidate received the best performance sheet, and the value -1 when the similar candidate received the worse performance sheet.

Overall, this means that when we fit the following model, the effect of similarity is captured by *the intercept term*. Positive (resp., negative) values of the intercept terms denote a preference for the similar (resp., dissimilar) candidate independently of the performance sheet assigned to this candidate. Accordingly, we can determine if there is a preference for the similar candidate if the 95% confidence interval for the intercept term does not include the value zero.

Table 8: Linear Regression Model Study 1

```
lm(ChoiceSimilar ~ 81Sheet  + Age + Youngest, data = NepPosA)
```

As displayed in table 9, the regression model (R^2=0.429, F(3,136)=34.4, $p<0.01$) shows that the sheet assigned to the similar candidate strongly affects the preference for this candidate, which is unsurprising: all other things being equal, participants expect the candidate with the best performance to be promoted. More importantly, the intercept term is positive but not significantly different from zero, with a 95% CI of [-0.59; 1.76].

While we could have ended our investigation with this null result, we opted to conduct enough studies (and variations on these studies) to allow for a meta analysis. Accordingly, we will postpone the interpretation of our results until after we have presented our full set of studies and their meta-analysis.

Table 9: Linear Regression Model Study 1 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.589
	(0.602)
PerformanceSimilar	1.623^{***}
	(0.160)
Age	-0.018
	(0.019)
YoungestNo	0.108
	(0.326)
Observations	140
R^2	0.432
Adjusted R^2	0.419
Residual Std. Error	1.880 (df = 136)
F Statistic	34.430^{***} (df = 3; 136)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

6.5.2 *Study 2*

6.5.2.1 *Methods*

Study 2 is the natural continuation of our first study. By expanding our subject pool (most notably, to female participants), we hoped that we would be able to produce a clearer picture of what effect, if any, our manipulation of similarity had on our participants. Only having male participants in our first study was a clear weakness that needed to be corrected. In addition, replicating our study allows for a more precise estimation of our effect size.

As with all of our studies, 163 participants were recruited through the Crowdflower platform, ages 19-74 years (M=32.28), 123 Men and 40 Women out of which 41% were the youngest sibling. Participants were asked to participate in a study related to decision making and moral judgments being remunerated with 0.30 euro after completing the task. The participant overall satisfaction and pay satisfaction remained in line with our first study.

Participants were taken through the same steps outlined in the previous study before being presented with the boss, scenario, candidates and their respective review sheets. The review sheets assignment again determined the two groups to which the participants were randomly assigned. The experimental question remained "Who do you think John Smith is going to choose?" their answers being recorded on a 6-point response scale with 1-*Most certainly Candidate 1* and 6-*Most certainly Candidate 2*. In addition, they were asked to justify their choice and indicate which of the candidates had the higher overall performance review score.

6.5.2.2 *Results*

We recoded our data using the same procedure as seen our previous experiment. As a result our independent variable, was recoded into positive/negative values, so that the 81Sheet variable was coded +1 when the similar candidate received the best performance sheet, and the value -1 when the similar candidate received the worse performance sheet. The dependent variable, ChoiceSimilar, codes the dissimilar candidate denoted by *strictly negative* values (-1, -2, and -3) and the similar candidate denoted by *strictly positive values* (+1, +2, +3). Our interest remains in the the effect of similarity captured by *the intercept term.* of our model.

In our model (R^2=0.380, F(4,158)=24.22, p<0.01), as expected again the major driving force behind our participants' choices remains how

Table 10: Linear Regression Model Study 2

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepPosB)
```

performant the similar candidate is ($p<0.01$). On the other hand, the strength of our similarity manipulation effect has increased with an intercept value of 0.99, 95% CI [-0.06, 2.06]. As outlined in the previous experiment, a positive value of the model's intercept means that the similar candidate is preferred, independently of his performance sheet.

Table 11: Linear Regression Model Study 2 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.997^{*}
	(0.544)
PerformanceSimilar	1.427^{***}
	(0.148)
SexWomen	0.096
	(0.350)
Age	-0.020
	(0.016)
YoungestNo	-0.067
	(0.304)
Observations	163
R^2	0.380
Adjusted R^2	0.364
Residual Std. Error	1.875 (df = 158)
F Statistic	24.229^{***} (df = 4; 158)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

The fact that our desired effect has maintained the expected direction in addition to "becoming" marginally significant ($p<0.1$) would indicate that in fact our manipulation of similarity we implement in these studies does have an impact on the predictions and expectation our participants make in the context of our experiment. Nevertheless, given the discrepancy in terms of significance between our studies so far the wisest course of action is to reinforce these findings by conducting a direct replication. Thus, our goal in Study 3 will remain the same as Study 2, in that it will provide us with stronger dataset possibly reinforcing our findings thus far.

6.5.3 *Study 3*

6.5.3.1 *Methods*

For this replication study we recruited 199 participants, ages 18-65 years (M=31.88), 145 Male and 54 Female, 31% being the youngest sibling. As we had success in our previous studies in recruiting and maintaining our participants overall satisfaction with the task, we kept the remuneration and general Crowdflower recruitment process identical to our previous studies. In terms of experimental procedure, this study was identical to our previous study as per our objective of replicating our results.

6.5.3.2 *Results*

We maintained the recoding procedure from the previous experiments and our independent variable was recoded into positive/negative values (i.e., 81Sheet coded as +1 when the similar candidate received the 81-point sheet). While our dependent variable was recoded as *strictly negative* values (-1, -2, and -3) when participants are choosing the dissimilar candidate and when the similar candidate is chosen this is denoted by *strictly positive values* (+1, +2, +3).

Table 12: Linear Regression Model Study 3

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepPosC)
```

Our model ($R^2=0.503$, $F(4,194)=49.09$, $p<0.01$) again shows that, as we expected, the major driving force behind our participants' choices remains how performant the similar candidate is ($p<0.01$). When our participants are not in the group in which the similar candidate is assigned the higher performance sheet they are reliably choosing our dissimilar candidate, while opposite remains true of the group in which the similar candidate does have the higher performance sheet assigned to him .

Table 13: Linear Regression Model Study 3 - Results

	Dependent variable:
	ChoiceSimilar
Constant	-0.173
	(0.505)
PerformanceSimilar	1.701^{***}
	(0.121)
SexWomen	0.249
	(0.275)
Age	-0.0002
	(0.015)
YoungestNo	0.040
	(0.262)
Observations	199
R^2	0.503
Adjusted R^2	0.493
Residual Std. Error	1.709 (df = 194)
F Statistic	49.099^{***} (df = 4; 194)
Note:	*p<0.1; **p<0.05; ***p<0.01

However, as opposed to our other studies, here we are presented with a major departure in terms of the effect of our similarity manipulation. With an intercept value of -0.173, 95% CI [-1.16, 0.81], we see that the direction of our effect has been flipped, in addition to the intercept not being significant (p>0.1). Once more, we defer the interpretation of the results until after we will report a meta-analysis of all our studies.

One methodological concern with our protocol, though, is that similarity effects might be suppressed by the strong effect of the performance sheets. That is, participants may zero in on the difference in performance to such an extent that they only use this cue in order to make their prediction. If this is correct to some extent, then we may have a better chance to detect similarity effects by narrowing the difference in performance between the two candidates. This is what we tried in the next experiment.

6.5.4 *Study 4*

6.5.4.1 *Methods*

To investigate the possibility that the performance cues we present through our review sheets might be suppressing a similarity effect, we equalized the two performance sheets overall score while maintaining differences in individual scores present on the sheets (see table 14). Participants were still randomly assigned into two groups (corresponding to the sheet assigned to the similar candidate), only now this should not affect the predictions our participants make.

Table 14: Equal Performance Review Sheets

Performance Review	Performance Sheet 1	Performance Sheet 2
Job Knowledge	75	74
Productivity	79	72
Work Quality	74	72
Technical Skills	81	77
Work Consistency	76	74
Enthusiasm	78	76
Cooperation	75	76
Attitude	71	73
Initiative	73	78
Work Relations	72	73
Creativity	73	75
Punctuality	72	73
Attendance	73	75
Dependability	76	81
Communication Skills	77	76
Overall Rating	**75**	**75**

We recruited 200 participants, ages 18-65 years (M=31.38), 154 Male and 46 Female, 36% being the youngest sibling. We again kept the remuneration and general Crowdflower recruitment process identical to our previous studies. As before, participants were randomly distributed into one of two groups based on the performance review sheet assigned to the candidates. An important note to make here is that our attention check question reveals that 61% of our participants from both groups were aware that the review sheets were in fact equal.

6.5.4.2 *Results*

Preliminary analysis of the data revealed a surprising turn of events, showing that even though our review sheets were purposely designed to be equal in terms of overall score and a majority of our participants (61%) were aware of this fact, they ultimately do affect our participants predictions ($p<0.001$). As such we recoded the data as per the procedure used in our previous experiments, though we replaced the 81Sheet designation of our independent variable with 75Sheet, which takes the value +1 when the similar candidate is assigned Sheet 1 (i.e., the sheet that our analysis revealed to be the "best" one). Our linear regression model remained the same.

Table 15: Linear Regression Model Study 4

```
lm(ChoiceSimilar ~ 75Sheet + Sex + Age + Youngest, data=NepPosD)
```

As mentioned, within this model we would have expected that the review sheet the participant see attached to our candidates should not have a significant impact on their ultimate prediction. Nevertheless, the results of our model (R^2 =0.155, $F(4,195)$=8.95, $p<0.01$) reinforces our preliminary analysis revealing that again the primary force influencing our participants predictions remains performance($p<0.01$).

Interpreting this result in isolation is difficult as there are a number of possible explanations as to why we see this equal variant in fact mimicking the results of our other studies. One explanation might be that even though the review sheets were equal in terms of overall score, participants latched on to one or more of the individual traits that were not equal between them and based their performance analysis on them. What is clear is that it seems participant's main objective in creating their predictions is analyzing cues related to performance and when the main cue (i.e., overall score) is equal they use other values to determine which of the candidate is more performant. In this case they seem to be drawn towards Sheet 1 and subsequently predicting that the boss will promote the candidate that is assigned this

Table 16: Linear Regression Model Study 4 - Results

	Dependent variable:
	ChoiceSimilar
Constant	-0.599
	(0.561)
PerformanceSimilar	0.805***
	(0.144)
SexWomen	0.158
	(0.341)
Age	0.016
	(0.017)
YoungestNo	0.244
	(0.302)
Observations	200
R^2	0.155
Adjusted R^2	0.138
Residual Std. Error	2.016 (df = 195)
F Statistic	8.957*** (df = 4; 195)
Note:	*p<0.1; **p<0.05; ***p<0.01

sheet. Given the work context of our experiments participants propensity to search for meaningful cues of performance is a straightforward efficient route to take in producing one's predictions. Ultimately our equal variant becomes another positive context scenario like our previous experiments, in which this pseudo-performance plays the major role in their predictions. Given this, we move to determine if in this context our manipulation of similarity had any effect. The intercept value in our model, -0.599 95% CI [-1.69, 0.50], continues the trend of contradictory results by maintaining negative direction (i.e., compared to Study 1 and 2) and not meeting a significance threshold (i.e., compared to Study 2).

At this point it has become clear that a meta-analysis of our studies is needed if we are to create a clear picture and see whether or not our manipulation of similarity plays a role in directing our participants predictions. We discussed previously that for our manipulation of similarity to represent an interaction between facial stimuli and an inbuilt mechanism of third-party kin recognition the results of this interaction (i.e., the direction of our effect) should be context-dependent. Thus, if we need to produce a meta-analysis of our studies and we have this requirement of context-dependents, our next step must be to create a series of studies that increase the number of data points for our meta-analysis and provide a negative context for our scenarios. I will now present this series of experiments.

6.5.5 *Study 5*

6.5.5.1 *Methods*

The major departure in this series of experiments comes from the modifications made to the scenarios presented to our participants. In this study and the following ones using this negative context, participants are instructed that John is the manager of a local business and certain circumstances have made it infeasible to keep two of his employees that occupy a similar position, so it is necessary that he should fire one of them. For this study we recruited 200 participants, ages 18-60 years (M=31.06), 159 Male and 41 Female, 36% being the youngest sibling. We again kept the remuneration and general Crowdflower recruitment process identical to our previous studies. The experimental procedure remained the same. Participants were primed by presenting the pictures of the 3 agents in our scenario asking the participants if they reminded them of anyone. They were presented with the modified scenario in which they we informed them that extensive reviews were undertaken and subsequently John needs to make a decision. After being presented with the scenarios they were randomly distributed to one of the two groups as with all of our other experiments. The review sheets assignment determining the two groups to which the participants were randomly assigned, this is reflected in the 81Sheet variable which was coded +1 when the similar candidate received the 81-point sheet and -1 when the dissimilar candidate received the sheet with the higher overall score. After which they were asked the main experimental question: "Who do you think John Smith is going to choose?". Their answers being recorded on a 6-point response scale with 1-*Most certainly Candidate 1* and 6-*Most certainly Candidate 2*. In addition, the same justification and attention check questions were asked following the experimental question.

Note that we decided to maintain the same phrasing as in the previous studies, by asking: *"Who do you think John Smith is going to choose?"* While this increases consistency between studies, a potential drawback is that the question might be ambiguous. It is meant as "who will John choose *to fire*", but the verb "choose" can also denote a positive choice, here, a choice of who is going to stay. this ambiguity will be addressed in the next studies.

6.5.5.2 *Results*

Recoding our data was done using the same procedures used in the positive context experiments. This was done to maintain the consistency between these two different contexts. The choices our participants made (i.e., our dependent variable) were coded so as the higher values represent our participants' increased certainty that John would choose the similar candidate in lieu of the dissimilar candidate irrespective of condition. Responses were coded -3, -2, -1, +1, +2, and +3. The choice in this context meant that this candidate will be fired. We would expect that the similarity manipulation would have a negative impact on the predictions our participants have as opposed to the positive impact in the previous studies this effect being captured by *the intercept term.* of our model. Additionally, as with our other studies, our independent variable was recoded into positive/negative values (i.e., 81Sheet -1 dissimilar candidate 81-point sheet; +1 similar candidate 81-point sheet). Only 16 participants (8%) were unable to respond correctly to the attention check.

Table 17: Linear Regression Model Study 5

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepNegA)
```

As one might expect by now, fitting the data to our model (R^2=0.089, F(4,195)= 4.74, p<0.01) revealed that performance had a significant impact on the predictions our participants made (p>0.001). On the other hand analysis of the effect of our similarity manipulation had a very interesting result. The intercept value, 1.107; 95% CI [-0.17, 2.39] in our model revealed that the effect of our manipulation was marginally significant (p<0.1). However, the more interesting aspect of this result is the direction of this effect. It seems that our similarity manipulation increases the likelihood that our participants will predict the boss firing the similar candidate.

This of course presents a problem in terms of the context-dependent nature that this effect should have, if it is to be linked to a kin-recognition mechanism. As a first indication of a lack of a context-dependent effect, this result reinforces the need gather more data and in the end produce a meta-analysis of all our data. However, to achieve this our first step is for our next study to achieve a direct replication that also tries to correct the methodological issue present in this study. That is to say, we will modify the experimental question to remove any ambiguity present here.

Table 18: Linear Regression Model Study 5 - Results

	Dependent variable:
	ChoiceSimilar
Constant	1.107*
	(0.655)
PerformanceSimilar	0.609***
	(0.166)
SexWomen	0.385
	(0.416)
Age	−0.022
	(0.020)
YoungestNo	−0.506
	(0.347)
Observations	200
R^2	0.089
Adjusted R^2	0.070
Residual Std. Error	2.336 (df = 195)
F Statistic	4.747*** (df = 4; 195)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.5.6 *Study 6*

6.5.6.1 *Methods*

To correct some of the methodological weakness that might have surfaced in addition to replicating the results of our previous study, we reworded our experimental question. Where participants were asked "Who do you think John Smith is going to choose?", we added "choose to fire". This simple solution removed any of the possible ambiguity that the question posed. We recruited 202 participants from the Crowdflower platform, ages 18-58 years (M=31.90), 151 Male and 51 Female, 42% being the youngest sibling. Following the same exact procedure the modified scenario was presented and they were informed that extensive reviews were undertaken and subsequently John needed to make a decision.

After being presented with the scenarios they were randomly distributed to one of the two groups as with all of our other experiments. The review sheets determining the two groups to which the participants were assigned. As noted, they were presented with the modified main experimental question: "Who do you think John Smith is going to choose to fire?". Their answers being recorded on a 6-point response scale with 1-*Most certainly Candidate 1* and 6-*Most certainly Candidate 2*. On the attention check question only 20 participants did not respond correctly (9.9%).

6.5.6.2 *Results.*

Fitting the data to our model (R^2=0.022, F(4,197)=1.11), showed the main force impacting the predictions our participants made was performance (p>0.001).

Table 19: Linear Regression Model Study 6

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepNegB)
```

In the case of our manipulation effect, the intercept value, 0.372; 95% CI [-0.90, 1.65], reveals the same direction of the manipulation effect as our previous experiment, which does not meet requirement of being context-dependent. As opposed to our previous study these results were not significant (p>0.1). Additionally, the rewording we implemented in this experiment did not affect the direction of our effect, as such the results of our previous study cannot be solely attributed to ambiguity present in our experimental question.

While the continued direction of the effect of our manipulation in this study brings strength to the interpretation that perhaps we are not in fact dealing with a context-dependent kin recognition response, the

Table 20: Linear Regression Model Study 6 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.372
	(0.653)
PerformanceSimilar	-0.327^{**}
	(0.163)
SexWomen	0.004
	(0.376)
Age	0.001
	(0.019)
YoungestNo	-0.213
	(0.329)
Observations	202
R^2	0.022
Adjusted R^2	0.002
Residual Std. Error	2.308 (df = 197)
F Statistic	1.119 (df = 4; 197)
Note:	*p<0.1; **p<0.05; ***p<0.01

fact that the effect itself is not significant requires us to be cautious and hold off on expanding on these results here. Hopefully our next studies will offer us a better picture. To that end, our next step will be to replicate the results of this study using the new modified experimental question.

6.5.7 *Study 7*

6.5.7.1 *Methods*

As a direct replication of the previous study, all procedures and techniques were identical. 199 participants were recruited through the Crowdflower platform, ages 18-65 years (M=31.66), 147 Male and 51 Female, 32% being the youngest sibling. They were randomly distributed to the two groups based on which candidate was assigned the 81-point Sheet. Only 2% were unable to correctly remember which candidate had the higher overall score on their respective review sheet.

6.5.7.2 *Results*

Fitting the data (R^2=0.092, $F(4,193)$=4.87, $p<0.01$) to our model showed the main force impacting the predictions our participants made was performance ($p<0.001$).The intercept value, 1.069; 95% CI [-0.15, 2.29], shows us that the significant effect of our manipulation in the same direction as seen in our previous experiment.

Table 21: Linear Regression Model Study 7

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepNegC)
```

This consistent trend we see in all of our negative context experiments strongly reinforce the need to analyse this issue through the lens of a meta-analysis and bring together the full weight of our studies. This might bring to light if indeed our manipulation of similarity is having an effect and if this effect is context dependent. As such the following studies are fully geared towards precisely this goal.

Table 22: Linear Regression Model Study 7 - Results

	Dependent variable:
	ChoiceSimilar
Constant	1.069^{*}
	(0.626)
PerformanceSimilar	0.678^{***}
	(0.167)
SexWomen	0.287
	(0.382)
Age	-0.025
	(0.018)
YoungestNo	-0.318
	(0.359)
Observations	198
R^2	0.092
Adjusted R^2	0.073
Residual Std. Error	2.342 (df = 193)
F Statistic	4.878^{***} (df = 4; 193)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.5.8 *Study 8*

6.5.8.1 *Methods*

Study 4 was designed to use equal performance sheets so as to exclude possibility that our participants preference for performance was obfuscating the effect of our similarity manipulation. The results in that experiment showed that even though the overall score was equal on both review sheets, perhaps participants still used some of the individual values to determine that one of those sheets was "superior" mimicking the results of our previous positive context studies. To reinforce our ability to reach this interpretation, this study is designed to use the same performance review sheets as Study 4 in addition to the same type of scenario as the other negative context studies. For this investigation we recruited through the Crowdflower platform 198 participants, ages 19-57 years (M=31.71), 151 Male and 45 Female, 33% being the youngest sibling. Remuneration and general Crowdflower recruitment process identical to our previous studies. Participants were randomly distributed into one of two groups based on the performance review sheet assigned to the candidates. 57% of our participants from both groups were aware that the review sheets were in fact equal or at least they were able to indicate as such through the attention check question.

6.5.8.2 *Results*

Fitting the data to our model (R^2=0.048, F(4,191)=2.39, $p<0.01$) showed the main force impacting the predictions our participants made was performance ($p<0.001$). However, analyzing our data revealed that in fact in this study participants flipped their review sheet preference. In Study 4, participants seem to be drawn towards Sheet 1, however in this experiment we see that they predict that the boss will keep the candidate that is assigned Sheet 2. In addition, we see that the effect of our manipulation, intercept=-0.131; 95% CI [-1.40, 1.14] , is no longer significant and switches direction. However, given the nature of the intercept value one would be hard pressed to give any meaningful interpretation other than, in this case, our manipulation seems to have no effect.

Table 23: Linear Regression Model Study 8

```
lm(ChoiceSimilar ~ 81Sheet + Sex + Age + Youngest, data=NepNegD)
```

Our final three studies in this series will be tackling the issue of whether or not our participants will take into account information related to kinship if this is present in an explicit form. One interpretation would be, given our results so far, that our manipulation of similarity is too subtle or that in fact participants are not attentive towards cues of kinship as this information does not generally have an impact on their predictions at least not when this information is not at an explicit level. As such our next studies will investigate this issue by presenting our participants with this information.

Table 24: Linear Regression Model Study 8 - Results

	Dependent variable:
	ChoiceSimilar
Constant	−0.131
	(0.650)
PerformanceSimilar	−0.452***
	(0.156)
SexWomen	0.151
	(0.369)
Age	0.005
	(0.019)
YoungestNo	0.203
	(0.331)
Observations	196
R^2	0.048
Adjusted R^2	0.028
Residual Std. Error	2.166 (df = 191)
F Statistic	2.396* (df = 4; 191)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.5.9 *Study 9*

6.5.9.1 *Methods*

This study will be first in a series of three investigations in which we modify our experimental procedure by presenting our participants with varying degrees of information related to the kinship relationships between the agents in our scenario. In this study participants are only presented with partial information being informed that one of the candidates is part of the boss's family. Would this information be enough to influence our participants in making use of the facial cues related to resemblance to inform their predictions? Our pretest clearly showed that this information is accessible to them, in that they were able to determine which of the faces we present were similar to each other at an above chance level.

We used Crowdflower to recruit another pool of participants, 203 participants, ages 18-61 years (M=31.15), 153 Male and 50 Female, 34% being the youngest sibling. They were randomly distributed to the two groups as before based on the review sheet the saw assigned to the two candidates after being presented with scenario, now having the additional information that one of the candidates was a member of John's family.

6.5.9.2 *Results*

Our model (R^2=0.452, F(4,198)=40.79, $p<0.01$) showed that performance played an important role in directing our participants predictions ($p<0.001$). Given the fact that our scenario presented participants with partial kin information, one might expect that participants would be more attentive to the cues of similarity we manipulated and subsequently this would increase the strength of this effect. Analysing the intercept value in our linear regression model, -0.069; 95% CI [-1.40, 1.14], reveals somewhat surprisingly, this is not in fact the case ($p>0.1$). Thus, even though our participants had access to this additional information, our similarity manipulation still fails to be a significant factor in their predictions.

Table 25: Linear Regression Model Study 9

```
lm(ChoiceSimilar ~ 81Sheet+Sex+Age+Youngest, data=NepPosExpA)
```

One possible critique of this experimental design would be that even with this partial kinship information, participants are not able to connect this information to the cues of resemblance present in the facial stimuli. That is to say, they know that one of them is related to the boss and they know that one of them resembles the boss, nevertheless these two bits of data are not connected either because cues of resemblance do not inform kinship relationships or that information related to kinship relationships is irrelevant in the mechanisms responsible for these predictions.

Table 26: Linear Regression Model Study 9 - Results

	Dependent variable:
	ChoiceSimilar
Constant	−0.069
	(0.497)
PerformanceSimilar	1.568***
	(0.126)
SexWomen	0.169
	(0.297)
Age	0.013
	(0.015)
YoungestNo	−0.288
	(0.266)
Observations	203
R^2	0.452
Adjusted R^2	0.441
Residual Std. Error	1.786 (df = 198)
F Statistic	40.793*** (df = 4; 198)
Note:	*p<0.1; **p<0.05; ***p<0.01

To address this issue our following experiments will present explicit and complete information on the kin relationships of our agents in both positive and negative contexts. Allowing us to address this issue and see whether or not our participants predictions are informed by the kinship relationship of the agents

6.5.10 *Study 10*

6.5.10.1 *Methods*

As discussed, the major departure in both this study and our last one in this series, we will explicitly tell our participants that candidate 1 is a member of the boss's family. In every other regards this study is identical to our other studies. The 197 participants, ages 18-58 years (M=33.4), 139 Male and 58 Female, 39% being the youngest sibling went through the same stages as our other participants. They were presented with the scenario in which they were told that John needs to promote one of his employees and after extensive reviews he is left with too choices. The major difference came in the form of sniped under the scenario and John's picture which informed the participants that "EMPLOYEE 1 IS A MEMBER OF JOHN'S FAMILY."

They were randomly distributed as before; however, in this case the fact that participants had access to the full kin information slightly modifies the basis of our experiment. While our similarity manipulation remains, this new information reinforces it and in some way completely replaces it by making it, in a way, irrelevant as the participants know (i.e., in the full sense of the word) that Candidate 1 is a relative of John. That is to say that if there is an effect of kinship we cannot disentangle if it's origin is the explicit information, the similarity manipulation or a combination of them. Nevertheless, this is not a weakness of our design, but it fact a strength as our goal here is to see if kinship information plays a role irrespective of the shape it comes in.

Moving on, in this study we used the positive context, as such our experimental question was: "Who do you think John Smith is going to choose?". The participants answers being recorded on a 6-point response scale with 1-*Most certainly Candidate 1* and 6-*Most certainly Candidate 2*. The same attention and justification questions as before were posed at this point.

6.5.10.2 *Results*

Fitting the data to our model (R^2=0.354, F(4,192)=26.28, p<0.01) showed what we have come to expect that the performance played a key role in informing our participants predictions (p<0.001). As mentioned, this experiment provides full kin information as such the value of our intercept in this model does not only encompass our manipulation of similarity, but also the impact of the explicit information provided, both acting in the same direction. Thus, it is rather interesting that we do not have a significant effect of kinship, intercept= 0.526; 95% CI [-0.52, 1.57].

Table 27: Linear Regression Model Study 10

```
lm(ChoiceSimilar ~ 81Sheet+Sex+Age+Youngest, data=NepPosExpB)
```

Participants seem to not take into consideration this information or if they do we are not able to detect it. It seems like our findings in the previous studies are vindicated in part by the results of this experimental design. I will now present our last experiment in this series, in which our goals will be to further reinforce this finding and show its robustness by using the negative context design in which participants will be asked to predict which candidate John will fire.

Table 28: Linear Regression Model Study 10 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.526
	(0.537)
PerformanceSimilar	1.438***
	(0.141)
SexWomen	0.234
	(0.317)
Age	0.004
	(0.016)
YoungestNo	−0.236
	(0.292)
Observations	197
R^2	0.354
Adjusted R^2	0.340
Residual Std. Error	1.983 (df = 192)
F Statistic	26.282*** (df = 4; 192)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.5.11 *Study 11*

6.5.11.1 *Methods*

Using Crowdflower we recruited 394 participants, ages 18-69 years (M=32.9), 279 Male and 115 Female, 39% being the youngest sibling. The experimental design here was identical to our last experiment. The only difference being that in this case we used the negative context scenario and as such participants were asked who they thought John would fire. Again, they were offered the full information related to the kin relationships of our agents.

6.5.11.2 *Results*

The results of our linear regression model (R^2=0.023, $F(4,389)$=2.29, $p<0.1$) showed that performance remained the primary force informing the participants predictions ($p<0.001$). In regards to kinship, intercept= 0.677; 95% CI [-0.23, 1.59], we replicated the results of our previous study, the effect maintaining its direction and remaining statistically insignificant ($p>0.1$).

Table 29: Linear Regression Model Study 11

```
lm(ChoiceSimilar ~ 81Sheet+Sex+Age+Youngest, data=NepNegExpA)
```

Ignoring for the moment that this result was not significant, the direction remaining positive would again point towards this effect not being context-dependent. This means that it's connexion to kinship would be highly doubtful.The sometimes contradictory and varying results we have gathered in this series of experiments must be cleared by analyzing all our studies in a meta-analysis. Even without it, what is clear is that manipulations of similarity, when dealing with third-parties and not using self-resemblance, are not equivalent. Cues of kinship play an important role in determining an individual's action when they are a result of direct kin recognition. However, it seems that this mechanism does not extend to third-party cues. Even though participants are able to detect similarity in third-parties this information is not processed in the same way as cues of self-resemblance and as such do not lead to context-dependent responses. Our meta-analysis will allow us to see if given this, our manipulation of similarity has an effect with any semblance of stability.

Table 30: Linear Regression Model Study 11 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.677
	(0.467)
PerformanceSimilar	0.330***
	(0.120)
SexWomen	0.216
	(0.272)
Age	−0.018
	(0.013)
YoungestNo	−0.048
	(0.247)
Observations	394
R^2	0.023
Adjusted R^2	0.013
Residual Std. Error	2.377 (df = 389)
F Statistic	2.294* (df = 4; 389)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.6 meta-analysis of nepotism studies

Within the social sciences experimental research has increasingly relied upon meta-analysis as framework to analyze a large body of studies (Vemer et al., 1989; Waldorf et al., 2005; Amato et al., 1991; Wagner et al., 2006). This type of framework allows greater systematization in the way previous literature is analyzed and used to contribute to an ever growing body of work in a field. Traditionally, this type of process begins with creating a set of requirements for inclusion of a study and a meticulous process of recoding previous studies for the inclusion in the eventual analysis. However, in our case this process is wholly simplified by the fact that we are analyzing our own body of work. All studies used the same experimental design, posed the same experimental question and are interested in the same thing, the constant of the models used to detect the effect of our manipulation. The constant or intercept in our model captures whether the similar candidate is preferred, independently of his performance sheet. This systematic approach we took in analyzing our data allows us to easily transfer our findings in each individual study to our meta-analysis. Thus, fitting our data to a multi level regression with a mixed model becomes a rather straightforward endeavor with the main steps becoming simply coding the traits that make each study unique. As such, three variables were added to our merged dataset:

1. Study: which defined the provenance of the data being included.
2. Variant: what type of context our participants were presented with in that study (i.e., Positive, or Negative)
3. Type: denoted the type of kin information that was provided to our participants (i.e., Explicit or Standard)

Let us review a few key concepts and how our studies, individually, have shaped our understanding of similarity and its effect on our participants' predictions. Similarity as putative cue of relatedness has been a staple that has allowed experiments in direct kin-recognition experiments (i.e., using self-morphs) to delve deeper and investigate human kin recognition. In our case, if similarity is treated by our participants as e cue of relatedness than we would expect context-dependent application of this information that is consistent with kin selection and mechanisms that allow for kin recognition. Kin selection is not a context-independent process. We do expect increased cooperation and altruism towards kin, we also expect that individuals will prefer to select kin as targets of beneficial behavior. This means that there has to be a high degree of interaction between this preference and the valence of the behavioral question being posed. Here the distinction between promoting and hiring is the heart of this valence analysis. Individuals endowed with a ability to recognized third-party kinship should expect others to promote their kin and conversely not fire them driving their predictions. An overriding similarity bias clearly fails to account for this valence change. The results of our 11 experiments have provided us with a wealth of information in regards to these predictions that our participants make. Our initial positive context experiments (i.e., Study I, II, III, IV) provided a glimpse at possible minor effect of similarity as a cue of relatedness. However, this glimpse of a connection to kin recognition seemed to disappear as more and more of our negative context experiments (i.e., Study V, VI, VII, VIII) showed the same trend of the similar candidate being predicted to be fired. The final three experiments in which we presented participants with varying levels of explicit information regarding the kin relationships between the actors in our scenarios only serves to reinforce this notion. These results shape our hypothesis that what we are actually detecting is in fact the effect of a perceptual matching bias unrelated to any sort of third-party kin recognition mechanism. The fact that our participants are seemingly unresponsive to any change in valence within our experimental design or even explicit kin information makes a similarity bias a far more fitting explanation for the pattern of results we see in our experiments.

Our goal here, with this meta-analysis is to further investigate this hypothesis in addition to further revealing the trends and patterns of predictions our participants make when looking at the whole of our experimental data. The systematic design of our studies allows us to seamlessly import our results into a meta-analysis that would provide greater power of detecting the nature of the effect of our manipulation of similarity. To this end, we used four models in our meta-analysis with increasing number of covariates added so as to make evident the robustness of our results. We first fitted our data to a simple multilevel regression model (I) that only included the study variable as a random factor. Our model, lmer(ChoiceSimilar 81Sheet + (1+81Sheet|Study), dat), allows for a varying intercept within each study and varying slope for 81Sheet within each study. As was the case in our analysis of our other experiments the effect of similarity is captured by *the intercept term* of our model. Positive (resp., negative) values of the intercept terms denote a preference for the similar (resp., dissimilar) candidate independently of the performance sheet assigned to this candidate. As such we can determine if there is a preference for the similar candidate if the 95% confidence interval for the intercept term does not include the value zero.

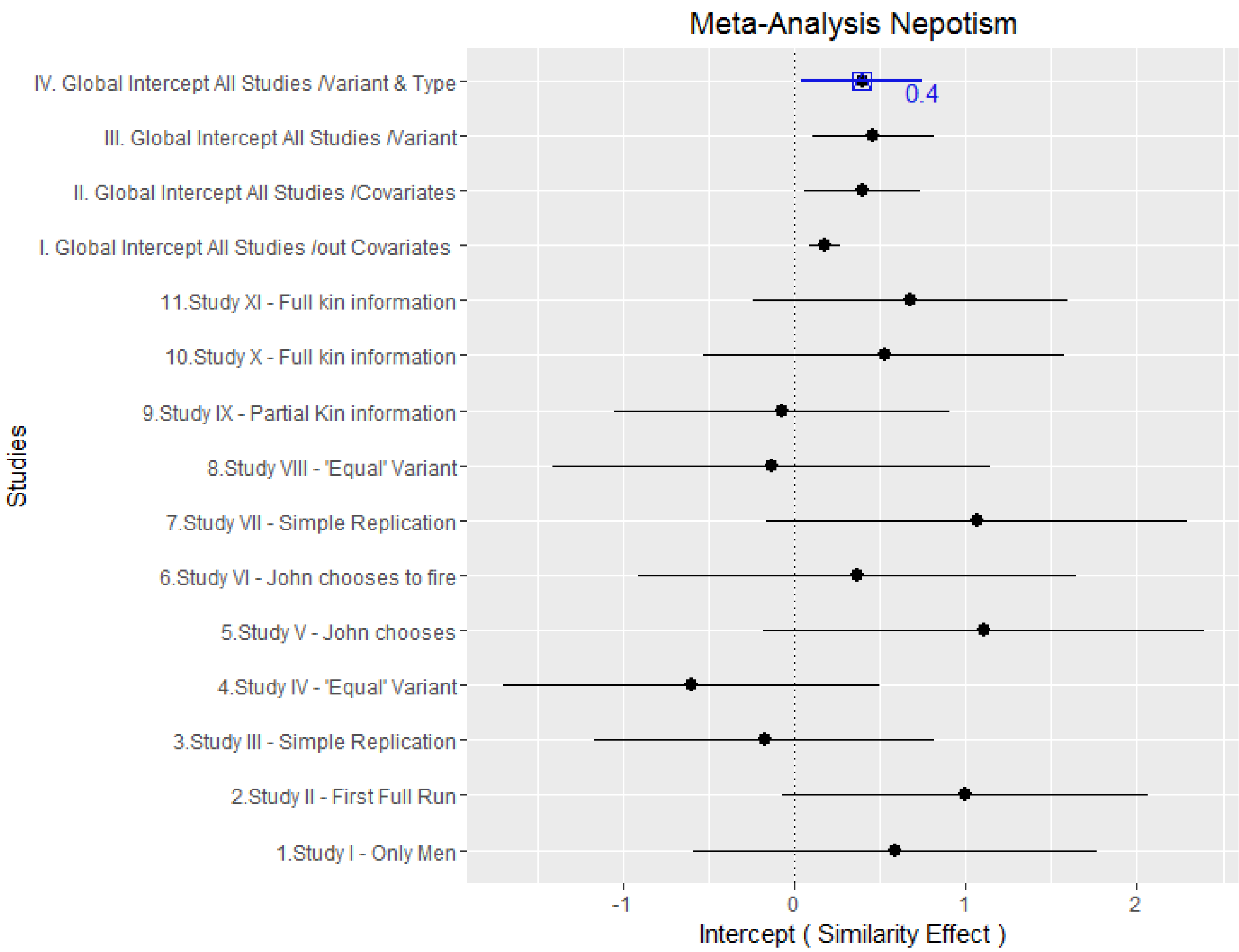

Figure 9: Nepotism Meta-Analysis Forest Plot - All Studies

Using this model we indeed see that in fact our manipulation of similarity is actually having a significant effect ($p<0.05$) on our participants' predictions. However, it is driving what seems to be a context-independent preference for similarity. This result strengthens our hypothesis that our participants predictions are being directed, in part, by a perceptual matching bias wholly unrelated to any effect that might arise from output of a mechanism that incorporates third-party cues of kinship. To our subsequent models we gradually added additional covariates. The second model incorporated demographic covariates such as age, sex and whether the participant was the youngest sibling (II). This model maintained this pattern of results again revealing a strong context-independent effect of our manipulation. This pattern persisted even when type (III) and variant (IV) were added to the model (see table **??**).

Table 31: Nepotism Multi Level Regression Models Results - All Studies

	Dependent variable:			
	ChoiceSimilar			
	(I)	(II)	(III)	(IV)
Constant	0.179***	0.400**	0.461**	0.396**
	(0.044)	(0.172)	(0.181)	(0.179)
PerformanceSimilar	0.855***	0.858***	0.857***	0.858***
	(0.233)	(0.233)	(0.233)	(0.233)
VariantNegative			−0.120	−0.034
			(0.100)	(0.088)
TypeExplicit				0.080
				(0.093)
SexWomen		0.187*	0.189*	0.183*
		(0.105)	(0.105)	(0.105)
Age		−0.006	−0.006	−0.006
		(0.005)	(0.005)	(0.005)
YoungestNo		−0.113	−0.110	−0.112
		(0.092)	(0.092)	(0.092)
Observations	2,292	2,292	2,292	2,292
Log Likelihood	−4,988.375	−4,992.542	−4,993.665	−4,995.044
Akaike Inf. Crit.	9,988.750	10,003.080	10,007.330	10,012.090
Bayesian Inf. Crit.	10,023.170	10,054.720	10,064.700	10,075.200

Note: *p<0.1; **p<0.05; ***p<0.01

This pattern of results is made even clearer when separating studies by valence (i.e., negative & postive scenarios) and type (see figure 10). The patterns we glimpsed during our initial experiments are not simply a fluke. Direction of the effect of our manipulation of similarity is a positive one in all three cases and significant when analyzing the negative context studies ($p<0.05$). The case of the negative variants is highly important as the positive direction and significance fundamentally lays to rest any notion that our effects could be tied to a kin recognition mechanism. The fact that our participants are predicting the firing of the similar candidate is in stark contrast to predictions of nepotistic and kin directed behavior that one would expect if the connection between similarity and relatedness was made.

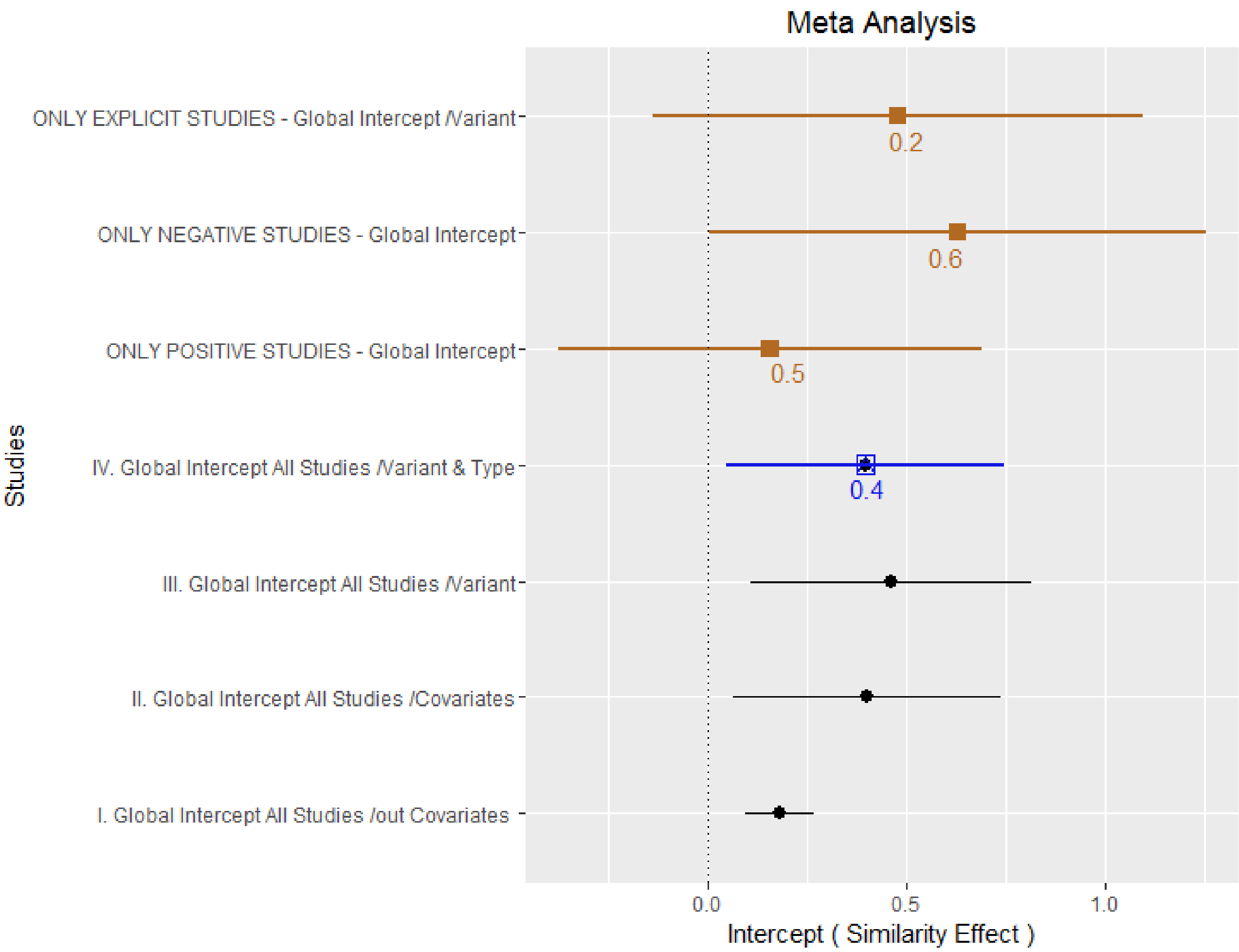

Figure 10: Nepotism Meta-Analysis Forest Plot - Context Analysis

Thus, after eleven studies with a total of 2292 participants, investigating explicit and non-explicit scenarios with negative or positive contexts we can review and recap our findings. First of all, through our pretests we can safely say that participants are fully able to de-

Table 32: Nepotism Multi Level Regression Models Results - Context Analysis

	Dependent variable:		
	ChoiceSimilar		
	(Positive)	(Negative)	(Explicit)
Constant	0.16	0.63*	0.48
	(0.27)	(0.32)	(0.32)
PerformanceSimilar	1.39***	0.13	1.11**
	(0.20)	(0.30)	(0.40)
SexWomen	0.14	0.18	0.20
	(0.18)	(0.19)	(0.17)
Age	−0.004	−0.01	−0.005
	(0.01)	(0.01)	(0.01)
YoungestNo	0.06	−0.22	−0.16
	(0.15)	(0.17)	(0.16)
Observations	702	796	794
Log Likelihood	−1,445.20	−1,795.47	−1,741.53
Akaike Inf. Crit.	2,908.40	3,608.94	3,501.07
Bayesian Inf. Crit.	2,949.39	3,651.06	3,543.16
Note:	*p<0.05; **p<0.01; ***p<0.001		

tect facial similarity with a good degree of accuracy between agents unrelated to themselves (see figure 2.2). This should come as no surprise, as a number of studies have used cues of facial similarity as part of their third party kinship experimental design (e.g., Maloney et al., 2006; DeBruine et al., 2009; Alvergne et al., 2007; Alvergne et al., 2009b; Alvergne et al., 2009a). However, as opposed to direct kin recognition experiments and more importantly third party kin recognition experiments our participants do not seem use similarity as cue for relatedness. In fact our results show that similarity does not trigger expectations of nepotistic behavior. So, how is this possible? Facial similarity undeniably captures some aspect of relatedness and studies such as those of Maloney et al. (2006) or DeBruine et al. (2009), consistently find a high correlation (.92) between perceptions of similarity and perceptions of kinship. Additionally, the two studies from 2009 of Alvergne and their colleagues show that unrelated third-parties can and do use facial appearance to accurately detect kin relationships both within their own culture and cross-culturally (Alvergne et al., 2009b; Alvergne et al., 2009a). Both our pretest results and the literature reinforce only this notion, humans are able to detect similarity in third-parties. However, our results do not stop there and we come back again to the fundamental need to disentangle similarity from relatedness. The fact that similarity captures genetic relatedness confounds our interpretation of the results within the literature. Yes, participants in all of the studies mentioned are able to accurately determine kin relationships. But are they able to do so because of a third-party recognition system or is the fact that their ability to detect similarity directs their predictions the root cause of this accuracy? In each case from the literature, the participants were asked only questions with positive valence. In the studies by Maloney et al. (2006) or DeBruine et al. (2009), participants are tasked with rating the similarity of the facial cues in one group vs another group tasked with rating relatedness. The other procedure used is the one from the Alvergne studies, the "true father" task expanded from Hill et al. (1995). Here participants are required to select the true parent of a child from 3 possible options. The issue should be clear by now, all of the studies mentioned do not modify the valence of the question being posed to their participants. All of these results point towards answering the question of if participants are to detect similarity. Whether or not similarity in these cases is same as an ability to detect relatedness in third-parties remains unanswered. Context matters, and it matters especially in this case as a third-party kin recognition mechanism is bound to kin selection and its predictions. Our literature review revealed that when context is starved participants seem to able to detect relatedness, on the other hand our experimental work, by re-inserting context offers us a different story altogether. Within scenarios in which valence is modified and our participants analysis is bounded by predictions made by kin selection their choices do no reflect a connection between similarity and relatedness. What they do point towards is a perceptual matching bias affecting their predictions. From a methodological standpoint it is clear that future experiments will have to take into account the fact that in investigating third-party kin recognition, it is highly likely that without negative variants their findings might actually be a result of a context-independent similarity bias and not third-party kin recognition. This meta-analysis and body of work offers a clear pattern of results. The significance of the effect of this perceptual matching bias is maintained in our meta-analysis through gradual addition of covariates in our multilevel regression models which provides a clear picture of the mechanism affecting our participants predictions. Our work here provides a robust finding of a perceptual matching bias as opposed to an expected third-party kin recognition effect.

7

THIRD-PARTY EXPECTATIONS OF MATING PREFERENCES

7.1 INTRODUCTION

The second line of inquiry in our research has been focused on third-party expectations of mating preferences. Much of what we discussed in terms of nepotism in the previous chapter remains true when it comes to this topic. We touched upon how recognizing kin affects behavior in very specific ways and for any effect to be linked to kin recognition one must show that this effect is context-dependent. If anything, the results of our first line of research has shown how important this aspect is in determining the presence of actual third-party kin recognition mechanisms.

This shift to the domain of mating choices modifies what patterns of behavior we should expect from individuals endowed with an ability to recognize kin. Both in terms of expectations of third-parties and direct decisions of individuals, kin selections requires that this new decisional context be taken into account. Individuals should be drawn to cooperate with kin and generally adopt a more altruistic attitude towards them while being disincentivized to approach and engage them in terms of mating related behaviour. This results in nepotistic behaviour on one side and reinforces incest avoidance on the other. When it comes to third-parties it has been widely shown that negative moral judgments are a norm when it comes to individual's reactions to mating relationships that include close kin. While, of course, there is some variation in terms of how closely related these individuals must be for their relation to elicit this type of moral condemnations, the same can be said for the actual detrimental fitness effect that incestual relationship have on the offspring from this type of mating behaviour. This being a result of the fine balance that has to be maintained between inbreeding and outbreeding depression. The presence of this predisposition to morally condemn incest brings some weight to the proposition that individuals might be able to recognize third-party cues of kinship and "appropriately" react to them. One possibility is that these third-party cues of kinship inform individuals as to which relationships are a risk of being incestual and thus "deserving" of moral condemnation. Of course, this type of moral judgment can exist in the absence of any such mechanism, being purely based on information from other sources and wholly unrelated to kin recognition, or more specifically, in our case phenotype matching.

Thus, from the vantage point of our research objectives, it seems clear that if there is a mechanism that is able to recognize cues of kinship in third-parties, this mechanism should fuel third-party expectations of mating behaviour. If this were the case, that individuals would be drawn towards relationships between kin that encompass both cooperation, altruistic behaviour and mating, then this mechanism would be incompatible with direct-kin recognition and subsequently would not be an extension of it. Our research tries to disentangle this issue and through it extend the findings of our previous line of research.

To achieve these goals we implemented an experimental design similar to our previous experiments. To briefly summarize, participants were asked to participate in a survey on decision making and dating on the Crowdflower platform and were subsequently taken to our Qualtrics survey. Here they were presented with a picture of Jack and our vignette in which they were told that Jack was in search of a new short term relationship, used a dating app, and narrowed down his options to two women. The picture of the two women were presented to our participants in addition to two personality cards and an overall match rating. These personality cards played a role akin to our performance review sheets in the nepotism experiments, in that our participants were separated into two groups, one group had the performance card with higher overall match rating assigned to the similar woman while the other group had the same card assigned to the dissimilar woman.

As with our nepotism experiment, we took care to create a highly controlled experimental environment by pre-testing our participants' ability to recognize the similar woman, and their perception of the two women in terms of attractiveness, healthiness, trustworthiness, mood and youth. In addition we pre-tested our personality cards to be sure that we would be able to fully account for their influence on our participants' choice.

7.2 FACIAL STIMULI PREPARATION

The procedures and techniques used to create our stimuli in this line of research were very similar to our previous nepotism experiments. Only two male composites were used in these experiments and they were transferred from the nepotism studies. The face used for John (i.e., the boss) in the previous experiments was used here for Jack and the face designated as Control Boss (CB) played a similar role in these experiments. Additionally, the average male prototype created from the majority of white male faces in the database was again used in this series of studies. Finally the new female faces were created specifically for these experiments.

The images used here in creating the female composites were again randomly selected from the CFD and these were limited to being of the same-sex and same-race. As we discussed in our nepotism studies, the initial steps in creating our stimuli involves producing a series of distinct composites of multiple faces. This is done both as a control for any idiosyncratic traits of any one individual's face, which might have unintended effects on our participants' perception of the stimuli and as a technical requirement in creating high quality stimuli. Initially, this process involved the selection of 30 female faces as our procedure involved creating 6 distinct composites and then using these composites we would create our final stimuli. This process was identical to our nepotism studies. In the end, this technique proved to be problematic in creating the stimuli for these mating studies as facial characteristics that were being inadvertently manipulated here were related to mating preferences.

The selection process resulted in 10 female faces that were randomly assigned to one of two groups. These groups consisting of 5 faces each were used to create two distinct female composites which were each assigned a specific proto-identity, proto-Anna and proto-Emily. The male faces imported from the nepotism studies were also assigned their own new designation within these studies, the boss agent's image being assigned to Jack identity while the CB maintained its control role as Control Jack (CJ). Using the two female base identities in addition to the two male ones and the average male prototype, we created the final stimuli using the following procedure:

Table 33: Mating Manipulation Procedure Steps

1.	Jack= Boss Agent from the Nepotism Studies
2.	Emily= pE-90%(Jack - aMale)
3.	Anna=pE-90%(CJ - aMale)

Notes: Proto-Emily (pE); proto-Anna (pA); control-Jack (pCB); average male prototype (aMale); 90% (the strength level of the similarity manipulation)

While we avoided using other-sex images in creating our proto-identities through the averaging technique to avoid the result being androgynous and unrealistic faces, the morphing procedure used here bypasses this issue. Let's take for example the final face of Emily: the identity of proto-Emily is maintained during this process as only a percentage of what makes Jack's face unique, in terms of shape, is applied to this image. That is to say that we are only changing the proto-Emily image by moving its fiducial points xy-coordinates by a percentage of the difference between the same fiducial points in Jack's face and the average male prototype. This difference in terms of xy-coordinates between Jack and the aMale pinpoints exactly to what makes Jack's face unique and by modifying the proto-Emily image according to this value we are creating a final stimuli of Emily that resembles Jack. The same remains true for Anna, however in this case we are using the difference between the aMale and the CJ to modify this image. Resulting in two female faces that are distinct from each other while the process used to create them being identical. This means that any artifacts of our manipulation procedure are nullified.

As with our nepotism experiments we ran a series of pre-tests to verify that our manipulation procedure was having the intended effect. As such, the first of these involved testing the different manipulation strength levels and which of them would best suit our needs.

7.2.1 *Pretest: Manipulation Strength*

We conducted this pre-test at the same time as our nepotism study. And as with the nepotism studies we ended up using the 90% manipulation strength level. However, to resolve the issue of what level to use and verify whether the common theme of 50% or 60% manipulation strength in other studies would apply to the transition we conducted from direct-kin recognition experiments of the past to our third-party kin-recognition paradigm, we first conducted this study. We recruited 150 participants, 115 male and 35 female from the Crowdflower platform between the ages of 19 and 65 (mean age 31.45 years). These subjects also participated in the nepotism pretest. We randomly distributed them to 3 groups with varying strength levels of our manipulation of similarity:

- 50% increase in similarity (n=50),
- 70% increase (n=50),
- 90% increase (n=50)

Subjects in all 3 groups were shown the picture of Jack and they were asked to pay attention and familiarize themselves with him, noting that, in the following sections of the task, they would be asked a series of questions directly related to his appearance and others' likeness to him. Following this, they saw the picture of Jack and Anna's (dissimilar mate choice) picture next to the picture of Emily (i.e., the similar mate choice) and they were asked which of these individual look like the boss. (see figure 11). The appearance of Emily varied by group according to the remembrance manipulation strength level used in each of these 3 groups (i.e., 50%, 70 & 90%). Participants responses were recorded on a standard 6-point response scale with 1-Person 1 is very similar to John and 6-Person 2 is very similar to John, where Person

2 was the designation assigned to the face that was manipulated to resemble John.

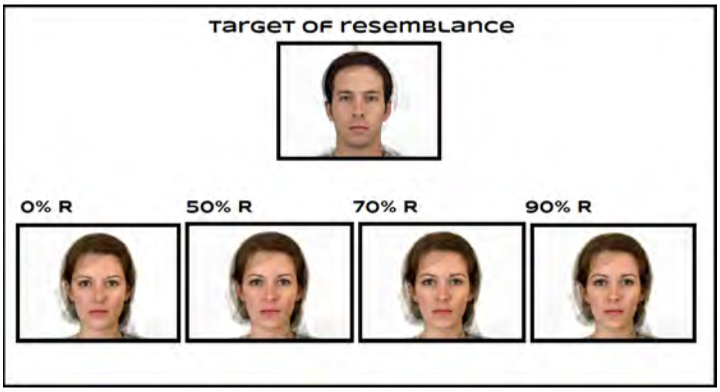

Figure 11: Mating Similarity (R) Manipulation Sequence

Table 34: Mating Linear Regression and Logistic Regression Models

```
lm(MatingQ~Condition+Sex+Age+Siblings+Youngest,data)
```

```
glm(MatingBinary~Condition+Sex+Age+Siblings+Youngest,
family = binomial(link="logit"),data)
```

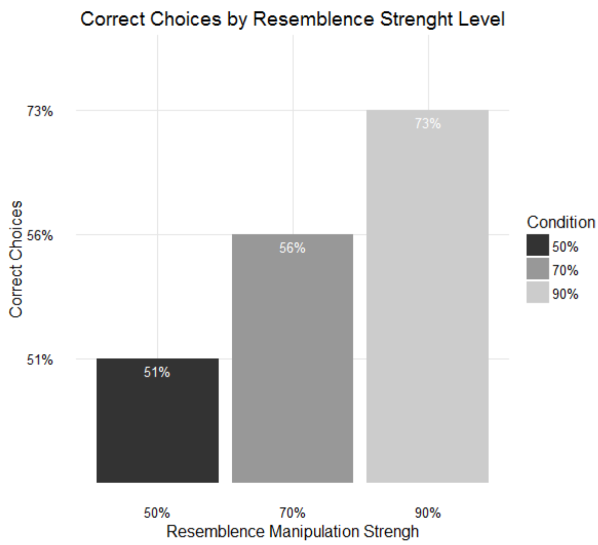

Figure 12: Mating Correct Choices by Similarity Threshold

We analyzed the responses of our participants and the results show that at all similarity manipulation strength levels, participants are able to select the correct response at an above chance level. By fitting our data to linear regression and logistic regression models, even given our switch to other-sex morphs, the most reliable strength manipulation level remains 90%. At this level, 73% (figure 12), of all our participants are able to choose the correct response (i.e., our manipulated candidate), this combined with the strong effect of our manipulation effect ($p<0.001$) that we see in our models (see table 35 and 36) reinforces our continued choice of manipulation strength.

Table 35: Pretest Mating Experiments - Linear Regression Results

	Dependent variable:		
	MatingQ		
	(1)	(2)	(3)
Condition	0.013*	0.014**	0.014*
	(0.007)	(0.007)	(0.007)
SexWomen		0.310	0.323
		(0.275)	(0.279)
Age		−0.006	−0.006
		(0.013)	(0.013)
Siblings			−0.017
			(0.081)
YoungestYes			−0.248
			(0.429)
YoungestNo			−0.202
			(0.435)
Constant	2.593***	2.621***	2.877***
	(0.499)	(0.628)	(0.719)
Observations	150	150	150
R^2	0.023	0.032	0.036
Adjusted R^2	0.017	0.013	−0.004
Residual Std. Error	1.382 (df = 148)	1.385 (df = 146)	1.397 (df = 143)
F Statistic	3.536* (df = 1; 148)	1.632 (df = 3; 146)	0.897 (df = 6; 143)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 36: Pretest Mating Experiments - Logistic Regression Models Results

	Dependent variable:		
	MatingBinary		
	(1)	(2)	(3)
Condition	0.023**	0.027**	0.027**
	(0.010)	(0.011)	(0.011)
SexWomen		0.838*	0.856*
		(0.440)	(0.444)
Age		0.003	0.003
		(0.019)	(0.020)
Siblings			0.003
			(0.120)
YoungestYes			−0.223
			(0.652)
YoungestNo			−0.147
			(0.662)
Constant	−1.189	−1.730*	−1.570
	(0.743)	(0.963)	(1.088)
Observations	150	150	150
Log Likelihood	−98.501	−96.463	−96.393
Akaike Inf. Crit.	201.002	200.927	206.786
Note:	*p<0.1; **p<0.05; ***p<0.01		

Additional Notes: (1), (2), (3) represent the different similarity manipulation strength levels of 50%, 70% and 90%; Ordinary Least Squares (OLS) & logistic represent the two different models used in our analysis, the standard regression model and logistic regression model respectively.

7.3 SCENARIO DEVELOPMENT

In creating the scenarios for our mating related studies we focused on two major goals. First, these studies should follow in the footsteps of our previous studies and present participants with vignettes that explore both positive and negative contexts. If we are to put forward any kind of interpretations based on our results, the experimental environment from which we draw them must allow us to explore both contexts. In the nepotism studies the choices were to either predict who would be fired or promoted. In the case of this series of studies the choices that our participants have to predict rely on which mate they expect Jack to choose in positive context or reject in the negative context.

Table 37: Positive Mating Scenario

Recently, Jack has decided to use a dating app to find a new short-term relationship. This app offers information on its users in the form of a personality card coupled with a "Match Rating", based on a series of tests conducted at signup. POSITIVE: After using the app for a time, Jack has narrowed down his options to two women suggested as good matches by the app. He can only choose to date one of them.

The choice being predicted here is rather more simple than the fire/promote one in the previous studies. Both types of scenario contexts in essence ask our participants to predict which mate will be chosen by Jack. By responding that one is rejected they are implicitly saying that the other one would be chosen. As such one would imagine that the results from both contexts would remain consistent and maintaining the context-dependent nature of this effect would be even more straightforward and easy to achieve.

The second goal was to reinforce the validity of our findings by investigating whether the nature of the information being presented to our participants played a role in shaping our participants predictions. While in the nepotism studies we achieved this by explicitly giving participants the relevant kin relationship information within the vignette, here we chose to adopt a different approach and modify their access to this information by removing its source, the image of Jack. While we will discuss this further in the study that used this approach, suffice it to say this plays the same role as our explicit variants in the nepotism experiments through a different approach.

Table 38: Negative Mating Scenario

Recently, Jack has decided to use a dating app to find a new short-term relationship. This app offers information on its users in the form of a personality card coupled with a "Match Rating", based on a series of tests conducted at signup. NEGATIVE: After using the app for a time, Jack has narrowed down his options to two women suggested as good matches by the app. He can only date one of them, which means he must reject one.

7.4 PRETEST: PERSONALITY CARD

In the same way that our nepotism studies used performance review sheets, this set of studies use personality cards and linked match ratings that serve the same role as overall score on the performance review sheets. We created two personality cards with 8 traits: friendliness, intelligence, sense of humor, creativity, work ethic, romance, trustworthy and cheerfulness. All of these traits were equal except intelligence and work ethic (see table 39) which were higher on card 1 as compared to card 2 (i.e., 78 vs. 76 intelligence and 75 vs. 77 work ethic). These two traits were picked as previous research (facial traits ref.) and our own pre-test showed work ethic is not a major factor in people's decisions as to whom to date, while we found intelligence to be associated with a higher dating success (i.e., individuals prefer more intelligent dating partners). This distinction is important as we wanted to create one personality card that was preferred, while remaining very close to the other. By selecting intelligence to be higher in one while the other remained higher in work ethic, makes it so that the intelligence one is overall preferred (see table 39 for the two performance cards).

We ran a pretest to verify if this preference for one card can be found in an experimental setting recruiting 124 participants from the Crowdflower platform participants, 86 male and 38 female between the ages of 18 and 60 (mean age 31.58 years). Running a one-sample t-test revealed that the personality card with a higher intelligence rating (1-*Definitely higher Intelligence Card*, 6-*Definitely higher Work Ethic Card*) is selected at a significantly higher rate than the higher work ethic card ($M=3.13$, test value = 3.5, $p=.004$). .

Table 39: Personality Cards (PC) and Match Rating

Higher Performance Sheet		Lower Performance Sheet	
SAMPLE TRAITS		SAMPLE TRAITS	
FRIENDLYNESS	78	FRIENDLYNESS	78
INTELLIGENCE	78	INTELLIGENCE	76
SENSE OF HUMOR	75	SENSE OF HUMOR	75
CREATIVITY	76	CREATIVITY	76
WORK ETHIC	75	WORK ETHIC	77
ROMANCE	78	ROMANCE	78
TRUSTWORTHY	75	TRUSTWORTHY	75
CHEERFULNESS	75	CHEERFULNESS	75
MATCH RATING		MATCH RATING	
			

7.5 PRETEST: FACIAL STIMULI TRAITS

We ran two pretests in which we tested a number of traits of our final two female stimuli, Emily and Anna. The basic issue starting out has been that, if there is a third-party extension of the inbuilt incest avoidance mechanism seen in direct kin-recognition, our participants should predict that Jack should select Anna, not Emily. As such, if our participants are drawn too much towards Anna, as a result of our manipulation procedure enhancing certain facial traits that are correlated to general attractiveness (i.e., not only physical attractiveness), then we would be unable to disentangle this effect from any effect arising from our specific similarity manipulation. Finally, as discussed in the facial stimuli preparation section, during the beginning stages of this research we used a technique almost identical to the one used in the nepotism studies to create our female stimuli. We ran a series of pretests using a within-subject design to investigate this very issue. During this first run, we used the technique from the nepotism series to create our stimuli. We will discuss the reason why this was later changed after presenting the results from this run.

We recruited 59 participants from the Crowdflower platform, 41 male and 18 female, between the ages of 18 and 57 (mean age 32.55 years). They were asked a series of questions related to the appearance of the two female faces (i.e., Emily and Anna). For each of the faces the participants were required to rate them on a 7-point scale on attractiveness, health and trustworthiness. In addition they were asked how they would rate Emily and Anna's mood on a 6-point scale (1-*Extremely Happy*, 6-*Extremely Unhappy*) and finally they were asked approximate their age.

Analysis of the data revealed that there was no significant difference in attraction (4.57 Emily vs. 4.40 Anna; $p=.317$) and trustworthiness (4.71 Emily vs. 4.59 Anna; $p=.472$). However, our participants rated Anna as being slightly more healthy (4.66 Emily vs 5.06 Anna; $p=0.18$), younger (28.57 Emily vs. 26.67 Anna; $p<.000$) and happier (-.39 Emily vs .03 Anna; $p<.000$). To further investigate these findings we ran a number of tests on the correlations between how participants rate Anna's traits as compared to how they rate the same traits in Emily. The perceived age of Anna does not significantly correlate with any of Emily's traits (i.e., except age .438 correlation, $p>.01$). Thus, it seems that while there is a significant difference in perceived age this is not really affecting the other judgements the participants make. In addition, even though the difference is significant, the actual ages 28 vs. 26, (Emily vs. Anna) do not pose a significant issue as they are both well within the young-adult age bracket. On the other hand, this still leaves us with the issues of health and mood which both should and do play a role in how our participants view Emily and Anna. The issue of health seems to be an artifact of our composite creation procedure. While the images that make up each composite were selected at random, the combination of images resulted in one final stimuli being inadvertently made to look slightly healthier. To correct this we modified our manipulation procedure by reducing the number of female composites to two and using these composites as the proto-identities of Anna and Emily. However, the issue of mood is slightly more problematic as, by all accounts, as the base pictures used to create the composite all have neutral expressions which were verified and tested as part of the development of the Chicago Face Database. The very nature of the manipulation does result in this type of unintended modification and realistically this type of result can only be corrected by recreating the stimuli and testing the results. For this reason we implemented the new manipulation procedure we outlined in the stimuli preparation section.

Using these new stimuli we recruited a new batch of 65 participants from the Crowdflower platform, 45 male and 20 female, between the ages of 18 and 60 (mean age 30.70 years). As with the first run we used a within-subject design and we asked our participants a series of questions related to the appearance of the two female faces (i.e., Emily and Anna). For each faces the participants were required to give a rating on a 7-point scale of attractiveness, healthy and trustworthiness. In addition they were asked how they would rate Emily and Anna's mood on a 6-point scale (1-*Extremely Happy*, 6-*Extremely Unhappy*) and finally they were asked approximate their age.

This new dataset using the new face stimuli revealed that now Emily is slightly more attractive (4.49 Emily vs. 4.18 Anna, p=.034), more trustworthy than Anna (4.71 Emily vs. 4.59 Anna; p=.472) and slightly younger (Emily 28.29 vs. Anna 29.61, p=0.013). On the other hand, there is no longer any significant difference in healthiness (Emily 4.76 vs. Anna 4.67, p=.501). The only metric that remains in favour of Anna is mood, with Emily being rated slightly less happy than Anna (Emily -.15 vs. Anna .06, p=0.12). This transformation from an overall higher set of ratings for Anna in the first run to the results of this run allows for greater flexibility in analysis the effect of our similarity manipulation by using this new set of stimuli. As noted, the fact that subjects would be drawn to Anna would make disentangling the effect of our similarity manipulation from a general preference for Anna unfeasible. With the new changes the only rating that remains slightly problematic is mood, though is significantly reduced as a result of our change in the manipulation technique.

7.6 experimental studies

Drawing inspiration from our nepotism studies, our goal here is to investigate the context-dependence of a possible third-party kin recognition mechanism. When it comes to direct kin recognition one can produce concrete testable hypotheses that are consistent with inbuilt mechanism that is responsible for kin recognition. From using a number of cues, such as sensory information in the form of phenotype matching, olfactory and auditory stimuli to more complex nature-nurture interactionist perspectives such as the Westermarck's hypothesis, the predictions one makes, based on an ability to recognize kin, remain constant. Individuals should maintain a context-dependent response that reinforces a tendency to prefer to cooperate with kin, favouring them over non-kin in terms of pseudo-altruistic interactions (i.e., pseudo because this type of behaviour has an indirect fitness benefit) and on the flip side avoid inflicting "losses" on kin. In terms of mating behaviour the trend is reversed with a greater inbuilt incentive to avoid choosing mates that might be close kin thus avoiding negative fitness impact such unions might entail. If we are to grant that this type of kin-recognition mechanism extends and is able to affect predictions of behaviours of third-parties, then it follows that the same hypothesis should be maintained. As such, in the first series of experiments we investigated the notion that agents that look like each other are expected to cooperate more, while now we will investigate mating related behaviour and see whether or not participants expect agents that look like each other to avoid each other as mates. We test this by asking participants to predict the dating choices of one agent when presented with a potential similar vs. dissimilar mate.

7.6.1 *Study 1*

7.6.1.1 *Introduction*

Attitudes related to incestous behaviour have been widely researched. Some of these studies, such as those of Lieberman et al. (2003) or Fessler et al. (2004), have involved reactions to third-party fictional cases of sibling incest. These types of studies have shown that co-residence with an opposite-sex individuals is a reliable predictor of increased lack of tolerance and disgust towards this type of behaviour. Though these studies investigate co-residence and by extension the Westermarck's hypothesis, their results provide a crucial piece of information that informs our studies. People generally condemn and are disgusted by incestual behavior. In these studies, the information that these relationships are incestual is directly provided, the important aspect to keep in mind is that cues of relatedness (in this case actual direct information) plays a key role in directing moral judgments applied to said relationships. As such, one might hypothesize that if one has the ability to recognize facial cues of relatedness in others and this information would also help in directing moral judgments applied to third-party relationships, as does direct knowledge of the kinship connections. It naturally follows that individuals should have the same aversion and disgust reaction towards those incestual relationships as if they were directly told of kinship of those mates. Following from this, their predictions as to which mate a third-party might be inclined to select when presented with a similar-mate and dissimilar-mate would naturally be skewed in favor of the mate that would not elicit a moral condemnation and disgust effect (i.e., the dissimilar mate).

7.6.1.2 *Methods*

We again recruited our participants through the Crowdflower platform asking them to take part in a short task related to decision making and dating. In all our studies participants were informed as to the steps they will follow and that they would receive 0.30 euro for participation. The exit survey conducted by Crowdflower revealed that trend of satisfaction with the remuneration, time, and difficulty of the task continued from our nepotism studies as such nothing was modified in this regard. To maintain the integrity of our sample and avoid participants "cheating" the system and enrolling multiple times in our task, we implemented both IPs filtering through Qualtrics and relied on Crowdflower own security measures.

The initial section of the survey ensured that participants would pay close attention to our stimuli. For this participants were presented with all 3 picture of the agents that would be part of the scenarios (i.e., Emily, Anna & Jack). However, their "identity" was not revealed at this stage, instead they were simply asked "[Do] these people remind you of someone?". Participants were then presented with the picture of Jack and the scenario in which they learn that he is in search a new short-term relationship and he has decided to use a new app that offers information on its matches in the form of a personality card and a "Match Rating". This study used the positive context: participants were informed that after using the app for a time, Jack has narrowed down his choices to two women and he must make a choice as to which one to date. Similar to the nepotism series, these studies followed a between-subject design when participants saw different personality cards and match ratings assigned to each potential mate based on the group they were part of, creating the basis of our experimental design.

We recruited 200 participants from the Crowdflower platform, 145 male and 55 female, between the ages of 18 and 65 (mean age 32.25 years). Participants were randomly distributed into one of two groups, as per our between subject design. Half of our participants were randomly distributed to the condition in which the mate that resembled Jack (i.e., Emily,) was assigned the personality card with the higher overall intelligence score, 78 vs. 76 intelligence. The rest of our participants being assigned to the other group, which had the higher work ethic score on the personality card assigned to Emily. Finally, all participants were asked the same experimental question: "Which one do you think Jack will ask out on a date?" Their responses were recorded on a standard 6-point response scale with 1-*Definitely Emily* and 6-*Definitely Anna*. In addition, they were asked to justify their choice "Why do you believe that Jack will make this choice?" and as an attention check "Which of two women had the higher overall Match Rating?".

7.6.1.3 *Results*

Analysis of our data required us to first recode our variables. Our dependent variable, the choices our participants made (i.e., ChoiceSimilar), was coded so that the higher values represented the choice of the similar mate denoted by *strictly positive values* (+1, +2, +3) and the choice of the dissimilar mate denoted by *strictly negative* values (-1, -2, and -3). Our independent variable (i.e., Intellcard) representing which mate is assigned the higher intelligence personality card was denoted by +1 when it was assigned to the similar mate and -1 when it was assigned to the dissimilar mate. We fit the following model, ChoiceSimilar Intellcard (+ Demographic Covariates), as with the nepotism experiments the effect of similarity is captured by *the intercept term.*. Positive (resp., negative) values of the intercept terms denote a preference for the similar (resp., dissimilar) mate independently of the personality card assigned to this candidate. Our demographic covariates included biological sex, age and whether the participant was the youngest sibling. The sibling covariate was included to account for an increased ability to recognize kin as a function of their co-residence with with a sibling and possible effects of MPA. Additionally, the mating framework of this experiment also opens up the possibility of additional biases arising from an increased predisposition to condemn and elicit disgust effects from incestual relationships.

Table 40: Linear Regression Model Study 1

```
lm(ChoiceSimilar ~ IntellCard+Sex+Age+Youngest, data=MatPosA)
```

In contrast with our nepotism experiments, a preference for the similar mate would actually be in opposition to predictions made on the basis of an inbuilt third-party kin recognition mechanism. As such, in our analysis we are mainly interested in verifying if the constant of the model is significantly different from zero, thus suggesting that our manipulation of similarity is significantly impacting the predictions our participants make, the main difference being that this value should be negative.

Our model (R^2=0.023, $F(4,195)$=1.148) revealed that one of the factors impacting our participants decision is the personality card assigned to each mate, though this effect is only marginally significant ($p<0.1$). In a similar way to the overall performance effect seen in our nepotism experiment, here we see that increased intelligence plays the same role, though not nearly as strong. Barring this, we were interested in the constant of our model, 0.99, 95% CI [-0.002, 1.98] , which revealed that in fact our similarity manipulation is having a significant impact on our participants decisions (p=0.05). However, the direction of the constant shows us that participants are preferring the similar mate, which means that this effect is in fact more in line with predictions made on the basis of a homogeneity preference as opposed to any predictions based on kin recognition. Given this result we are left with an interesting conundrum. Is this effect a result of our similarity manipulation, a coincidental marginal effect or is this simply the result of the difference in attractiveness between the female faces we presented to our participants? Our pretest revealed that Emily is slightly more attractive (4.49 Emily vs. 4.18 Anna, p=.034), more trustworthy than Anna (4.71 Emily vs. 4.59 Anna; p=.472) and slightly younger (Emily 28.29 vs. Anna 29.61, p=0.013). When developing the stimuli this slight advantage that Emily had in terms of these traits did not pose an issue as the effect of our similarity manipulation should have drawn our participants towards Anna and the positive traits of Emily would only moderate this effect. However, these results raises doubts as to the validity of this interpretation. The preference seen here for Emily could be a simply a results of her slight advantage over Anna in terms of attractiveness and youth and not a result of her similarity to Jack.

Table 41: Linear Regression Model Study 1 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.991*
	(0.507)
IntelligenceSimilar	0.231*
	(0.130)
SexWomen	−0.201
	(0.294)
Age	−0.010
	(0.015)
YoungestNo	−0.135
	(0.281)
Observations	200
R^2	0.023
Adjusted R^2	0.003
Residual Std. Error	1.821 (df = 195)
F Statistic	1.148 (df = 4; 195)
Note:	*p<0.1; **p<0.05; ***p<0.01

7.6.2 *Study 2*

7.6.2.1 *Methods*

Our second study focuses on disentangling the issue of the preference for Emily (i.e., the similar mate) seen in our previous experiment. To separate the effect of Emily's slight advantage in terms of facial traits vs. an effect of homogeneity arising from our similarity manipulation we decided to proceed with a new study in which we removed Jack's face from our participants' prediction equation. This study was identical to the previous one in every regards except that here the participants are never shown Jack's face. As such there is no effect of similarity as the compare point, Jack's face, is not there to inform participants in any way in terms of similarity of any of the potential mates. In this situation if the results of our previous experiment are simply Emily's advantage in terms of attractiveness directing our participants predictions, then we should see no significant changes in terms of the constant of our model. On the other hand, if our similarity manipulation is the main factor influencing our participants prediction our the constant of the model should reflect and its significance should evaporate.

196 participants were recruited from the Crowdflower platform, 136 male and 60 female, between the ages of 18 and 69 (mean age 33.51 years). They were randomly distributed into one of two groups. For half of our participants had the mate that resembled Jack (i.e., Emily,) assigned the personality card with the higher overall intelligence score, 78 vs. 76 intelligence. While the rest of our participants had the higher work ethic score on the personality card assigned to Emily. All were asked the same experimental question: "Which one do you think Jack will ask out on a date?" With their responses being recorded on a standard 6-point response scale with 1-*Definitely Emily* and 6-*Definitely Anna*. They were asked two additional question, first to justify their choice "Why do you believe that Jack will make this choice?" and second, as an attention check, "Which of two women had the higher overall Match Rating?

7.6.2.2 *Results*

The choices our participants made (i.e., ChoiceSimilar), were coded as -3, -2, -1, +1, +2, and +3. The higher values represented the choice of the similar mate denoted by *strictly positive values* (+1, +2, +3) and the choice of the dissimilar mate denoted by *strictly negative* values (-1, -2, and -3). As before, Intellcard represents which mate is assigned the higher intelligence personality card denoted by +1 when it was assigned to the similar mate and -1 when it was assigned to the dissimilar mate. These variable were then utilized in our regression model ChoiceSimilar Intellcard (+ Demographic Covariates).

Table 42: Linear Regression Model Study 2

```
lm(ChoiceSimilar ~ IntellCard+Sex+Age+Youngest, data=MatPosB)
```

The results of our previous study showed that one of the factors influencing our participants predictions is a slight preference for intelligence, however in this study our model (R^2=0.015, F(4,191)=0.74) shows that this preference does not meet the same requirement of significance (p>0.1). While the difference in terms of intelligence is not in fact that pronounced, 78 vs. 76, our pretest showed that it does in fact play a role and our previous study reinforced this notion. As such, any deviation from this can be seen a normal result of the experimental design. The difference in personality cards was not designed to be an overriding major factor in our participants predictions, it was designed to provide a recognizable difference that influences our participants predictions while maintaining our ability to detect the effect of our similarity manipulation. On the other hand analysis of the constant in our model does not pose any such difficulties.

As we expected removing the image necessary for the comparison in terms of similarity, Jack, removed any significance related to a preference towards any of the mates. While there is no Jack to compare to and thus the constant, 0.121; 95% CI [-0.78, 1.02], is not connected to similarity, it still allows us to see that there is no longer any significant preference for one of them independent of the personality card. This result reinforces our interpretation that the results of our first study are an effect arising from a possible homogeneity preference and clearly not a third-party kin recognition mechanism.

Table 43: Linear Regression Model Study 2 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.121
	(0.462)
IntelligenceSimilar	0.189
	(0.125)
SexWomen	0.217
	(0.276)
Age	−0.001
	(0.013)
YoungestNo	0.093
	(0.257)
Observations	196
R^2	0.015
Adjusted R^2	−0.005
Residual Std. Error	1.744 (df = 191)
F Statistic	0.742 (df = 4; 191)
Note:	*p<0.1; **p<0.05; ***p<0.01

7.6.3 *Study 3*

7.6.3.1 *Methods*

To follow up on the results from the first two studies we ran a full replication study to see if our results remain consistent and thus reinforce any interpretations we made. This study was identical to Study 1 in every regard: 204 participants were recruited from the Crowdflower platform, 136 male and 64 female, between the ages of 18 and 66 (mean age 31.91 years). Half of our participants had Emily assigned the personality card with the higher overall intelligence score, 78 vs. 76 intelligence. While for the rest of our participants Anna had higher work ethic score on the personality card. Every participant was asked the same experimental question: "Which one do you think Jack will ask out on a date?" With their responses being recorded on a standard 6-point response scale with 1-*Definitely Emily* and 6-*Definitely Anna*. They were asked two additional question, first to justify their choice "Why do you believe that Jack will make this choice?" and second, as an attention check, "Which of two women had the higher overall Match Rating?

7.6.3.2 *Results*

Responses were coded -3, -2, -1, +1, +2, and +3 (i.e., ChoiceSimilar). So that the higher values represented the choice of the similar mate denoted by *strictly positive values* (+1, +2, +3), while the choice of the dissimilar mate were denoted by *strictly negative* values (-1, -2, and -3). Second, we created a variable named Intellcard, which took the value +1 when the similar mate received the higher intelligence personality card, and the value -1 when the similar mate received the lower intelligence personality card. Thus, when we fit the following model, ChoiceSimilar Intellcard (+ Demographic Covariates), the effect of similarity is captured by *the intercept term.* Positive (resp., negative) values of the intercept terms denote a preference for the similar (resp., dissimilar) mate independently of the personality card assigned to this mate.

Table 44: Linear Regression Model Study 3

```
lm(ChoiceSimilar ~ IntellCard+Sex+Age+Youngest, data=MatPosC)
```

Analysis of our regression model (R^2=0.043, F(4,195)=2.16) revealed that, as expected, participants have a significant preference for intelligence ($p<0.05$), confirming the fact that the lack of significance in our previous study might be a result of the relative size of the effect that makes it possible for it to remain undetectable in some samples. The constant of our model, 0.614; 95% CI [-0.35, 1.58], shows that, though not significant here, the positive trend and thus preference for the similar mate continues.

Table 45: Linear Regression Model Study 3 - Results

	Dependent variable:
	ChoiceSimilar
Constant	0.614
	(0.496)
IntelligenceSimilar	0.315^{**}
	(0.126)
SexWomen	0.255
	(0.269)
Age	-0.015
	(0.014)
YoungestNo	-0.050
	(0.263)
Observations	200
R^2	0.043
Adjusted R^2	0.023
Residual Std. Error	1.765 (df = 195)
F Statistic	2.167^{*} (df = 4; 195)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

This trend of a preference for the similar mate signals that this effect is highly unlikely to be connected to a kin recognition mechanism and is instead a good candidate for a general homogeneity preference. With these results in hand, our objective becomes to verify this hypothesis. So, our commitment to context-dependent scenarios becomes even more important. Switching to negative context in which Jack has to choose whom to reject we can verify if this is in fact an effect of a homogeneity preference or only a simple perceptual matching bias.

7.6.4 *Study 4*

7.6.4.1 *Methods*

Our positive context studies investigating predictions on mating behaviour only allow us go so far in shining a light on the mechanism behind our results. As it stands , the preference for the similar mate seen in Study 1 and 3 offers us two possibilities. First, as with the nepotism experiments what we are seeing might be a result of a perceptual matching bias that directs our participant towards the similar choice independent of context. If this were the case, implementing a negative context scenario in this series of experiments should give us a similar pattern of results as seen in the nepotism studies with participants maintaining a preference for the similar mate even when the choice is to reject them. On the other hand, a preference for homogeneity would mean a context dependent response following the opposite pattern of response guided by a kin-recognition mechanism. As such we would expect participants to choose the similar mate as a date in the positive context, while preferring to reject the dissimilar mate in the negative context. This would be the polar opposite pattern of results compared to a kin-recognition direct prediction pattern. As mentioned, a mechanism utilizing kin-recognition would dictate that participants prefer the dissimilar mate in the positive context, while rejecting the similar mate in the negative one. For this reason we ran a new study using the negative context scenario.

This study was identical in design to the previous mating studies with the only exception been the negative context provided by the scenario. We recruited 204 participants from the Crowdflower platform, 136 male and 64 female, between the ages of 18 and 66 (mean age 31.91 years). Half of our participants had Emily assigned the personality card with the higher overall intelligence score, 78 vs. 76 intelligence. While for the rest of our participants Anna had higher work ethic score on the personality card. Every participant was asked the same experimental question: "Which one do you think Jack will reject?" With their responses being recorded on a standard 6-point response scale with 1-*Definitely Emily* and 6-*Definitely Anna*. They were asked two additional question, first to justify their choice "Why do you believe that Jack will make this choice?" and second, as an attention check, "Which of two women had the higher overall Match Rating?"

7.6.4.2 *Results*

We followed the same procedure as our other experiments with the choices our participants coded so that the higher values represented the choice of the similar mate denoted by *strictly positive values* (+1, +2, +3), while the choice of the dissimilar mate were denoted by *strictly negative* values (-1, -2, and -3). Intellcard, represents which mate is assigned the higher intelligence personality card. This is denoted by +1 when it was assigned to the similar mate and -1 when it was assigned to the dissimilar mate. These variable were then utilized in our now standard regression model, ChoiceSimilar Intellcard (+ Demographic Covariates).

Table 46: Linear Regression Model Study 4

```
lm(ChoiceSimilar ~ IntellCard+Sex+Age+Youngest, data=MatNegA)
```

Interestingly, our regression model (R^2=0.009, F(4,190)=0.44) reveals that in the negative context our participants' predictions are not significantly affected by the personality card ($p<0.1$). This might again be a result of the relative size of the effect or, if the pattern persists, this might indicate that in the case of predicting the rejection of a potential mate, it does not play a significant role in directing their choices. On the other hand the constant of our model, -0.120; 95% CI [-0.99, 0.75], reveals that in the case of this negative context, participants do not seem to be influenced in either direction and with both of these results (i.e., lack of preference for intelligence and similarity) they seem to be predicting Jack's choice at random or at the very least not being influenced by these two factors. This series of results requires us to run a replication study to verify that this pattern remains consistent before we can present any meaningful interpretation.

Table 47: Linear Regression Model Study 4 - Results

	Dependent variable:
	ChoiceSimilar
Constant	−0.120
	(0.448)
IntelligenceSimilar	0.023
	(0.122)
SexWomen	0.139
	(0.258)
Age	−0.003
	(0.013)
YoungestNo	0.301
	(0.248)
Observations	195
R^2	0.009
Adjusted R^2	−0.011
Residual Std. Error	1.693 (df = 190)
F Statistic	0.449 (df = 4; 190)
Note:	*p<0.1; **p<0.05; ***p<0.01

7.6.5 *Study 5*

7.6.5.1 *Methods*

As we discussed, the primary focus of this study will be to replicate the findings of Study 5 and see if the lack of influence exerted by our manipulation of similarity or the difference in intelligence remains. As such this study is identical in every regard to Study 5, 201 participants from the Crowdflower platform, 131 male and 70 female, between the ages of 18 and 66 (mean age 32.81 years). Half of our participants had Emily assigned the personality card with the higher overall intelligence score, 78 vs. 76 intelligence. While for the rest of our participants Anna had higher work ethic score on the personality card. Every participant was asked the same experimental question: "Which one do you think Jack will reject?" With their responses being recorded on a standard 6-point response scale with 1-*Definitely Emily* and 6-*Definitely Anna*. They were asked two additional question, first to justify their choice "Why do you believe that Jack will make this choice?" and second, as an attention check, "Which of two women had the higher overall Match Rating?

7.6.5.2 *Results*

ChoiceSimilar was coded so that the higher values represented the choice of the similar mate denoted by *strictly positive values* (+1, +2, +3), while the choice of the dissimilar mate were denoted by *strictly negative* values (-1, -2, and -3). Intellcard took the value +1 when the similar mate received the higher intelligence personality card, and the value -1 when the similar mate received the lower intelligence personality card. These variable were then utilized in our regression model, ChoiceSimilar Intellcard (+ Demographic Covariates), focusing on the effect of our similarity manipulation which is captured by *the intercept term.*.

Table 48: Linear Regression Model Study 5

```
lm(ChoiceSimilar ~ IntellCard+Sex+Age+Youngest, data=MatNegB)
```

Analysis of the linear regression model (R^2=0.004, F(4,196)=0.175) revealed that the pattern seen in Study 4 continues with our participants' predictions not being affected by either the personality card (p>0.1) or by our manipulation of similarity which we can analyze through the constant in our model, -0.091; 95% CI [-1.02, 0.84]. Given these results it seems that the switch from positive context to negative changes the equation in terms of what factors our participants take into account when making their prediction.

Table 49: Linear Regression Model Study 5 - Results

	Dependent variable:
	ChoiceSimilar
Constant	−0.091
	(0.476)
IntelligenceSimilar	0.022
	(0.124)
SexWomen	0.095
	(0.259)
Age	−0.005
	(0.013)
YoungestNo	0.177
	(0.263)
Observations	201
R^2	0.004
Adjusted R^2	−0.017
Residual Std. Error	1.744 (df = 196)
F Statistic	0.175 (df = 4; 196)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Neither the homogeneity preference nor the kin recognition based interpretation fully account for these results. However, one possibility is that the switch to the negative context modifies our participants' perception of the choice being made by Jack and as such obfuscates the effect of homogeneity behind a desire to avoid rejecting one mate over the other. The difference in terms of traits on the personality card might not be sufficient to warrant rejecting one the mates and the effect of our manipulation of similarity and by extension the preference for homogeneity might not be sufficient to drive our participants prediction in one way or the other. Our main goal has been to create an experimental environment in which the only factor that will be influencing our participants predictions would be our manipulation of similarity, as such, it is to be expected that if our participants are looking for reasons to expect the rejection one of the potential mates they would not find those reasons in any other place than our manipulation of similarity. However, this does not imply that our manipulation of similarity would be sufficient to expect this rejection if this stimuli is not in fact connected to mechanisms connected to third party kin-recognition. As with our previous set of experiments it becomes abundantly clear that any worthwhile interpretations can only be reached in these types of experiments by looking closely at a meta-analysis of all our data so as to disentangle these subtle differences that can make or break any interpretation of these results.

7.7 META-ANALYSIS MATING STUDIES

We approached the meta-analysis of our data regarding predictions of potential mates in a similar fashion to our nepotism dataset. As such, the advantages we presented in the case of our nepotism studies remain true for this series. And as our approach has been consistent, all our studies using the same experimental design, questions and coding norms, the framework upon which to base our analysis is rigorous. However, there are two distinction worth mentioning. Firstly, Study 2 was not included in the meta-analysis, because this study did not use Jack's picture, thus it's unrelated to any effect of similarity, and as such it would not be beneficial to add it to our analysis here. To recap, this experiment was designed as control run in which we could observe our participants' predictions without implementing our manipulation of similarity. Second, this series of studies did not vary by type, so there is no differentiation between the amount of kin information presented to our participants, that is to say, all experiments used a type setting akin to the standard implicit model used in our nepotism series. With this in mind, our main interest remains the intercept value in our models that represents the effect of our similarity manipulation. Positive (resp., negative) values of the intercept terms denote a preference for the similar (resp., dissimilar) mate independently of the personality card assigned to this mate. Accordingly, we can determine if there is a preference for the similar mate if the 95% confidence interval for the *intercept term* does not include the value zero. Considering all of these factors, fitting this data to our multilevel regression models only requires us to code a few essential moderators:

1. Study: which defined the provenance of the data being included (e.g., MatPosA for Study 1; MatPosB for Study 2; etc.).
2. Variant: what type of context our participants were presented with in that study (i.e., Positive, or Negative)

As before, we gradually added an increasing number of covariates to our model starting out with a multilevel regression, ChoiceSimilar IntellCard + (1 | Study) that only included the study as a random factor (I). The results of this model showed no significant effect of our manipulation of similarity ($p>0.1$). This pattern of results continued when we added demographic covariates such as age, sex or sibling status (II) with none of them significantly affecting the results. Our participants seemed to be unresponsive to our manipulation of similarity with this pattern of results remaining unchanged even when we finally added the Variant variable to our model (III), ChoiceSimilar IntellCard + (1 | Study) + Variant + Sex + Age+ Youngest (see table 50).

Table 50: Multi Level Regression Models Results - All Studies

	Dependent variable:		
	ChoiceSimilar		
	(1)	(2)	(3)
Constant	0.15	0.37	0.07
	(0.14)	(0.27)	(0.27)
IntelligenceSimilar	0.14	0.15	0.15
	(0.08)	(0.08)	(0.08)
VariantPositive			0.57^{**}
			(0.18)
SexWomen		0.07	0.06
		(0.13)	(0.13)
Age		−0.01	−0.01
		(0.01)	(0.01)
YoungestNo		0.07	0.08
		(0.13)	(0.13)
Observations	800	796	796
Log Likelihood	−1,586.58	−1,584.95	−1,584.62
Akaike Inf. Crit.	3,185.17	3,187.90	3,189.24
Bayesian Inf. Crit.	3,213.27	3,230.02	3,236.04
Note:	*p<0.05; **p<0.01; ***p<0.001		

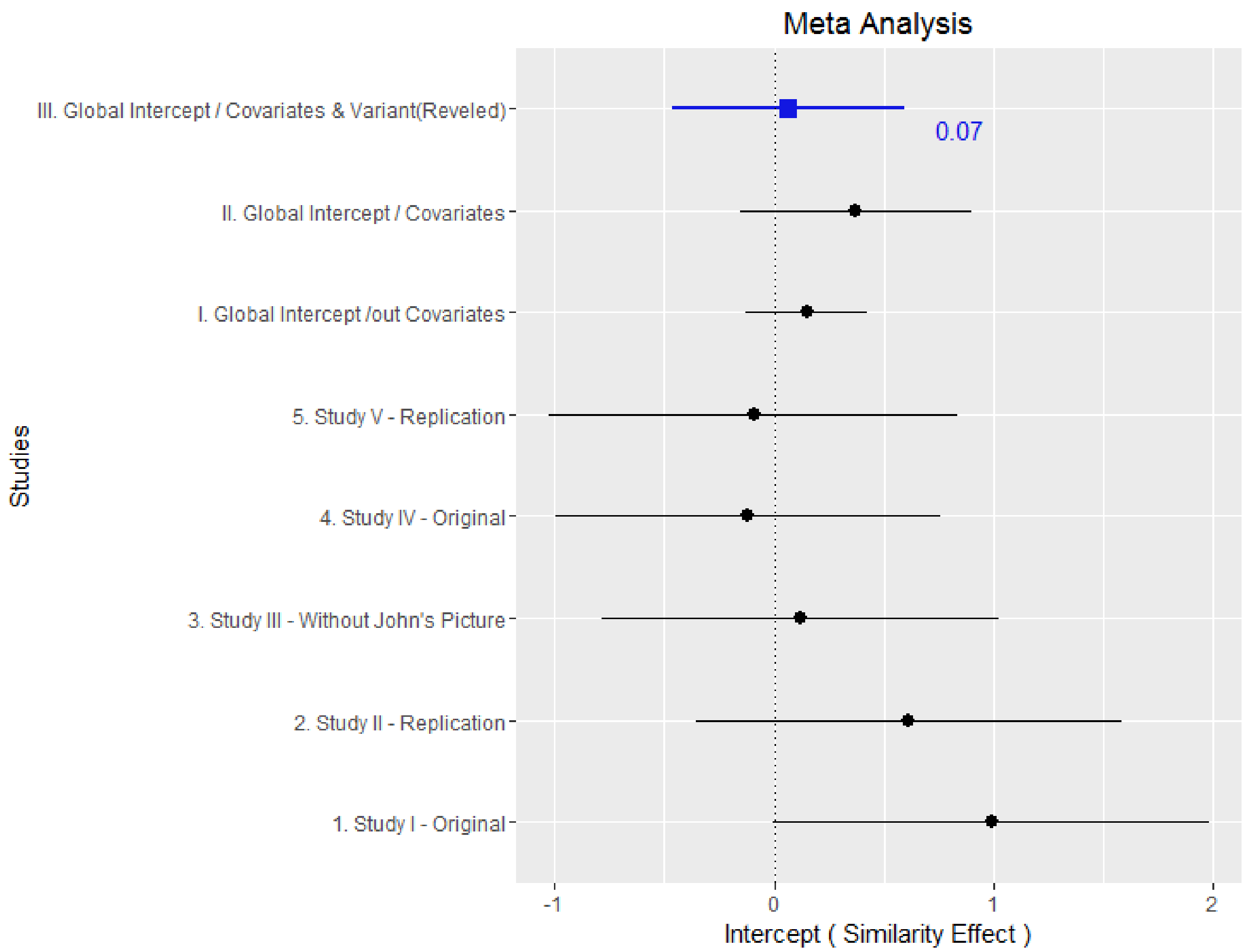

Figure 13: Mating Meta-Analysis Forest Plot - All Studies

The fact that our manipulation of similarity does not become significant even when context is accounted, raises a few questions and also shows us that our participants' decisions are clearly not influenced by the same perceptual matching bias or at the very least not to the same extent as in our nepotism series (see figure). One possibility is that our participants' predictions are, in part, significantly influenced by a preference for homophily that only fully reveals itself only in the positive context while being obfuscated in the negative context. This would be in line with predictions based on the reason-based choice framework proposed by Shafir et al. (1993). Previous research has consistently shown that there is a preference for homophily when individuals are asked to seek a positive reason to pair individuals (Griffiths et al., 1973; Zajonc et al., 1987; Hinsz, 1989; Bereczkei et al., 2002; Bereczkei et al., 2004). Conversely, this means that homophily will not be used as a "reason" to reject certain pairings, precisely because homophily is seen as positive factor.

Through these notions our participants' predictions start to become a lot clearer. Separating our data set by context should allow us to see precisely these interactions at work (see figure). In the positive context (4), using the same model as before, ChoiceSimilar IntellCard + (1|Study) + Sex + Age+ Youngest, without the Variant variable, we can clearly see the significant and positive ($p>0.1$) effect of our similarity manipulation. This shows that our participants are influenced by the similarity between the potential partners and is in line with the predictions made using the reason-based choice framework as our participants are using our similarity manipulations and by extension their preference for homophily as one of the main positive reasons behind their prediction, this is not the only possible explanation. Only looking at the positive context experiments does not allow us to disentangle origin of these findings as this could also be the same perceptual matching bias we saw in the nepotism series. However, as the results of the combined meta-analysis have shown, this is a very unlikely proposition. As it stands, a reason-based choice of our participants could be the foundation of their predictions.

Switching our focus on the analyzing the negative context (5) experiments in this series reveals a drastically different set of results. In this case our model reveals that there is in fact no significant ($p<0.1$) effect resulting from our manipulation of similarity and participants are conversely predicting the mate rejection choices seemingly at random. From the onset we can clearly see that this result is not in line with any predictions one might make based on the presence of perceptual matching bias as the effect of our manipulation should have maintained its significance while inverting its trend, signifying a preference for the prediction of the rejection of the similar mate (see table 51 and figure 14).

However, these results are in line with a reason-based choice framework, as our participants are searching for a positive reason upon which to base their choice in the positive valence and finding one in the form of a preference for homophily when paring individuals. On the other hand, when presented with the negative context, the search for a reason to base their rejection prediction reveals nothing that significantly impacts their choice. As our methodology is specifically designed to create an environment which we can eliminate any confounding variables that might direct our participants in one way or the other, we clearly show that they are not searching for positive reason to choose whom not to reject as this would imply that they would naturally be inclined to select Emily even without any effect of our manipulation of similarity as the face stimuli that represents Emily was revealed to be slightly more attractive than that of Anna in our pretest. Thus, if we would be dealing again with a perception matching bias or even an artifact of our manipulation technique we would

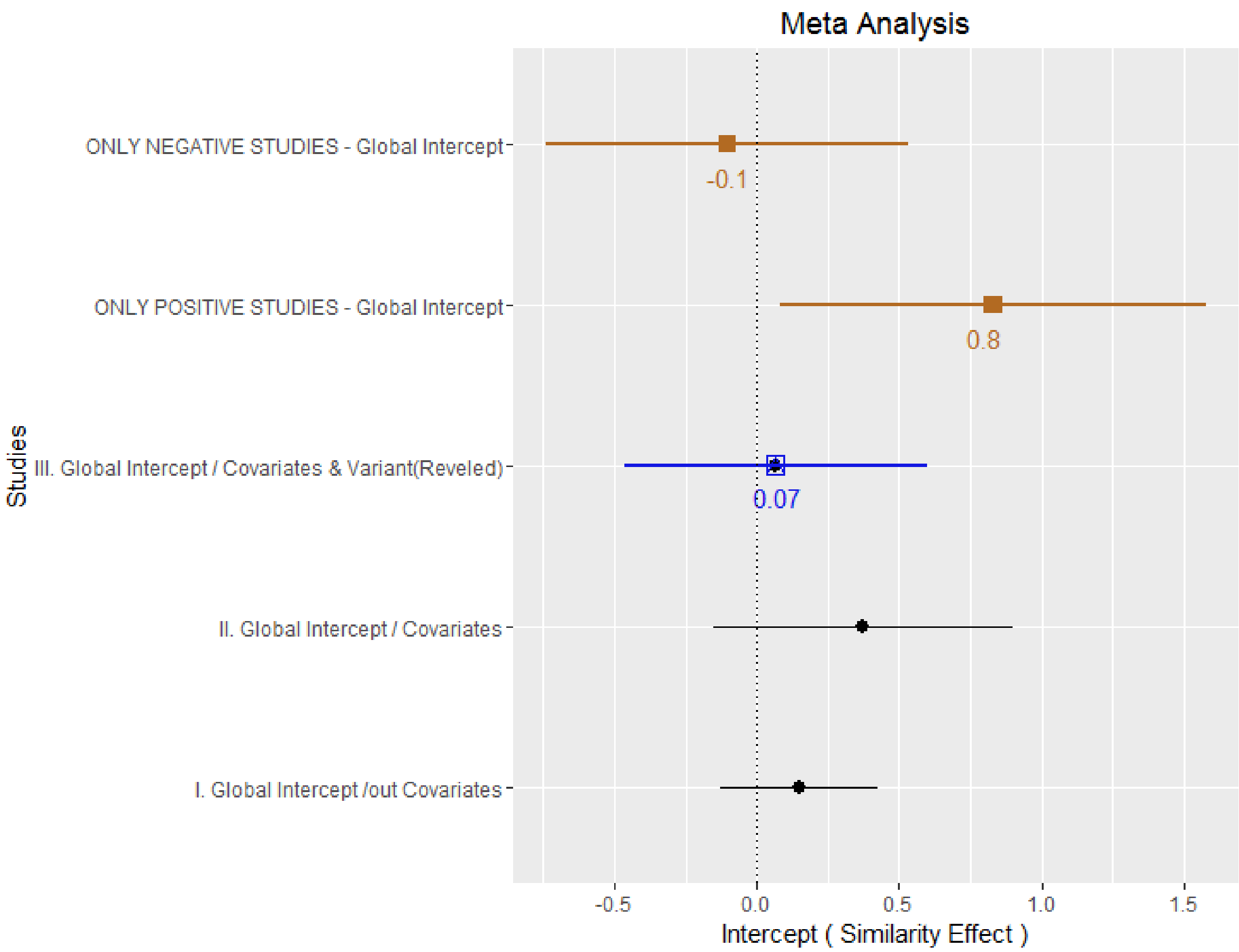

Figure 14: Mating Meta-Analysis Forest Plot - Context Analysis

have clearly seen this in the results. However, we have not seen such a trend and participants are not using straightforward context independent preference to direct their predictions. Instead they are searching for positive reasons for their predictions of Jack's mate choice in the positive context and negative reasons of their rejection predictions in the negative context where they find none. Of course this reason-based choice framework is a post-hoc interpretation of these results and would need further experiments that are specifically designed to test it. One possibility would be to conduct a follow-up study in which the same procedure is used with one exception. Implementing a negative manipulation of similarity. Previous research has shown that participants are able to detect negative similarity (Krupp et al., 2012) which gives us great leeway in constructing this alternate manipulation strategy. By using this variant we would expect the opposite trend with a significant effect in the negative context. Participants would search for a "reason" to predict the rejection of one mate and they would find it in the negative similarity manipulation we would have implemented.

Likewise, a case can also be made that these results reflect the fact that the perceptual matching bias we observed in our nepotism series of studies is not affecting our participants in the same way here, as a result of the cross-sex nature of the stimuli being presented to them. This would explain our null finding in the global meta-analysis and asymmetry of our results with the significant result in our positive variants being simply a statistical fluke. While either explanation has merits, the important take away message has to remain that our data here offer a clear and robust set of results that undermines the notion that a kin recognition mechanism is influencing our participants predictions. As with our nepotism series, these studies offer a clear and concise requirement for any future experiments to fully integrate and take into account the fact that manipulations of similarity may not be the putative cue of relatedness at least not in third-party scenarios.

Table 51: Mating Multi Level Regression Models Results - Context Analysis

	Dependent variable:	
	ChoiceSimilar	
	(Positive)	(Negative)
Constant	0.83*	−0.10
	(0.38)	(0.32)
IntelligenceSimilar	0.27**	0.03
	(0.09)	(0.09)
SexWomen	0.04	0.12
	(0.20)	(0.18)
Age	−0.01	−0.004
	(0.01)	(0.01)
YoungestNo	−0.09	0.23
	(0.19)	(0.18)
Observations	400	396
Log Likelihood	−806.08	−780.29
Akaike Inf. Crit.	1,630.16	1,578.58
Bayesian Inf. Crit.	1,666.08	1,614.41
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$	

8 CONCLUSION

"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" (I found it!) but "That's funny...""

— Isaac Asimov

This research started from a desire to expand our understanding of kin recognition and see if this ability extends to third-party scenarios. A great deal of attention has been directed towards using manipulations of similarity to investigate effects of kin recognition. The most successful of these studies used participants own faces in order to create facial stimuli that resemble themselves. These manipulated faces are then used in a number of scenarios to investigate if the participants perception of this similarity affects their choices and behavior. This type of direct kin recognition scenario, using self-resembling stimuli, have brought facial similarity to the forefront of kin recognition research. Revealing that the judgments that participants make as to the relatedness of facial stimuli, are associated with actual kinship (Kaminski et al., 2009) and in turn, facial similarity predicts these judgments of relatedness (Maloney et al., 2006; DeBruine et al., 2009). In addition, researchers like DeBruine (2002) and DeBruine (2005) and Krupp et al. (2008) have shown that in accordance with predictions one might make based on kin selection, participants are likely to be more trusting and cooperative with self-resembling interaction partners. When mating choice comes into play participants seem to find opposite-sex faces that resemble themselves as being less attractive (DeBruine, 2005; DeBruine et al., 2011).

One major theme flowing through this thesis has been the importance of acknowledging that categorization must be in service of action. Research using self-resemblance has provided a wealth of data regarding the nature of the effect of similarity manipulations on participants behavior and choices. Showing that manipulations of similarity affect participants behavior in context-dependent ways, both in terms of domain (i.e., prosocial vs. mating) and valence. Suggesting that by all accounts, in direct kin recognition experiments (i.e., using self-morphs or self-transforms), facial similarity is deserving of its status as a putative cue of relatedness.

A number of studies have also tried to use the same type of facial stimuli in third-party scenarios. Using these procedures, researchers have seemingly uncovered an ability of individuals to recognize others unrelated to themselves as kin (e.g., Maloney et al., 2006; DeBruine et al., 2009; Alvergne et al., 2007; Alvergne et al., 2009b; Alvergne et al., 2009a). Studies such as those of Maloney et al. (2006) or DeBruine et al. (2009), both consistently find a high correlation (.92) between perceptions of similarity and perceptions of kinship in third parties. Additionally, the two studies from 2009 of Alvergne and their colleagues show that unrelated third-parties can and will use facial appearance to accurately determine kin relationships both within their own culture and cross-culturally. Thus, both of our pretests and the literature reinforce this notion that, humans are able to detect similarity in third-parties. As such, our research does not dispute the presence of a perceptual mechanism responsible to detecting similarity between facial stimuli. Though, throughout our discussion here we keep coming back to one important aspect of kin recognition. Effects of kin recognition are bound by requirements of kin selection. And, as similarity undeniably captures some aspect of genetic relatedness, an ability to detect similarity does not, by itself, inform us in regards to the presence of an ability for kin recognition. *Categorization must serve action*. So, what would strengthen the notion of a presence of third-party kin recognition in humans? Facial similarity would need to have a context-dependent effect on participants predictions, susceptible to valence changes in scenarios and switches from the prosocial and mate choice domains.

This is precisely what we set out to do with our two lines of research. Our literature review revealed that when context is starved participants seem to able to detect similarity and seemingly connect it to relatedness. On the other hand, our nepotism series of experiments, by re-inserting context, offers us a different conclusion altogether. Within scenarios in which valence is modified and our participants analysis is bounded by predictions made by kin selection, their choices do no reflect a connection between similarity and relatedness. What they do point towards is a perceptual matching bias affecting their predictions. The direction of the effect of our manipulation of similarity is a positive one in all cases in our meta-analysis and significant when analyzing the negative context studies ($p<0.05$). The case of the negative variants is highly important as the positive direction and significance fundamentally lays to rest any notion that our effects could be tied to a kin recognition mechanism. The fact that our participants are predicting the firing of the similar candidate is in stark contrast to predictions of nepotistic and kin directed behavior that one would expect if the connection between similarity and relatedness was made.

Conversely, in our series of studies in the domain of mate choice, our participants seem to be influenced by a preference for homophily that fully reveals itself only in the positive context while being obfuscated in the negative context. This would be in line with predictions based on the reason-based choice framework proposed by Shafir et al. (1993). As previous research has consistently shown that there is a preference for homophily when individuals are asked to seek a positive reason to pair individuals (Griffiths et al., 1973; Zajonc et al., 1987; Hinsz, 1989; Bereczkei et al., 2002; Bereczkei et al., 2004). Of course this reason-based choice framework is a post-hoc interpretation of these results and would need further experiments that are specifically designed to test it. Likewise, a case can also be made that these results reflect the fact that the perceptual matching bias we observed in our nepotism series of studies is not affecting our participants in the same way here, as a result of the cross-sex nature of the stimuli being presented to them. This would explain our null finding in the global meta-analysis and asymmetry of our results with the significant result in our positive variants being simply a statistical fluke. Again, the important take away message has to remain that our data here offers a clear and robust set of results that undermines the notion that a kin recognition mechanism is influencing our participants' predictions.

Both of these lines of research bring us closer to fulfilling the fundamental need to disentangle similarity from relatedness. Our research through its 16 separate studies, investigating both the topics of nepotistic behavior and mating choice while presenting positive and negative valence scenarios offers a robust set of results that contest the status of facial similarity as a putative cue of relatedness in third-parties. In fact our results in the nepotism series of experiment provide a clear finding of a perceptual matching bias, while our mating series results points towards a reason-based framework in which participants are searching for a positive reason upon which to base their choice in the positive valence scenarios and finding one in the form of a preference for homophily when paring individuals.

The prospect for future research investigating the possible implications of a perceptual matching bias or a preference for homophily presents a tantalizing endeavor.

Though, the major strength of this work is the methodological improvements and changes that *must be made* as a result in this field. Given our findings, future studies using facial similarity manipulations must at the very least include positive and negative variants, to rule out perceptual matching bias effects. Additionally, our results in the mating series strongly support a need for further investigation on the effect of manipulations of similarity in third-party scenarios to disentangle a possible preference for homophily as a result of similarity manipulations. Finally, an understanding of the differences between effects of kin recognition in domains of prosociality and mate choice must be integrated within any framework of future experimental endeavors.

BIBLIOGRAPHY

Alberts, Susan C (1999). "Paternal kin discrimination in wild baboons." In: *Proceedings of the Royal Society of London B: Biological Sciences* 266.1427, pp. 1501–1506 (cit. on p. 22).

Alvergne, Alexandra, Charlotte Faurie, and Michel Raymond (2007). "Differential facial resemblance of young children to their parents: who do children look like more?" In: *Evolution and Human behavior* 28.2, pp. 135–144 (cit. on pp. 85, 124).

Alvergne, Alexandra, Ryo Oda, Charlotte Faurie, Akiko Matsumoto-Oda, Valérie Durand, and Michel Raymond (2009a). "Cross-cultural perceptions of facial resemblance between kin." In: *Journal of Vision* 9.6, pp. 23–23 (cit. on pp. 28, 85, 124).

Alvergne, Alexandra, Charlotte Faurie, and Michel Raymond (2009b). "Father–offspring resemblance predicts paternal investment in humans." In: *Animal Behaviour* 78.1, pp. 61–69 (cit. on pp. 85, 124).

Amato, Paul R and Bruce Keith (1991). *Parental divorce and the well-being of children: a meta-analysis.* (Cit. on p. 79).

André, Jean-Baptiste and Nicolas Baumard (2011). "The evolution of fairness in a biological market." In: *Evolution* 65.5, pp. 1447–1456 (cit. on p. 25).

Antfolk, Jan, Mira Karlsson, Anna Bäckström, and Pekka Santtila (2012). "Disgust elicited by third-party incest: the roles of biological relatedness, co-residence, and family relationship." In: *Evolution and Human Behavior* 33.3, pp. 217–223 (cit. on p. 31).

Axelrod, R (1984). *The Evolution of Cooperation* (cit. on p. 12).

Baumard, Nicolas, Jean-Baptiste André, and Dan Sperber (2013). "A mutualistic approach to morality: The evolution of fairness by partner choice." In: *Behavioral and Brain Sciences* 36.1, pp. 59–78 (cit. on pp. 12, 25).

Baumard, Nicolas and Mark Sheskin (2015). "Partner choice and the evolution of a contractualist morality." In: *The moral brain: a multidisciplinary perspective* 20 (cit. on p. 26).

Bennedsen, Morten, Kasper Meisner Nielsen, Francisco Pérez-González, and Daniel Wolfenzon (2007). "Inside the family firm: The role of families in succession decisions and performance." In: *The Quarterly Journal of Economics* 122.2, pp. 647–691 (cit. on p. 30).

Bereczkei, T, P Gyuris, P Koves, and L Bernath (2002). "Homogamy, genetic similarity, and imprinting; parental influence on mate choice preferences." In: *Personality and Individual Differences* 33.5, pp. 677–690 (cit. on pp. 117, 125).

Bereczkei, Tamas, Petra Gyuris, and Glenn E Weisfeld (2004). "Sexual imprinting in human mate choice." In: *Proceedings of the Royal Society B: Biological Sciences* 271.1544, p. 1129 (cit. on pp. 117, 125).

Birch, Jonathan and Samir Okasha (2014). "Kin selection and its critics." In: *BioScience* 65.1, pp. 22–32 (cit. on p. 17).

Blaustein, Andrew R, Marc Bekoff, and Thomas J Daniels (1987). "Kin recognition in vertebrates (excluding primates): empirical evidence." In: *Kin recognition in animals*, pp. 287–331 (cit. on p. 17).

Bressan, Paola and Guendalina Zucchi (2009). "Human kin recognition is self-rather than family-referential." In: *Biology Letters* 5.3, pp. 336–338 (cit. on p. 23).

Busse, Curt D (1978). "Do chimpanzees hunt cooperatively?" In: *The American Naturalist* 112.986, pp. 767–770 (cit. on p. 8).

Clutton-Brock, Tim (2002). "Breeding together: kin selection and mutualism in cooperative vertebrates." In: *Science* 296.5565, pp. 69–72 (cit. on p. 12).

Daly, Martin, Margo Wilson, et al. (1988). "The Darwinian psychology of discriminative parental solicitude." In: *Nebraska symposium on motivation*. Vol. 35, pp. 91–144 (cit. on p. 28).

Darwin, Charles (1871). "The descent ofman." In: *The Great Books ofthe Western World* 49, p. 320 (cit. on p. 9).

Dasgupta, Nilanjana, Mahzarin R Banaji, and Robert P Abelson (1999). "Group entitativity and group perception: Associations between physical features and psychological judgment." In: *Journal of personality and social psychology* 77.5, p. 991 (cit. on p. 30).

Dawkins, Richard et al. (1976). "selfish gene." In: (cit. on p. 23).

DeBruine, L. M. (2002). "Facial resemblance enhances trust." In: *Proceedings of the Royal Society B: Biological Sciences* 269.1498, pp. 1307–1312. ISSN: 0962-8452, 1471-2954. DOI: 10.1098/rspb.2002.2034. (Visited on 10/03/2013) (cit. on pp. 27, 31, 123).

DeBruine, L. M (2005). "Trustworthy but not lust-worthy: context-specific effects of facial resemblance." In: *Proceedings of the Royal Society B: Biological Sciences* 272.1566, pp. 919–922. ISSN: 0962-8452, 1471-2954. DOI: 10.1098/rspb.2004.3003. (Visited on 10/03/2013) (cit. on pp. 27, 31, 123).

DeBruine, L. M., B. C. Jones, C. D. Watkins, S. C. Roberts, A. C. Little, F. G. Smith, and M. C. Quist (2011). "Opposite-sex siblings decrease attraction, but not prosocial attributions, to self-resembling opposite-sex faces." In: *Proceedings of the National Academy of Sciences* 108.28, pp. 11710–11714. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1105919108. (Visited on 10/03/2013) (cit. on pp. 27, 123).

DeBruine, Lisa M. (2004). "Facial resemblance increases the attractiveness of same–sex faces more than other–sex faces." In: *Proceedings* *of the Royal Society of London. Series B: Biological Sciences* 271.1552, pp. 2085–2090. (Visited on 10/03/2013) (cit. on p. 35).

DeBruine, Lisa M, Finlay G Smith, Benedict C Jones, S Craig Roberts, Marion Petrie, and Tim D Spector (2009). "Kin recognition signals in adult faces." In: *Vision research* 49.1, pp. 38–43 (cit. on pp. 27, 28, 44, 85, 123, 124).

Ebner, Natalie C (2008). "Age of face matters: Age-group differences in ratings of young and old faces." In: *Behavior Research Methods* 40.1, pp. 130–136 (cit. on p. 42).

Emlen, Stephen T (1997). "Predicting family dynamics in social vertebrates." In: *Behavioural ecology: an evolutionary approach* 4, pp. 228–253 (cit. on p. 12).

Fairbanks, Lynn A (1990). "Reciprocal benefits of allomothering for female vervet monkeys." In: *Animal Behaviour* 40.3, pp. 553–562 (cit. on p. 11).

Fehr, Ernst and Klaus M Schmidt (2006). "The economics of fairness, reciprocity and altruism–experimental evidence and new theories." In: *Handbook of the economics of giving, altruism and reciprocity* 1, pp. 615–691 (cit. on p. 25).

Fessler, Daniel MT and C David Navarrete (2004). "Third-party attitudes toward sibling incest: Evidence for Westermarck's hypotheses." In: *Evolution and Human Behavior* 25.5, pp. 277–294 (cit. on pp. 28, 31, 100).

Frank, Steven A (1998). *Foundations of social evolution*. Princeton University Press (cit. on p. 18).

Gardner, Andy and Stuart A West (2004). "Cooperation and punishment, especially in humans." In: *The American Naturalist* 164.6, pp. 753–764 (cit. on p. 12).

Gardner, Andy and SA West (2006). "Demography, altruism, and the benefits of budding." In: *Journal of evolutionary biology* 19.5, pp. 1707–1716 (cit. on p. 18).

Gardner, Andy, Stuart A West, and Geoff Wild (2011). "The genetical theory of kin selection." In: *Journal of evolutionary biology* 24.5, pp. 1020–1043 (cit. on p. 17).

Goodnight, Keith F (1992). "The effect of stochastic variation on kin selection in a budding-viscous population." In: *The American Naturalist* 140.6, pp. 1028–1040 (cit. on p. 18).

Gould, Stephen J (1976). "Biological potential vs. biological determinism." In: *Natural History* 85.5, pp. 18–20 (cit. on p. 21).

Grafen, Alan (1984). "Natural selection, kin selection and group selection." In: *Behavioural ecology: An evolutionary approach* 2, pp. 62–84 (cit. on p. 18).

Griffin, Ashleigh S and Stuart A West (2002). "Kin selection: fact and fiction." In: *Trends in Ecology & Evolution* 17.1, pp. 15–21 (cit. on p. 18).

Griffiths, R Wayne and Phillip R Kunz (1973). "Assortative mating: A study of physiognomic homogamy." In: *Social biology* 20.4, pp. 448–453 (cit. on pp. 117, 125).

Grosberg, Richard K and James F Quinn (1986). "The genetic control and consequences of kin recognition by the larvae of a colonial marine invertebrate." In: *Nature* 322.6078, pp. 456–459 (cit. on p. 15).

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982). "An experimental analysis of ultimatum bargaining." In: *Journal of economic behavior & organization* 3.4, pp. 367–388 (cit. on p. 25).

Hamilton, W. D. (1964). "The genetical evolution of social behaviour." In: *Journal of Theoretical Biology* 7.1, pp. 1–52 (cit. on pp. 1, 15).

Hauber, Mark E, Paul W Sherman, and Dóra Paprika (2000). "Self-referent phenotype matching in a brood parasite: the armpit effect in brown-headed cowbirds (Molothrus ater)." In: *Animal Cognition* 3.2, pp. 113–117 (cit. on p. 22).

Hauber, Mark E and Paul W Sherman (2001). "Self-referent phenotype matching: theoretical considerations and empirical evidence." In: *Trends in neurosciences* 24.10, pp. 609–616 (cit. on p. 41).

Hauert, Christoph and Michael Doebeli (2004). "Spatial structure often inhibits the evolution of cooperation in the snowdrift game." In: *Nature* 428.6983, p. 643 (cit. on p. 18).

Heth, Giora, Josephine Todrank, and Robert E Johnston (1998). "Kin recognition in golden hamsters: evidence for phenotype matching." In: *Animal Behaviour* 56.2, pp. 409–417 (cit. on p. 22).

Hill, Emily A and N. J. S. Christenfeld (1995). "Whose baby are you?" In: *Nature* 378, p. 14 (cit. on p. 85).

Hinsz, Verlin B (1989). "Facial resemblance in engaged and married couples." In: *Journal of Social and Personal Relationships* 6.2, pp. 223–229 (cit. on pp. 117, 125).

Hintze, Arend and Ralph Hertwig (2016). "The Evolution of Generosity in the Ultimatum Game." In: *Scientific reports* 6, p. 34102 (cit. on p. 26).

Holmes, Warren G and Paul W Sherman (1982). "The ontogeny of kin recognition in two species of ground squirrels1." In: *American Zoologist* 22.3, pp. 491–517 (cit. on p. 20).

Ip, Wai-man, Chi-yue Chiu, Ching Wan, et al. (2006). "Birds of a feather and birds flocking together: physical versus behavioral cues may lead to trait-versus goal-based group perception." In: *Journal of personality and social psychology* 90.3, p. 368 (cit. on p. 30).

Kaminski, Gwenaël, Slimane Dridi, Christian Graff, and Edouard Gentaz (2009). "Human ability to detect kinship in strangers' faces: effects of the degree of relatedness." In: *Proceedings of the Royal Society of London B: Biological Sciences* 276.1670, pp. 3193–3200 (cit. on pp. 27, 123).

Kaplan, Hillard S, Paul L Hooper, and Michael Gurven (2009). "The evolutionary and ecological roots of human social organization." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364.1533, pp. 3289–99. DOI: 10.1098/rstb.2009.0115 (cit. on p. 7).

Kaplan, Hillard, Michael Gurven, Kim Hill, and A Magdalena Hurtado (2005). "The natural history of human food sharing and cooperation: a review and a new multi-individual approach to the negotiation of norms." In: *Moral sentiments and material interests: The foundations of cooperation in economic life*, pp. 75–113 (cit. on p. 13).

Keller, Laurent and Kenneth G Ross (1998). "Selfish genes: a green beard in the red fire ant." In: *Nature* 394.6693, p. 573 (cit. on p. 23).

Kelly, John K (1992). "Restricted migration and the evolution of altruism." In: *Evolution* 46.5, pp. 1492–1495 (cit. on p. 18).

Kelly, John K (1994). "The effect of scale dependent processes on kin selection: mating and density regulation." In: *Theoretical population biology* 46.1, pp. 32–57 (cit. on p. 18).

Killingback, Timothy, Jonas Bieri, and Thomas Flatt (2006). "Evolution in group-structured populations can resolve the tragedy of the commons." In: *Proceedings of the Royal Society of London B: Biological Sciences* 273.1593, pp. 1477–1481 (cit. on p. 18).

Krebs, John R, Nicholas B Davies, and Jan Parr (1993). *An introduction to behavioural ecology*. Blackwell Scientific Publications (cit. on p. 12).

Krupp, D. B., L. M. DeBRUINE, B. C. Jones, and M. L. LalumièRe (2012). "Kin recognition: evidence that humans can perceive both positive and negative relatedness: Humans recognize negative relatives." In: *Journal of Evolutionary Biology* 25.8, pp. 1472–1478. ISSN: 1010061X. DOI: 10.1111/j.1420-9101.2012.02553.x. (Visited on 10/03/2013) (cit. on p. 119).

Krupp, D, L Debruine, and P Barclay (2008). "A cue of kinship promotes cooperation for the public good." In: *Evolution and Human Behavior* 29.1, pp. 49–55. ISSN: 10905138. DOI: 10.1016/j.evolhumbehav.2007.08.002. (Visited on 10/03/2013) (cit. on pp. 27, 41, 123).

Krupp, Daniel Brian, Lisa M. DeBruine, and Benedict C. Jones (2011). "Cooperation and Conflict in the Light of 20 Kin Recognition Systems." In: *The Oxford Handbook of Evolutionary Family Psychology*, p. 345. (Visited on 10/01/2013) (cit. on pp. 19, 22).

Kümmerli, Rolf, Andy Gardner, Stuart A West, and Ashleigh S Griffin (2009). "Limited dispersal, budding dispersal, and cooperation: an experimental study." In: *Evolution* 63.4, pp. 939–949 (cit. on p. 18).

Langlois, Judith H and Lori A Roggman (1990). "Attractive faces are only average." In: *Psychological science* 1.2, pp. 115–121 (cit. on p. 35).

Lee, PC (1987). "Allomothering among African elephants." In: *Animal Behaviour* 35.1, pp. 278–291 (cit. on p. 11).

Lehmann, Laurent and Laurent Keller (2006). "The evolution of cooperation and altruism–a general framework and a classification of models." In: *Journal of evolutionary biology* 19.5, pp. 1365–1376 (cit. on p. 18).

Lieberman, Debra, John Tooby, and Leda Cosmides (2003). "Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest." In: *Proceedings of the Royal Society of London B: Biological Sciences* 270.1517, pp. 819–826 (cit. on pp. 31, 100).

Lieberman, Debra, John Tooby, and Leda Cosmides (2007). "The architecture of human kin detection." In: *Nature* 445.7129, p. 727 (cit. on pp. 26, 28).

Lieberman, Debra, Robert Oum, and Robert Kurzban (2008). "The family of fundamental social categories includes kinship: Evidence from the memory confusion paradigm." In: *European Journal of Social Psychology* 38.6, pp. 998–1012 (cit. on pp. 27–29).

Ma, Debbie S, Joshua Correll, and Bernd Wittenbrink (2015). "The Chicago face database: A free stimulus set of faces and norming data." In: *Behavior research methods* 47.4, pp. 1122–1135 (cit. on p. 33).

Maloney, Laurence T and Maria F Dal Martello (2006). "Kin recognition and the perceived facial similarity of children." In: *Journal of Vision* 6.10, pp. 4–4 (cit. on pp. 27, 28, 44, 85, 123, 124).

Mateo, Jill M and Robert E Johnston (2000). "Kin recognition and the 'armpit effect': evidence of self–referent phenotype matching." In: *Proceedings of the Royal Society of London B: Biological Sciences* 267.1444, pp. 695–700 (cit. on p. 22).

Minear, Meredith and Denise C Park (2004). "A lifespan database of adult facial stimuli." In: *Behavior Research Methods, Instruments, & Computers* 36.4, pp. 630–633 (cit. on p. 42).

Mitteldorf, Joshua and David Sloan Wilson (2000). "Population viscosity and the evolution of altruism." In: *Journal of Theoretical Biology* 204.4, pp. 481–496 (cit. on p. 18).

Moore, Harry, Katerina Dvorakova, Nicholas Jenkins, and William Breed (2002). "Exceptional sperm cooperation in the wood mouse." In: *Nature* 418.6894, p. 174 (cit. on p. 23).

Murray, Martyn G and Russell Gerrard (1984). "Conflict in the neighbourhood: models where close relatives are in direct competition." In: *Journal of theoretical biology* 111.2, pp. 237–246 (cit. on p. 18).

Nowak, Martin A and Robert M May (1992). "Evolutionary games and spatial chaos." In: *Nature* 359.6398, pp. 826–829 (cit. on p. 18).

Nowak, Martin A, Sebastian Bonhoeffer, and Robert M May (1994). "Spatial games and the maintenance of cooperation." In: *Proceedings of the National Academy of Sciences* 91.11, pp. 4877–4881 (cit. on p. 18).

Nowak, Martin A, Karen M Page, and Karl Sigmund (2000). "Fairness versus reason in the ultimatum game." In: *Science* 289.5485, pp. 1773–1775 (cit. on p. 25).

Nowak, Martin A, Corina E Tarnita, and Edward O Wilson (2010). "The evolution of eusociality." In: *Nature* 466.7310, pp. 1057–1062 (cit. on p. 17).

Ohtsuki, Hisashi, Christoph Hauert, Erez Lieberman, and Martin A Nowak (2006). "A simple rule for the evolution of cooperation on graphs." In: *Nature* 441.7092, p. 502 (cit. on p. 18).

Pollock, Gregory B (1983). "Population viscosity and kin selection." In: *The American Naturalist* 122.6, pp. 817–829 (cit. on p. 18).

Puurtinen, Mikael (2011). "Mate choice for optimal (k) inbreeding." In: *Evolution* 65.5, pp. 1501–1505 (cit. on p. 11).

Puurtinen, Mikael, Tarmo Ketola, and JS Kotiaho (2009). "The good-genes and compatible-genes benefits of mate choice." In: *The American Naturalist* 174.5, pp. 741–752 (cit. on p. 11).

Queller, David C (1992). "Does population viscosity promote kin selection?" In: *Trends in Ecology & Evolution* 7.10, pp. 322–324 (cit. on p. 18).

Queller, David C and Keith F Goodnight (1989). "Estimating relatedness using genetic markers." In: *Evolution* 43.2, pp. 258–275 (cit. on p. 18).

Ratnieks, FLW (2006). "The evolution of cooperation and altruism: the basic conditions are simple and well known." In: *Journal of evolutionary biology* 19.5, pp. 1413–1414 (cit. on p. 12).

Reid, Vincent M, Kirsty Dunn, Robert J Young, Johnson Amu, Tim Donovan, and Nadja Reissland (2017). "The human fetus preferentially engages with face-like visual stimuli." In: *Current Biology* (cit. on p. 22).

Richards, Robert J (1987). "Darwin and the emergence of evolutionary theories of mind and behavior." In: (cit. on p. 9).

Rosa-Salva, Orsola, Lucia Regolin, and Giorgio Vallortigara (2010). "Faces are special for newly hatched chicks: evidence for inborn domain-specific mechanisms underlying spontaneous preferences for face-like stimuli." In: *Developmental science* 13.4, pp. 565–577 (cit. on p. 21).

Rousset, François (2013). *Genetic structure and selection in subdivided populations (MPB-40)*. Princeton University Press (cit. on p. 18).

Sachs, Joel L, Ulrich G Mueller, Thomas P Wilcox, and James J Bull (2004). "The evolution of cooperation." In: *The Quarterly Review of Biology* 79.2, pp. 135–160 (cit. on pp. 12–15).

Santos, Fernando P, Francisco C Santos, Ana Paiva, and Jorge M Pacheco (2015). "Evolutionary dynamics of group fairness." In: *Journal of theoretical biology* 378, pp. 96–102 (cit. on p. 29).

Shafir, Eldar, Itamar Simonson, and Amos Tversky (1993). "Reason-based choice." In: *Cognition* 49.1, pp. 11–36 (cit. on pp. 117, 125).

Sherman, Paul W (1977). "Nepotism and the evolution of alarm calls." In: *Science* 197.4310, pp. 1246–1253 (cit. on p. 11).

Shorey, Lisa, Stuart Piertney, Jon Stone, and Jacob Hoglund (2000). "Fine-scale genetic structuring on Manacus manacus leks." In: *Nature* 408.6810, p. 352 (cit. on p. 22).

Silk, Joan B (2002). "Kin selection in primate groups." In: *International Journal of Primatology* 23.4, pp. 849–875 (cit. on p. 17).

Taylor, Peter D (1992a). "Altruism in viscous populations—an inclusive fitness model." In: *Evolutionary ecology* 6.4, pp. 352–356 (cit. on p. 18).

Taylor, Peter D (1992b). "Inclusive fitness in a homogeneous environment." In: *Proceedings of the Royal Society of London B: Biological Sciences* 249.1326, pp. 299–302 (cit. on p. 18).

Teleki, Geza (1973). *The predatory behavior of wild chimpanzees*. Bucknell University Press (cit. on p. 8).

Tiddeman, Bernard, Michael Burt, and David Perrett (2001). "Prototyping and transforming facial textures for perception research." In: *IEEE computer graphics and applications* 21.5, pp. 42–50 (cit. on pp. 31, 35).

Tomasello, M. (2009). *Why We Cooperate*. Boston Review Books. MIT Press. ISBN: 978-0-262-25849-4 (cit. on p. 7).

Traulsen, Arne and Martin A Nowak (2006). "Evolution of cooperation by multilevel selection." In: *Proceedings of the National Academy of Sciences* 103.29, pp. 10952–10955 (cit. on p. 18).

Trivers, Robert L (1971). "The evolution of reciprocal altruism." In: *The Quarterly review of biology* 46.1, pp. 35–57 (cit. on p. 12).

Van Baalen, Minus and David A Rand (1998). "The unit of selection in viscous populations and the evolution of altruism." In: *Journal of theoretical biology* 193.4, pp. 631–648 (cit. on p. 18).

Van Schaik, Carel P and Charles H Janson (2000). *Infanticide by males and its implications*. Cambridge University Press (cit. on p. 18).

Vemer, Elizabeth, Marilyn Coleman, Lawrence H Ganong, and Harris Cooper (1989). "Marital satisfaction in remarriage: A meta-analysis." In: *Journal of Marriage and the Family*, pp. 713–725 (cit. on p. 79).

Wagner, Michael and Bernd Weiß (2006). "On the variation of divorce risks in Europe: Findings from a meta-analysis of European longi-

tudinal studies." In: *European sociological review* 22.5, pp. 483–500 (cit. on p. 79).
Waldorf, Brigitte and Pillsung Byun (2005). "Meta-analysis of the impact of age structure on fertility." In: *Journal of Population Economics* 18.1, pp. 15–40 (cit. on p. 79).
Wang, Ze, Xin He, and Fan Liu (2016). "From similitude to success: The effects of facial resemblance on perceptions of team effectiveness." In: *Journal of Experimental Psychology: Applied* 22.1, p. 48 (cit. on pp. 30, 41, 52).
Warneken, Felix, Brian Hare, Alicia P Melis, Daniel Hanus, and Michael Tomasello (2007). "Spontaneous altruism by chimpanzees and young children." In: *PLoS biology* 5.7, e184 (cit. on p. 8).
Warneken, Felix and Michael Tomasello (2009). "Varieties of altruism in children and chimpanzees." In: *Trends in cognitive sciences* 13.9, pp. 397–402 (cit. on p. 8).
Weisfeld, Glenn E, Tiffany Czilli, Krista A Phillips, James A Gall, and Cary M Lichtman (2003). "Possible olfaction-based mechanisms in human kin recognition and inbreeding avoidance." In: *Journal of experimental child psychology* 85.3, pp. 279–295 (cit. on p. 31).
West, Stuart A, Ido Pen, and Ashleigh S Griffin (2002). "Cooperation and competition between relatives." In: *Science* 296.5565, pp. 72–75 (cit. on p. 18).
Westermarck, Edward (2006). "The Origin and Development of Moral Ideas (1906)." In: *INSTINCT IN PSYCHOLOGY* 179 (cit. on p. 31).
Wilson, David Sloan, GB Pollock, and Lee A Dugatkin (1992). "Can altruism evolve in purely viscous populations?" In: *Evolutionary ecology* 6.4, pp. 331–341 (cit. on p. 18).
Zajonc, Robert B, Pamela K Adelmann, Sheila T Murphy, and Paula M Niedenthal (1987). "Convergence in the physical appearance of spouses." In: *Motivation and emotion* 11.4, pp. 335–346 (cit. on pp. 117, 125).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Final Version as of September 15, 2017 (`classicthesis`).

Abstract:

Our relation to our kin shapes much of our social world. It's no surprise then, that how we recognize and react to our own kin has been a widely investigated topic. In particular, when tackling direct kin recognition, facial similarity has emerged as a putative cue of relatedness. In this thesis, I investigate whether or not the same can be said for third party kin recognition. Split between two lines of research, we explore individuals' predictions of nepotistic and mating behavior in third party scenarios using facial stimuli. These two domains provide the backbone of our research. *Categorization must serve action*. So, what would strengthen the notion of a presence of third-party kin recognition in humans? Facial similarity *must have* a context-dependent effect on participants predictions, susceptible to valence changes in scenarios and switches from the prosocial and mate choice domains. This is precisely what we set out to do with our two lines of research. Though our literature review revealed that when context is starved participants seem to be able to detect similarity and seemingly connect it to relatedness. Our nepotism and mating series of experiments, by re-inserting context, offers us a different conclusion altogether. Within scenarios in which valence is modified and our participants analysis is bounded by predictions made by kin selection, their choices do no reflect a connection between similarity and relatedness.

Keywords: *kin recogntion, third parties, inclusive fitness, phenotype matching, facial similarity.*

Résumé:

Notre relation avec nos apparentés forme une grande partie de notre monde social; et la façon dont nous reconnaissons et traitons nos apparentés a donné lieu à une importante somme de recherche. Lorsqu'il s'agit de reconnaître un apparenté direct, la similarité faciale est considérée comme un indice d'apparentement. Dans cette thèse, j'étudie si elle joue un rôle comparable lorsqu'il s'agit de reconnaître un apparentement entre des tiers, en menant deux lignes de recherche: les prédictions de comportement népotistiques et les prédictions de préférences de couple, par des tiers, en présence de stimuli faciaux. *La catégorisation devant servir l'action*, la similarité faciale doit avoir un effet dépendant du contexte sur ces prédictions, susceptible à des changements de valence et de domaine. En l'absence de contexte, les individus semblent pouvoir détecter la similarité faciale et la mettre en relation avec l'apparentement. Nos deux séries d'expériences offrent une confusion différente. Quand la valence du contexte change et que nous analysons les prédictions des participants en terme de kin selection, leurs choix ne semblent pas mettre en relation similarité faciale et apparentement.

Mots-Clés: *reconnaissance des apparentés, tierces parties, fitness inclusive, appariement de phénotype, similarité faciale*