No Evidence of an Effect of Resource Necessity and Unpredictability on Cognitive Mechanisms for Detecting Greediness and Stinginess

by

Andres Munoz Castro

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved October 2019 by the
Graduate Supervisory Committee:

Athena Aktipis, Chair
Steven Neuberg
Daniel Hruschka

ARIZONA STATE UNIVERSITY

December 2019

ABSTRACT

Resource transfers can confer many adaptive benefits such as specialization, helping genetically related individuals, future compensation, and risk-pooling. Need-based transfers are a risk-pooling mechanism in which partners mitigate unpredictable losses by transferring resources based on need. Need-based transfers are likely to be most useful for resources that are necessary and unpredictable because being unable to reliably obtain essential resources would be devastating. However, need-based transfers make people vulnerable to two types of exploitation: a person can be greedy by asking when not in need and a person with a surplus of resources can be stingy by not giving to someone in need. Previous research suggests that people might have cognitive mechanisms for detecting greediness and stinginess, which would serve to protect against exploitation by cheaters. This study investigated whether resources that are necessary and unpredictable are most likely to trigger greediness and stinginess detection mechanisms. Participants saw four types of rules. One rule could be violated through greedy behavior, another through stingy behavior, another by not paying a debt, and another was a descriptive rule that could be violated by not finding one type of resource near another type of resource. Then, participants saw information about events relating to one of the rules and indicated whether the rule in question could have been violated. Consistent with past research, participants were better at detecting greediness, stinginess, and debts not paid than at detecting violations of a descriptive rule. However, contrary to my predictions, the necessity and unpredictability of resources did not impact people's ability to detect greediness and stinginess. The lack of support for my hypothesis might be because the benefits of detecting greediness and stinginess might outweigh the costs even

for situations in which need-based transfer rules are unlikely to apply, because people might be able to consciously activate their greediness and stinginess mechanisms even for resources that would not naturally trigger them, or because of methodological limitations.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Imagine that you are on a camping trip in a dry region. During your time camping, a stranger walks up to you and says that they are almost out of water. Then, they ask you to pour some of your water into their canteen. Now imagine a different scenario. In this new scenario, you are walking down the main street of a metropolitan city while eating ice cream. This time, a stranger walks up to you and says that they are almost out of ice cream. Then, they ask you to scoop some of your ice cream into their ice cream bowl. A simple monetary maximization model would predict that if the value of the ice cream and of the water requested is equivalent, then your response to both requests should be the same. In reality, you might be willing to give some of your water in the first scenario, while you might consider the person requesting ice cream in the second scenario to be greedy and deny their request. Now imagine that the tables are turned and you are the one asking for water and ice cream from a stranger. You might think the stranger is stingy if they deny your request for water, while you might not take offense if the stranger denies your request for ice cream. Why do different resources lead to different perceptions of greediness and stinginess?

It is useful to consider an evolutionary perspective to investigate why decisions about whether someone is stingy or greedy depend on the type of resource involved. Evolutionary theory suggests that engaging in resource transfers could make people vulnerable to cheating. Throughout this paper, I use the word cheating to refer to a violation of resource transfer rules in the pursuit of personal fitness gain and at the expense of the transfer partner. Given this vulnerability to cheating, there should have

been selection pressure for cheater detection mechanisms that would allow people to identify instances of cheating. Consistent with this prediction, research has found that people have cognitive mechanisms for detecting debts not paid, greediness, and stinginess (Cosmides & Tooby, 1992; Munoz et al., in prep). Building on this line of research, I propose that greediness and stinginess detection mechanisms are not part of a general-purpose reasoning system, but rather specialized adaptive mechanisms. Specifically, I propose that greediness and stinginess detection mechanisms should be attuned to at least two characteristics of the resources being transferred: resource necessity and unpredictability. Resources that are high on necessity and unpredictability, such as water in the camping scenario described above, should be particularly likely to trigger greediness and stinginess detection. On the other hand, resources that are low on necessity and unpredictability, such as ice cream in the second scenario, should be particularly unlikely to trigger greediness and stinginess detection.

People would not need greediness and stinginess detection mechanisms if there were no resource transfers. It might seem counterintuitive that natural selection would favor resource transfer mechanisms that make organisms vulnerable to cheating. However, resource transfers, including dyadic resource transfers, can also confer immense fitness benefits.

There are Multiple Benefits to Engaging in Resource Transfers

One evolutionary benefit from engaging in resource transfers is the ability to save time and energy through specialization. Transferring resources can allow individuals to specialize in fulfilling one need while relying on other individuals to help them fulfill their other needs (Nowak, 2006). This can be especially useful if different necessary

resources require different skill sets that can be improved upon over time and if obtaining these resources is time consuming and effortful. In such a situation, attempting to gather all necessary resources oneself might not be as efficient as investing time to perfect the skills needed to gather a subset of resources and then trading some of those resources with people who have specialized in obtaining other kinds of resources.

From an evolutionary perspective, another advantage of dyadic resource transfers occurs among genetic relatives. Individuals who manage to increase the fitness of other individuals who are genetically related to them, i.e., genetic kin, are indirectly boosting their own fitness because genetic relatives share many of the same genes. Hamilton's rule is one theory that describes the kinds of circumstances under which we would expect to see altruistic behavior towards kin — including some types of dyadic resource transfers. Hamilton's rule suggests that inclusive fitness is adaptive if the benefits multiplied by the relatedness between the giver and the recipient are greater than the costs incurred by the giver (Hamilton, 1964). In humans, some evidence has been found that people tend to help close relatives even if this means incurring a cost to themselves and that they help close relatives more than non-relatives and distant relatives (Essock-Vitale & McGuire, 1985 and Kruger, 2003).

Kin selection explains the benefit of some types of resource transfers between genetically related dyads, but it does not explain how resource transfers between unrelated dyads could have evolved. Humans provide many different types of help to non-kin, including help in the form of resource transfers (Hruschka, 2010). One study found that people tended to link kinship cognitions to people with similar attitudes as their own and that people were more inclined to help people who had been linked to

kinship cognitions (Park & Schaller, 2004). One explanation for these findings is that altruism between non-kin represents a misfiring of cognitive mechanisms that evolved for recognizing kin. However, other research suggests that people across different societies apply kinship terminology on the basis of fitness interdependence, or the extent to which organisms influence each other's gene replication success (Cronk et al., 2019). This analysis of how people apply kinship terminology suggests that random misfiring might not be responsible for the linking of kinship cognitions to non-genetic-relatives. Instead, at least in the case of kinship terminology, it appears that people might co-opt kinship cognitions and link them to non-kin for adaptive purposes.

Evolutionary theorists have long proposed that helping non-kin might be an adaptation on its own rather than a misfiring of kin recognition. One of these researchers, Robert Trivers, proposed the theory of reciprocal altruism to explain how non-kin altruism can be adaptive (1971). Trivers proposed his theory of reciprocal altruism as an explanation for altruism in general, but it can be applied to dyadic resource transfers. Reciprocal altruism proposes that the evolution of resource transfers without immediate reciprocation can evolve when the short-term cost to the giver is less than the benefit to the recipient and less than the long-term benefit to the giver.

Reciprocal altruism goes beyond kin selection to shed light on the viability of altruism without the need for close genetic relatedness. This theory also revolutionized the study of altruism by ceasing to consider altruistic acts as occurring in a one-shot vacuum and instead recognizing that fitness costs and gains accumulate over repeated encounters. Therefore, reciprocal altruism lay the foundation for other researchers to

propose cognitive mechanisms through which resource transfers could take place without the need for immediate repayment.

Another theory, known as the need-based transfer framework, proposes that unilateral resource transfers between dyads can evolve through providing risk management benefits (Aktipis et al., 2016; Hao et al, 2015). In need-based transfer systems, a person who is in need can request resources from their partner (Aktipis, Cronk, & De Aguilar, 2011; Cronk et al., 2019). Their partner will then grant them the requested resources if enough resources are available to give, such that the giver would remain above the need threshold. Participation in need-based transfer relationships can enhance fitness by increasing the likelihood that people will have enough resources in times of need. Thus, the theory behind need-based transfers suggests that they should enable unilateral resource transfers between dyads to be evolutionarily stable, as long as the risk management benefits for the giver are greater than the short-term cost of the transfer.

To test the theory behind need-based transfers, researchers used agent-based models with parameters based on the Maasai system of transfers (Aktipis et al., 2016; Aktipis et al., 2011; Hao et al, 2015). Maasai pastoralists in East Africa have relationships known as “osotua,” which translates as “umbilical cord.” Osotua partners help each other in times of need without an expectation of repayment. These relationships allow the Maasai to survive when they experience unexpected losses due to events such as droughts. Extensive fieldwork has shown that need-based transfer relationships, in which risk is pooled to guard against unexpected catastrophes, are common across many societies in addition to the Maasai (Cronk et al., 2019).

Need-based transfer rules in the models are based on osotua relationships, such that agents are only allowed to ask when in need and when asked, agents must give as long as they have enough resources to do so without falling below a threshold for need (Aktipis et al., 2011). The first set of models compared need-based transfers to systems where there are no transfers and to probabilistic rules for giving and requesting resources. In the models, all agents were assigned 70 cattle and each round, there was a probability that some of those cattle would be eliminated. These parameters were based on the ecological reality of Maasai pastoralism as reported by Dahl and Hjort (1976). The results of the model showed that need-based transfers resulted in longer herd survival than any of the other rules tested.

Additional agent-based modeling studies have looked at the network features that result in the greatest herd survival (Hao et al., 2014). Researchers found that larger networks and a larger proportion of network connections switched at random — as long as the proportion is 30% or less — result in greater herd survival if the wealthiest partner in the network is asked for resources. Moreover, the models showed that networks with connections between more agents led to greater herd survival, regardless of whether asking is random or selective towards the wealthiest partner.

The need-based transfer literature has also contrasted need-based transfer rules with another set of potential resource transfer rules: debt-based transfers (Aktipis et al., 2016). In a debt-based transfer, the decision to transfer a resource depends on information about the recipient's likelihood to pay back (what has been often called reciprocity or 'tit-for-tat reciprocity'). Once a resource is given, the giver in a debt-based transfer keeps track of the amount given and the time it takes the recipient to pay for the benefit

received. The giver then uses this information to inform decisions about future transfers with that partner. In contrast, a need-based transfer does not involve a desire for repayment other than the expectation that if one of the partners is ever in need, the other partner will give if able. Debt-based transfers are based on the Maasai term “sile”, which indicates a transaction where debt is accrued and there is an expectation of repayment after a resource transfer.

Modeling work that compares need-based transfers with debt-based transfers used a five-step model schedule. In the first step, herds increase in size. Then, a disaster has a probability of striking, which would decrease the herd size. Then, agents are allowed to request and transfer resources using either need-based transfer or debt-based transfer rules. Finally, the round ends by checking whether agents survived. Agents who fail to stay above a minimum threshold for two consecutive rounds are considered to be no longer viable and are removed from the model.

The results of this modeling work show that need-based transfers led to greater herd survival than debt-based transfers (Aktipis et al., 2016). Need-based transfers also led to greater wealth equality. The results of the model suggest that need-based transfers are particularly useful in volatile ecologies where there are unpredictable events that result in losses. When the likelihood and severity of losses are very low (very low volatility), all individuals survive. When the likelihood and severity of losses is very high (very high volatility), all agents perish quickly. At any point between very low volatility (where all herds survive) and very high volatility (where all herds die), need-based transfers perform better than debt-based transfers. The researchers also found that need-based transfers are crucially dependent on high levels of generosity. Rules that are less

generous, such that agents with enough resources give to their partners in need only sometimes, do not outperform debt-based transfers and models with no transfers.

The need-based transfer rules that have been explored through this modeling work are consistent with Trivers' theory in that they both argue that sometimes it can be adaptive for organisms to incur costs by transferring resources to another organism. However, Trivers' theory of reciprocal altruism and the need-based transfer framework have different theoretical underpinnings. Reciprocal altruism focuses on the benefits that might be gained in the long-term by incurring a short-term cost. In contrast, the need-based transfer framework was proposed as a mechanism for risk management, so its focus is on the catastrophic losses that can be mitigated by forming long-term partnerships of mutual insurance.

In addition to their theoretical differences, these two frameworks lead to different predictions about the specific cognitive architecture that would be involved in resource transfers. A main difference between reciprocal altruism and need-based transfers is the consideration of symmetry of exchanges. Trivers suggests that reciprocal altruism should be most likely to occur when there are repeated interactions, when there are several situations in which the cost to one partner will be smaller than the benefit to another, and when partners are exposed to roughly symmetrical exchanges, such that both give at roughly equal cost to themselves and receive roughly equal benefits over the long term. The need-based transfer framework doesn't assume that symmetry of exchanges is necessary for the emergence of need-based transfers. Instead, resource transfers operate based on need as a mechanism for mitigating risk. As such, need-based transfers can be a stable strategy even when there is asymmetry of exchanges, if the risk of not maintaining

the partnership to mitigate the risk of catastrophic unpredictable losses would be too great. This notion is supported by agent-based modeling work that suggests that need-based transfers, which did not include any parameters requiring symmetry of exchanges in the models, can outcompete systems with no transfers (Aktipis et al., 2011; Aktipis et al., 2016), probabilistic transfer rules (Aktipis et al., 2011), and debt-based transfers (Aktipis et al., 2016).

Several Frameworks Suggest that Different Transfer Rules Apply in Different Situations

Similar to the literature that compares need-based transfers to debt-based transfers, other researchers have recognized that different transfer rules apply depending on the situation and that sometimes, there is no expectation for repayment. One framework argues that different types of relationships are characterized by different rules for how to transfer resources (Clark & Mills, 1979). In their framework, Clark and Mills make a distinction between communal and exchange relationships. Communal relationships are characterized by transfers that do not require repayment and are typically given with the aim of fulfilling the partner's needs and desires. Moreover, communal partners actively avoid keeping accounts of what has been given and what has been received. On the other hand, exchange relationships involve keeping accounts and expecting debts to be repaid.

Most of the research involving this framework has sought to establish that people in communal relationships avoid keeping accounts (Clark & Mills, 2012). For example, one study involved participants meeting a confederate and being placed in one of two conditions (Clark, 1984). In one condition, male participants were given a description of

the confederate that suggested that she was looking for friends and possibly a romantic partner. The researchers argued that this condition led participants to desire a communal relationship with the confederate. In another condition, participants read a description of the confederate that suggested that she would not be interested in becoming neither friends nor romantically involved with them. The researchers argued that this condition discouraged participants from desiring a communal relationship with the confederate.

Then, participants were asked to work on a task that involved working with a pen. They were instructed to alternate work on the task between themselves and the confederate. They were also told that they would be splitting the reward at the end of the session with the confederate and that they could split the amount however they liked. The participants had the option of working with two different colored pens.

The results of this experiment showed that participants that had a desire to form a communal relationship with the confederate chose to work with the same color of pen as the confederate more often than they chose to work with a different color of pen. They found the opposite pattern for the condition where participants did not have a desire to form a communal relationship with the confederate. The researchers argued that these results were evidence that people who desire communal relationships avoid keeping accounts since choosing the same color of pen as the confederate made it impossible to keep track of how much work each person had contributed.

Clark and Mills do not distinguish between different types of close relationships and include friend, mate, and kin relationships as part of a single category governed by the same transfer rules of communal relationships. Nonetheless, Clark and Mills' findings about people sometimes avoiding keeping accounts is consistent with research on

friendship. Studies on friendship suggest that friends are much less concerned with maintaining both short-term and long-term balance of transfers than are strangers and acquaintances (Hruschka, 2010). One cross-cultural study that included participants from the United States, Japan, and China found that not only are friends more tolerant of imbalances in payoffs than strangers, but people also monitor their friends less than they do strangers (Xue & Silk, 2012). Moreover, other studies have found that people tend to give more to friends than to casual acquaintances, that an evaluation of need is an essential component of deciding whether to give to a friend, and that people are more likely to give to friends to meet their needs than to choose an equal distribution of surplus (Schwinger & Lamm, 1981). These findings have led some researchers to note that while primatology researchers often use evidence of repayment of goods and services as an indication of friendship, human friendships seem to be characterized by a lack of both an expectation for repayment and account-keeping (Silk, 2003; Silk, 2002).

Some researchers have argued that an expectation of repayment is still present in friendships, but that the timescale is elongated, such that people don't expect immediate repayment from friends (Shackelford & Buss, 1996). They argue that the shadow of the future in exchange relationships is short because we don't know whether we will interact repeatedly with an exchange relationship partner over the long-run. In contrast, they argue that friendships have an extended shadow of the future, which leads to a longer-term tolerance for repayment of debts. However, this theory is incompatible with Clark and Mills' findings about how friends and people seeking a friendship obscure their contributions to joint tasks to make it harder for account-keeping to occur (Clark & Mills, 2012; Silk, 2003).

Another theory for understanding friendships from an evolutionary perspective is known as the Banker's Paradox (Tooby & Cosmides, 1996). The Banker's Paradox is that when someone needs help the most, they often have a bad credit rating and are less likely to receive a loan. Cosmides and Tooby view choosing friends much like a banker would choose who to loan money to. They argue that given that people have limited resources and time, the key to receiving help when you need it is to choose the right friends. Cosmides and Tooby explore the cognitive calculus that people might employ when choosing their friends. For instance, they predict that people are likely to assess potential friends based on whether their potential friends consider them to be irreplaceable. If your friends consider you to be irreplaceable, it is more likely that they will be willing to provide help when you need it.

Although the purpose of need-based transfers is not to provide a framework for understanding friendship, need-based transfers are likely to occur among friends and they might help us to understand one of the adaptive benefits of friendship. Recent theoretical work has suggested that it is beneficial to consider friendship through the lens of evolutionary theory and to specify the algorithms that might be involved in this type of relationship (Hruschka, Hackman, & Macfarlan, 2015; Hruschka & Silk, 2015). Need-based transfers propose an evolutionary benefit to engaging in these types of transfers: risk-pooling. Although need-based transfers can also occur among kin, expanding the size of the risk pooling network beyond kinship ties can provide a wider safety net that could further buffer against unpredictable events (Aktipis et al., 2018; Cronk et al., 2019; Hao et al, 2015).

Moreover, need-based transfers specify the algorithms that might be employed in this type of transfer: only ask when in need and give if able. These algorithms set need-based transfers apart from the evolutionary frameworks for friendship discussed thus far. Clark and Mills performed rigorous testing to establish that people in communal relationships, unlike people in exchange relationships, avoid keeping accounts. However, they didn't specify what the rules for transferring resources in communal relationships might be or whether these rules could be violated in specific ways.

The shadow of the future framework proposed by Shackelford also differs from the need-based transfer framework. The shadow of the future framework for friendship predicts that people should always keep accounts of debts. In contrast, the need-based transfer framework suggests that some transfers are made without an expectation of repayment other than the implicit agreement that if a need ever arises, help will be provided. Research on friendship, consistent with research on communal relationships, suggests that friends are less concerned with keeping accounts and balancing transfers than strangers (Hruschka, 2010; Clark and Mills, 2012). However, although less often than among strangers, some tracking of payoffs does seem to occur among friends (Xue & Silk, 2012). This finding is also consistent with the need-based transfer framework. The need-based transfer framework allows for different types of transfers to occur within the same relationship depending on the situation. For example, research finds that ranchers in the Malpai borderlands of Arizona and New Mexico invoke need-based transfer rules and rely on their neighbors for help when they sustain an injury that prevents them from working (Cronk et al., 2019). However, ranchers also often engage in

debt-based transfers with the same neighbors by engaging in transactions such as the purchase of equipment from each other or payments for specialized labor.

The Banker's Paradox proposed by Tooby and Cosmides leads to some of the same predictions as the need-based transfer framework, but the algorithms that are specified by each of these frameworks address different questions. Both need-based transfers and the Banker's Paradox recognize the value of having relationships with people that will help you when you're in need. However, the Banker's Paradox framework sees these relationships as a banker providing loans, which implies that account-keeping and debt collection would likely be components of these transfer relationships. In contrast, debt is not generated or tracked after a need-based transfer takes place. Moreover, the Banker's Paradox focuses on identifying the cognitive architecture underlying the process of choosing a friend, whereas the need-based transfer framework specifies rules of engagement during resource transfers.

Another influential framework that focuses on classifying different types of relationships was proposed by Alan Fiske (1992). Fiske's framework proposes that there are four main types of relationships. Fiske argues that many aspects of human social behavior, including planning, production, and coordination, can be evaluated using these four kinds of relationships. Therefore, resource transfers are one of many social behaviors that Fiske thought could be evaluated by recognizing four different types of relationships. The first type of relationship, called communal sharing, is characterized by common property, where people take what they need from a common pool of resources. The second kind of relationship is called authority ranking. In this kind of relationship, there is an imbalance of power. The more powerful individual can take what they want and the

less powerful individual can do little about it. The third type of relationship, equality matching, involves transfers that are aimed at maintaining egalitarian contributions and benefits received. People often keep track of transfers in this kind of relationship. The last kind of relationship is called market pricing. In this kind of relationship, people keep strict accounts and employ ratios to make sure a fair payment is made for resources received.

The need-based transfer framework differs from the Fiske framework, just as it does from the communal relationship framework, in that it does not claim that a particular dyadic relationship will consistently use the same rules for exchanging resources. The relationship literature has focused on classifying relationships by transfer format. In doing this, relationship frameworks don't recognize that the same relationship could employ different transfer rules at different times and in different contexts. Clark and Mills have briefly and partially addressed such challenges by stating that there might be other categories of hybrid relationships that are a mix between exchange and communal relationships. However, Clark and Mills have not elaborated on what rules would be followed in such hybrid relationships. Furthermore, even the addition of more categories, such as in Fiske's four-category framework, does not address the assumption that each relationship involves the use of transfer rules that are characteristic of that particular relationship. The need-based transfer framework does not claim that particular relationships only use one type of rule. Rather, the need-based transfer framework suggests that there are multiple transfer rules that can be employed by many different types of dyads depending on the situation.

Research on resource transfers is not limited to work with humans. Researchers in primatology have also recognized the need to identify the different rules underlying resource transfers. One framework proposes three different cognitive mechanisms for reciprocal altruism: symmetry-based, attitudinal, and calculated reciprocity (Brosnan & de Waal, 2002). Symmetry-based reciprocity involves the formation of a relationship between two individuals in which there is no need to keep track of transfers over time. Rather, give and take occurs in both directions, as long as the relationship remains satisfactory.

Attitudinal reciprocity involves exhibiting an altruistic predisposition and then mirroring the behavior of the other partner during subsequent transfers. Researchers have sought support for this mechanism through tray-pulling experiments with capuchin monkeys (de Waal & Berger, 2000). In one variation of these experiments, both monkeys in a pair have to pull on a rope attached to a tray in order for one of the monkeys to obtain a food reward. The researchers argue that if the monkeys didn't have cognitive mechanisms for attitudinal reciprocity, no reward would be obtained since one of the monkeys in the pair would not be motivated to pull on the rope. However, these studies found that often both monkeys would pull on the rope and that the monkey that received the reward would usually share the reward with the other monkey. This occurred even in cases when there was no prior acquaintance between the monkeys.

The third type of mechanism described by Brosnan and de Waal is known as calculated reciprocity. This type of reciprocity involves keeping track of previous exchanges with a specific individual in the long run to determine whether or not it would pay off to behave altruistically with that particular individual (Brosnan & de Waal, 2002).

This type of reciprocal altruism requires a high degree of cognitive complexity. It involves considering previous exchanges over long periods of time and using that knowledge to inform each decision about whether it would pay to behave altruistically or not towards a specific individual. This type of mechanism usually includes an expectation to receive a form of payment in return for favors granted in the past. The requirement for high cognitive capacity limits the number of species that could possess cognitive mechanisms for calculated reciprocity. In fact, it has mostly been observed in chimpanzees and in humans (Brosnan & de Waal, 2002). One study found that chimpanzees commonly rely on calculated reciprocity (de Waal, 1997). This study reported that chimpanzees that had groomed other chimpanzees in the morning received more food in the afternoon from the chimpanzees that they had groomed than from the chimpanzees that they had not groomed.

Brosnan and deWaal argue that symmetry-based reciprocity is the most common of these three mechanisms and is likely to be the most ancient. They also argue that calculated reciprocity is the rarest, the most complex, and probably the most recent of the three. However, this framework does not make specific predictions about whether each of these mechanisms serves a different evolutionary function.

Sahlins' typology in the field of anthropology is another framework that defines different types of dyadic transfers (1972). Sahlins argues that there are three different types of dyadic transfers: generalized reciprocity, balanced reciprocity, and negative reciprocity. Generalized reciprocity involves a transfer in which the giver does not expect immediate repayment. Rather, the time allowed for repayment is indefinite and the payment does not need to be equal in quality or quantity to the resource that is received.

Generalized reciprocity roughly corresponds to Brosnan et al.'s symmetry-based reciprocity in that both mechanisms allow for an agent to give repetitively without receiving a payment, although generalized reciprocity does not make an emphasis on the role of relationship satisfaction for the continuation of these types of transfers.

Unlike generalized reciprocity, balanced reciprocity involves resource transfers that require immediate repayment. Thus, balanced reciprocity is similar to attitudinal reciprocity in that failure to receive payment would result in termination of the transfers. However, balanced reciprocity is different in that it does not require an initial altruistic disposition. Finally, negative reciprocity involves one person taking resources from another without consent and without giving anything in return. Negative reciprocity does not have an equivalent in Brosnan et al.'s typology since Brosnan et al. focus on mechanisms that involve incurring a cost to increase the fitness of someone else. Sahlins' typology does not explain how these different types of transfers might have been shaped by evolution.

The Need-Based Transfer Framework Makes Specific Predictions About the Cognitive Architecture for Cheater Detection

Debt-based transfers are similar to calculated reciprocity, balanced reciprocity, and exchange relationships. Need-based transfers are similar to symmetry-based reciprocity, generalized reciprocity, and communal relationships. However, one major distinction between the need-based transfer framework and the rest of the frameworks discussed thus far is that the need-based transfer framework makes specific predictions about what constitutes a violation of the rules, because the rules are algorithmically specified. The Brosnan et al. framework argues that transfers that follow the rules of

symmetry-based reciprocity would continue if the relationship remains satisfactory. However, it is not clear what would cause the relationship to cease to be satisfactory. Therefore, it is not clear what would constitute a violation of symmetry-based reciprocity.

Similarly, Sahlins does not state what constitutes a violation of generalized reciprocity. Sahlins states that failure to pay back does not lead the giver to stop giving under generalized reciprocity transfers. Therefore, he makes clear that failure to pay back is not a violation, but does not state what is a violation. The Clark and Mills framework mainly focuses on how people avoid keeping accounts in communal relationships. Therefore, keeping accounts might be considered a violation in communal relationships. This framework also suggests that people in communal relationships are responsive towards their partner's needs and desires but does not discuss explicitly what would be considered a violation of communal rules.

In contrast to previous frameworks of resource transfer mechanisms, the need-based transfer framework explicitly states what counts as a violation of the rules. There are two possible violations in need-based transfers. First, a person with a surplus of resources could be stingy and refuse to give resources to someone in need. Second, a person could be greedy and ask for resources when not in need. Therefore, this framework suggests that even when people engage in transfers where there is no keeping of accounts, people must still be vigilant to the possibility of these types of cheating.

From an evolutionary standpoint, the need-based transfer framework's specificity about what constitutes a violation is a major strength because it recognizes that resource transfer systems can be vulnerable to cheating without adequate safeguards and it allows for predictions about potential cheater detection mechanisms. Even though cooperative

resource transfers can provide several adaptive benefits, cooperation is not always the preferred strategy. Following the rules of resource transfers can result in a large payoff in the long-run, but choosing the selfish option often provides a larger immediate reward (Trivers, 1971). Simulation games of a situation known as "cheater invasion" show that if cheaters cannot be identified and punished, a population of cooperators that rely on random interactions will be overtaken by the dominant cheating strategy (Axelrod, 1984).

Being able to detect cheaters is advantageous from an evolutionary standpoint because it allows people to retaliate against cheaters, thus reducing the likelihood that others would want to cheat them. Cheater invasion simulation games that show cheating to be the winning strategy over cooperation often rely on random interactions (Axelrod, 1984). However, if cheating can be identified in these simulation games and cooperators are able to either punish cheaters or selectively interact with cooperators, cooperation becomes a more successful strategy than cheating (Aktipis, 2011; Barclay & Willer, 2007). Following this logic, studies have investigated retaliatory behavior and found that people are likely to punish cheating even if it means incurring a cost themselves to do so (Fehr, Fischbacher, & Gächter, 2002). Additionally, models have shown that walking away from relationships can also serve to protect cooperators, as well as to indirectly inflict costs on cheaters (Aktipis, 2011).

Trivers recognized that cheating can be a threat to the viability of reciprocal altruism (Trivers, 1971). For Trivers, cheating meant a failure to act altruistically towards another person in a situation in which the immediate cost of behaving altruistically for the actor is lower than the immediate benefit to the recipient. If cheaters receive the benefits of altruistic acts without having to act altruistically themselves,

people possessing cheating alleles would be more successful than people possessing altruistic alleles. Therefore, he included as part of his model the need for altruists to cease to behave altruistically towards people who cheat (Trivers, 1971). This form of punishment towards cheaters would require that people possess cognitive mechanisms for identifying cheating. Although Trivers doesn't explicitly discuss this, reciprocal altruism suggests that a potential recipient of an altruistic act must engage in a cognitive process to assess the size of the immediate benefit to themselves and compare it to the immediate cost to the giver. If the potential recipient determines that the benefit to themselves is greater than the cost to the giver, and the giver refuses to help, the potential recipient should be able to identify this situation as an instance of cheating.

Cheating is different in the need-based transfer framework than in reciprocal altruism (Aktipis et al., 2016). There are two possible violations in need-based transfers. First, a person with a surplus of resources could be stingy and refuse to give resources to someone in need. Second, a person could be greedy and ask for resources when not in need. Thus, a potential recipient in a need-based transfer situation would need to determine whether the potential giver has enough resources to give without resulting in the giver being below the need threshold. Trivers focuses his discussion of cheating on the potential for the giver to cheat and does not discuss whether a recipient could also cheat. In contrast, the need-based transfer framework suggests that the giver would also require cheater-detection mechanisms to assess whether the recipient is in need before agreeing to give resources.

Given the importance of cheater detection for the viability of resource transfer mechanisms, it is likely that natural selection would have selected for cognitive

mechanisms to detect cheating. To test this hypothesis, Leda Cosmides and John Tooby conducted several studies involving the Wason selection task (Cosmides, 1989; Cosmides & Tooby, 1992; Cosmides, Barrett, & Tooby, 2010). As its name suggests, the Wason selection task was created by Peter Wason (1968) to test people's deductive reasoning.

In a Wason selection task, participants are presented with a logical proposition of the form "If P, then Q", where Q is the consequent that must be true if the antecedent is P (see *Figure 1* for an example of the Wason selection task). After seeing the rule, participants are presented with four cards. Participants are told that each of the cards has information about the antecedent on one side and information about the consequent on the opposite side. However, participants can only see the information on one side of the card. One of the cards shows a statement that the antecedent is true (P), another shows a statement that the antecedent is false (Not-P), another shows a statement that the consequent is true (Q), and another card shows a statement that the consequent is false (Not-Q). Each of the cards represents an independent instance in which the rule might or might not have been violated. The participants' task is to decide which of the cards must be turned over to determine whether the rule has been violated or not in each of the four instances.

The correct answer for this task is to select the card showing that the antecedent is true (P) and the card showing that the consequent is not true (Not-Q). For example, suppose that the rule in question is: "If I give Jim my sunglasses, then Jim gives me $40" (see *Figure 1*). In this case, the cards would read: "I gave Jim my sunglasses" (P), "I didn't give Jim my sunglasses" (Not-P), "Jim gave me $40" (Q), and "Jim didn't give me $40" (Not-Q). The cards that tell you whether I gave Jim my sunglasses have information

about whether Jim gave me $40 or not on the opposite side. The cards that tell you whether Jim gave me $40 or not have information about whether I gave Jim my sunglasses on the other side. The question is which of these four cards we must turn over to determine whether Jim violated the rule.

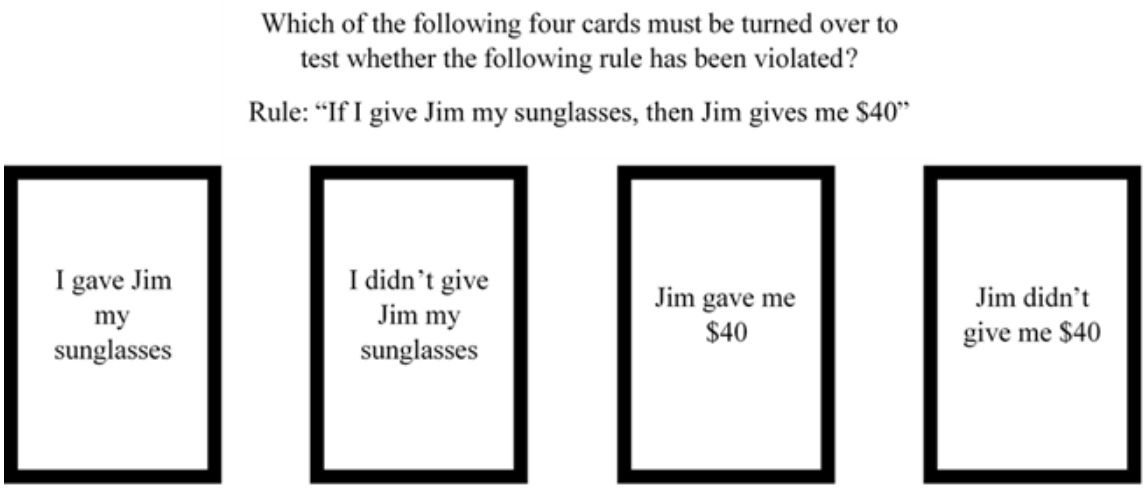

Figure 1. Wason selection task example. This figure shows a Wason selection task for which the correct answer would be to only turn over the first ("I gave Jim my sunglasses") and last ("Jim didn't give me $40") cards.

First, we would need to turn over the P card that says that I gave Jim my sunglasses to find out whether Jim paid for the sunglasses he received or not. If Jim paid for the sunglasses, then Jim didn't violate the rule. If Jim didn't pay for them, then he violated the rule. Second, we need to turn over the Not-Q card that says that Jim didn't give me $40 to find out whether Jim got the sunglasses or not. If I didn't give my sunglasses to Jim, then Jim didn't violate the rule; he didn't have to pay if he didn't get the sunglasses in the first place. If I gave my sunglasses to Jim, then he violated the rule; he was supposed to pay for the sunglasses after he got them and he didn't.

The other two cards should not be turned over since they are not necessary for determining whether Jim violated the rule. If we know that I didn't give my sunglasses to

Jim (Not-P), then it doesn't matter whether Jim gave me $40 or not; he didn't owe me anything. If we know that Jim gave me $40, then it doesn't matter whether I gave Jim my sunglasses or not; Jim paid the cost, so there is no way that Jim could have violated the rule.

The original set of experiments that used the Wason selection task involved abstract rules involving colors, numbers, shapes, and letters (Wason, 1968). For example, one rule was that: "If there is a D on one side of any card, then there is a 3 on its other side". Very few participants provided the correct answer to these rules, which led Wason to conclude that people are not very good at solving this type of conditional proposition. Following these experiments, the Wason selection task was adapted to test the hypothesis that humans possess specialized mechanisms for cheater detection (Cosmides & Tooby, 1992).

In a series of studies using the Wason selection task, the researchers presented a set of rules to participants that involved the same logic as the original Wason selection task (Cosmides, 1989). Although these rules had the same underlying logic as the original Wason selection task, they were framed as a situation in which participants had to determine whether someone had cheated. Cosmides and Tooby defined cheating in their studies as accepting a benefit without paying a cost. Performance in this task increased dramatically when it was framed in terms of detecting cheating. For example, in one experiment, 75% of participants chose the correct cards for a problem that was framed as detecting cheating, while only 21% of participants chose the correct cards for a problem that was framed as detecting a violation of a descriptive problem where there was no social contract (Cosmides, 1989).

Further studies also served to establish that familiarity with the content was not responsible for the increase in performance and that logical reasoning in general could not account for the results (Cosmides & Tooby, 1992; Cosmides, Barrett, & Tooby, 2010). Rather, the researchers concluded that a specialized module for cheater detection was responsible for allowing people to identify violations of rules where a cost must be paid when a benefit is received. Research has also been conducted to test whether the ability to test cheaters is culture-specific. In one study, researchers adapted the Wason selection task for use with non-literate Shiwiar participants in the Ecuadorian Amazonia (Sugiyama, 2002).

The researchers working with the Shiwiar knew that the traditional Wason selection task would pose problems for non-literate participants given that the task would have to be administered verbally instead of in written form. The traditional Wason selection task presents all four cards at the same time, but the researchers decided to present one card at a time to reduce memory load. In addition, the information about the antecedent and consequent were presented in picture form instead of in written form. Also, instead of having part of the information on one side of the card and the rest of the information on the other side of the card as in a traditional Wason selection task, the researchers had both the consequent and the antecedent on a single side of the card and covered one of them with a paper door that could be opened to reveal the photo underneath. These changes were all made to decrease memory load and make it easier for non-literate participants to understand the task. The researchers found that people who had little contact with modern Western civilization were also able to detect cheating and

performed better on the modified Wason selection task that was framed as a cheater detection situation than on one that was framed as a descriptive problem.

Additional research followed up on the findings by Cosmides and Tooby to investigate whether there is a brain area that is involved in this type of processing. In a study that tested this prediction, a subject referenced as R.M. with bilateral limbic damage to the orbitofrontal cortex, the temporal pole, and the amygdala, was tested on two types of Wason selection tasks (Stone, et al., 2002). Even though both tasks involved the same exact logic, one type of task was phrased as a cheater detection paradigm and the other as a precaution rules paradigm. The results of this experiment indicated that R.M.'s damage did not lead to general impairment of problem solving. However, R.M. was found to perform significantly worse than two control groups on the cheater detection tasks. This second finding suggested that some or all of the regions of R.M.'s brain that were damaged might have been responsible for cheater detection.

The work by Cosmides and Tooby focused on rules that involved paying a cost when a benefit is received. For example, one rule that they tested was: "If you give me your ostrich eggshell, then I'll give you duiker meat." This rule involves the obligation to repay a debt that is incurred after receiving a benefit from someone. If we were to place the logic of this type of rule within the Aktipis et al. (2016) resource transfer framework, we could classify the rules investigated by Cosmides and Tooby as debt-based transfers, in which people keep track of the balance of exchanges and expect repayment after giving someone a resource.

The research conducted by Cosmides and Tooby has contributed significantly to the field by suggesting that specialized cognitive mechanisms for cheater detection might

have developed through natural selection. This work represents the first step in identifying the specific cognitive architecture underlying cheater detection. However, as Cosmides and Tooby themselves recognized, "we have tested only a small number of hypotheses about one mode of social exchange out of many. We have not discussed, much less experimentally explored, the rest of the intricate and complex psychology of social exchange. Anyone examining his or her own human experience will immediately identify large areas of the psychology of social exchange, such as the psychology of friendship, that are not captured well by the models introduced so far" (Cosmides & Tooby, 1992). Therefore, more work is needed to investigate whether there might be multiple specialized cheater detection mechanisms attuned to identify different types of cheating. Recent work has built on the work by Cosmides and Tooby and suggested that people might have mechanisms for detecting violations of other kinds of rules, such as displaying submission to higher status individuals, helping injured allies, and displaying coalitional symbols (Sivan, Curry, & Lissa, 2018).

The recognition that there are different types of cheating led us to conduct a study to test whether people have cognitive mechanisms for detecting violations of need-based transfer rules (Munoz et al., in prep). Need based transfers can be violated in two ways: through greediness and stinginess. Therefore, need-based transfer norms would require two detection mechanisms to make need-based transfers a viable mechanism for reciprocal altruism. The first detection mechanism would allow the party that receives a request for resources (the giver) to detect greediness. The greediness detection rule stipulates that the requester should not ask when not in need. The rule to detect greediness is: "if not in need, then no request should be made". The second detection

mechanism would allow the party that requests resources (requester) to detect stinginess. The stinginess detection rule stipulates that the giver should give in response to a request for resources if a surplus of resources is available to the giver. The rule to prevent stinginess when the giver receives a request for resources is: “if a surplus of resources is available, then give”.

We conducted a within-subjects study in which all participants viewed four different Wason selection tasks (Munoz et al., in prep). In the control task, participants were asked to evaluate a descriptive rule with no resource transfers. In the debt-based transfer task, based on the work by Cosmides and Tooby, participants were asked to detect cheaters who had not paid a debt. The other two tasks involved detecting violations of need-based transfer rules. One task asked participants to detect stinginess and the other asked participants to detect greediness. We found that performance on the debt-based transfer task and on the two need-based transfer tasks was higher than performance on the descriptive task (see *Figure 2*). These results go beyond the findings of previous research by suggesting that people have mechanisms for detecting greediness and stinginess in addition to having mechanisms for detecting debts not paid.

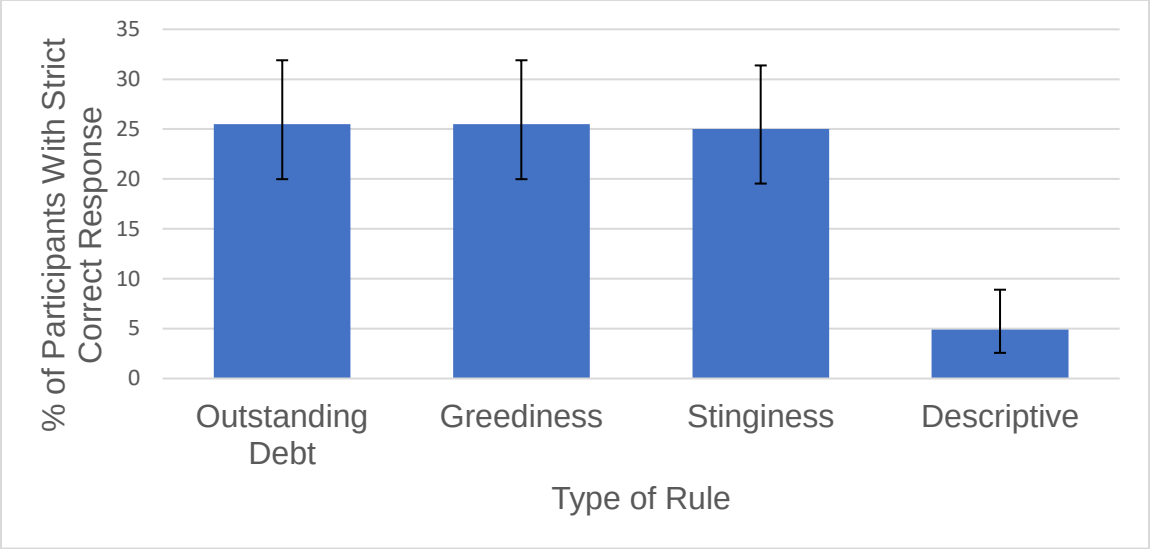

Figure 2. Strict correct responses by type of task (Munoz et al., in prep). More participants gave strict correct answers to questions involving debts not paid, greediness, and stinginess, than to questions involving rules with no resource transfers.

Situational Factors Might Influence the Likelihood of Activation of Greediness and Stinginess Detection Mechanisms

The research by Munoz et al. suggests that people might have greediness and stinginess detection mechanisms, but it does not establish whether these mechanisms are attuned to specific situations (in prep). It could be maladaptive for greediness and stinginess detection mechanisms to become active when need-based transfer rules are not in effect. For example, it would be inefficient to spend resources trying to determine whether your transfer partner is in need if you are exchanging oranges for bananas at the market and you are using debt-based transfer rules to do so. Need is not a relevant factor for debt-based transfers, so it would be wasteful to attend to it in this situation. Therefore, it is useful to think about when the rules of need-based transfers would apply in order to determine when greediness and stinginess detection mechanisms are likely to become active.

To understand when greediness and stinginess detection mechanisms are expected to become active, it is necessary to understand why need-based transfers might have been selected for. Need-based transfers can increase fitness by serving as an insurance policy (Aktipis et al., 2016; Hao et al, 2015; Aktipis, Cronk, & De Aguilar, 2011). Many people pay insurance companies in case an unexpected negative event occurs. For example, someone might make monthly payments to an insurance company in case their house ever catches on fire. The person paying the insurance company hopes that their house never catches on fire and therefore, that they never need the insurance company to pay them back. Similarly, a person who engages in need-based transfers hopes that they will never be in need and therefore, that their partner will never have to pay them back (Cronk et al., 2019). If an unexpected event does cause them to be in need, they can rely on their need-based transfer network to obtain the necessary resources. Therefore, just like insurance companies, need-based transfers can help by alleviating the costs of misfortune (Aktipis et al., 2016; Gurven, 2004). The idea of resource transfer networks serving as a successful form of social insurance against the risk of unpredictable losses or unsuccessful attempts at obtaining a resource has been supported by both modeling and empirical data (Aktipis et al., 2011; Aktipis et al., 2016; Cronk et al., 2019; Hames, 1990; Hao et al., 2014; Hill, Kaplan, Hawkes, Hurtado, 1987; Kaplan et al., 1985; Pietras, Cherek, Lane, Tcheremissine, 2006; Winterhalder, 1986). In fact, research suggests that social insurance can be a more successful form of risk reduction than other strategies such as dynamic foraging strategy modification (Winterhalder, 1986) and that people in different societies around the world successfully rely on social insurance (Cronk et al., 2019; Hames, 1990; Hill et al., 1987).

Although insurance companies are similar in many ways to need-based transfers, they are not a perfect analogy (Cronk et al., 2019). First, insurance companies tend to insure against very specific types of risks. For instance, an insurance policy for flood insurance only covers the risks associated with floods and would not cover the risks associated with car accidents. In contrast, need-based transfer networks include coverage of a wide range of unpredictable needs that do not need to be fully specified in advance (Cronk et al., 2019). Moreover, insurance companies are required to pay in the event of a misfortune regardless of whether the recipient is in need or not. A person who has six houses and has one of them burn down will still get money from their insurance company even though they are not in need. In contrast, need-based transfer partners are not obligated to give resources to a recipient who is not in need. Therefore, as the name suggests, need-based transfers are expected to occur only for resources that are necessary. For example, it is less likely that a need-based transfer would occur when a request for ice cream is made than when a request for water is made.

Research on vampire bats sheds light on another factor that might influence the likelihood of activation of need-based transfers: unpredictability of resources. Vampire bats feed on blood. Crucially, this species of bats cannot survive more than a couple of days without obtaining blood. Vampire bats have been observed to feed blood to hungry kin that were unsuccessful at obtaining blood on a particular night (Carter & Wilkinson, 2015; Wilkinson, 1984). From an evolutionary standpoint, this behavior can be easily explained by inclusive fitness. Kin share genes in common, so it would be in the fitness interest of an individual to give some of their resources to their hungry kin.

However, research suggests that individuals of this species often feed blood to non-kin (Carter & Wilkinson, 2015). Researchers find that bats rely on need and previous association when deciding whether to transfer blood to non-kin (Carter & Wilkinson, 2013; Wilkinson, 1984). Vampire bats incur a short-term cost by giving some of their food to another bat. However, they might enjoy a long-term fitness gain by expanding their risk-pooling network. A vampire bat that obtains a surplus of food might give some of their food to needy bats to increase the likelihood that they will receive help if they are ever in need. Therefore, these vampire bats appear to be relying on the rules of need-based transfers to engage in blood transfers with their risk-pooling partners.

For vampire bats, blood has two characteristics that are important to need-based transfers. First, blood is an essential resource. Without blood, a vampire bat will quickly starve. Second, blood is an unpredictable resource, such that the success rate of vampires at obtaining blood is highly variable. In other words, vampire bats will sometimes be unsuccessful at obtaining blood even if they invest substantial energy into trying to do so. These two factors combined would make it so that even bats with perfect work ethic would sometimes starve to death. To prevent this, bats have developed need-based transfer mechanisms for sharing food with other bats and developing risk-pooling networks.

The same logic that applies to vampire bats can be applied to humans. People are likely to have developed need-based transfer mechanisms that are especially attuned to resources that are both necessary and unpredictable. For example, it might be more likely for the rules of need-based transfers to become active when people are transferring water during a camping trip than when people are transferring ice cream at a restaurant.

Greediness and stinginess detection mechanisms should be most likely to become active when need-based transfers are likely to take place to allow people to detect cheating that is relevant to the rules that are in place in a given situation. Therefore, greediness and stinginess detection should also be most likely to become active when people are transferring resources that are necessary and unpredictable (see *Figure 3* for hypothesis). The experiment described in the methods section aims to test this prediction.

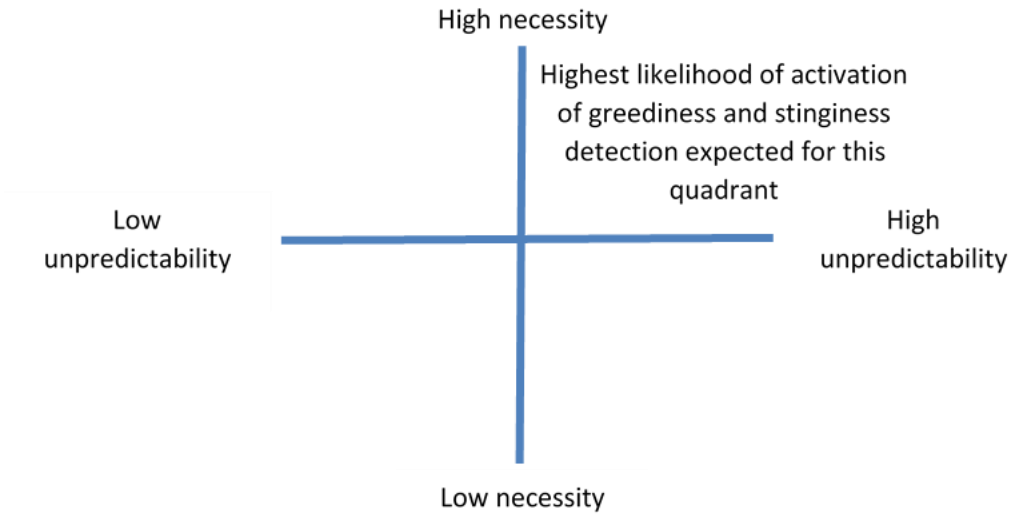

Figure 3. Predictions about activation of greediness and stinginess detection mechanisms based on resource necessity and unpredictability. I expected the highest likelihood of activation of greediness and stinginess detection mechanisms for resources that were high on both necessity and unpredictability.

CHAPTER 2

METHODS

Participants

The task was displayed on a computer to 184 Amazon Mechanical Turk participants who agreed to participate in the study. Participants were recruited to participate through a listing on Amazon Mechanical Turk that asked MTurk workers to participate in a study about rule violations. Out of the participants who agreed to participate, 32 participants did not complete the study and their data were excluded from the analyses, which brought the number of participants who completed the study to 152. Of these participants, 96 were male and 56 were female. The average age of the participants was 36.26 years old. Participants were paid eighty cents for participating in the study.

Materials and Procedure

Participants who were interested in participating in the study after reading the listing on Amazon Mechanical Turk were asked to click on a link that redirected them to the study on Qualtrics. Once participants opened the link on Qualtrics, they read some general information about the study and checked a box to confirm that they were at least eighteen years old and to consent to participating in the study. Then, they saw instructions asking them to imagine that they live on an island where there are coconuts. For this experiment, participants saw one of four versions of the instructions, depending on which condition they were randomly assigned to. In one version of the instructions, participants were given a paragraph suggesting that coconut water was both necessary for survival and another paragraph suggesting that coconut water was a predictable resource.

In another version, they were told that coconut water was necessary for survival and that it was unpredictable. The third version of the instructions told participants that coconut water was not necessary for survival and that it was predictable. Finally, the last version of the instructions told participants that coconut water is not necessary for survival and that it was unpredictable. The order in which the paragraph suggesting the level of necessity of coconut water and the paragraph suggesting the level of unpredictability of coconut water were presented was randomized, such that some participants saw the information about the level of necessity of coconut water first and some participants saw the information about the level of unpredictability of coconut water first.

Version of the instructions that suggests that coconut water is necessary for survival: "Coconut water is the only drinkable liquid available on the island. There are several ponds around the island, but they all contain dirty toxic water that is not safe to drink. In fact, you have a neighbor who died after drinking water from the ponds. Therefore, coconut water is essential for survival on the island."

Version of the instructions that suggests that coconut water is not necessary for survival: "People on the island can choose to drink fresh water instead of drinking coconut water. There are several ponds around the island that contain plenty of fresh water that is safe to drink. In fact, you have a neighbor who doesn't gather coconuts. Instead, he has been drinking water from the ponds exclusively throughout his entire life and it has kept him in perfect health. Coconut water is valued because it tastes better than water from the ponds, but it is not necessary for survival."

Version of the instructions that suggests that coconut water is predictable: "All of the coconuts on the island contain plenty of water inside. This means that people who go out

to gather coconut water always come back with plenty of coconut water. In fact, you have a neighbor who has gathered a surplus of coconut water every day for the past 40 years."

Version of the instructions that suggests that coconut water is unpredictable: "Most of the coconuts on the island have no water inside. This means that people who go out to gather coconut water might get unlucky and get only dry coconuts with no water inside. In fact, you have a neighbor who came home with many coconuts every day, but for an entire week all of his coconuts were dry."

After reading these initial instructions, participants read another set of instructions explaining the task. Then, they were presented with four rules, one at a time (See *Figures 4-7*). The order in which the rules were presented was randomized for each participant. Three of the rules were framed in terms of detecting a violation of a rule about transferring resources. One of these rules was about detecting stinginess, another about detecting greediness, and another about detecting an outstanding debt. The fourth rule was a descriptive rule that didn't involve any type of social contract about resource transfers. Instead, it was about detecting a logical violation of a rule about finding a resource when another resource has been found.

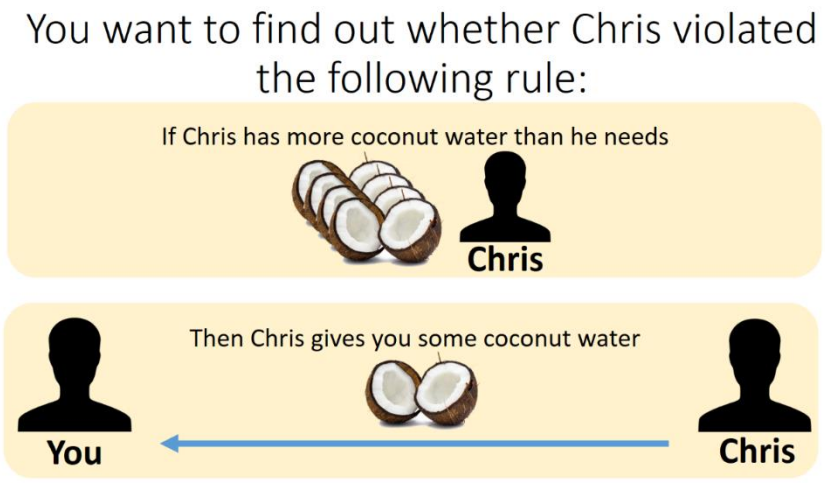

Figure 4. Stinginess detection rule. This figure shows the stinginess detection rule for the cheater detection task developed by Munoz et al. (in prep).

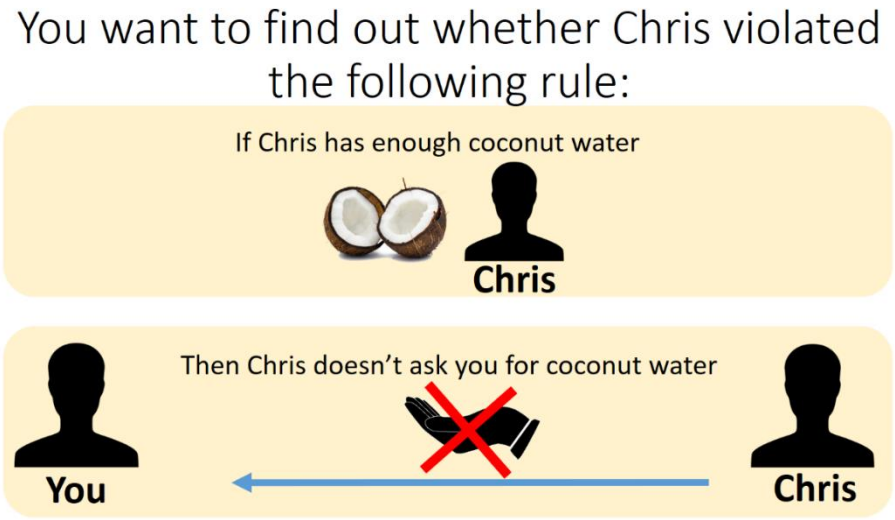

Figure 5. Greediness detection rule. This figure shows the greediness detection rule for the cheater detection task developed by Munoz et al. (in prep).

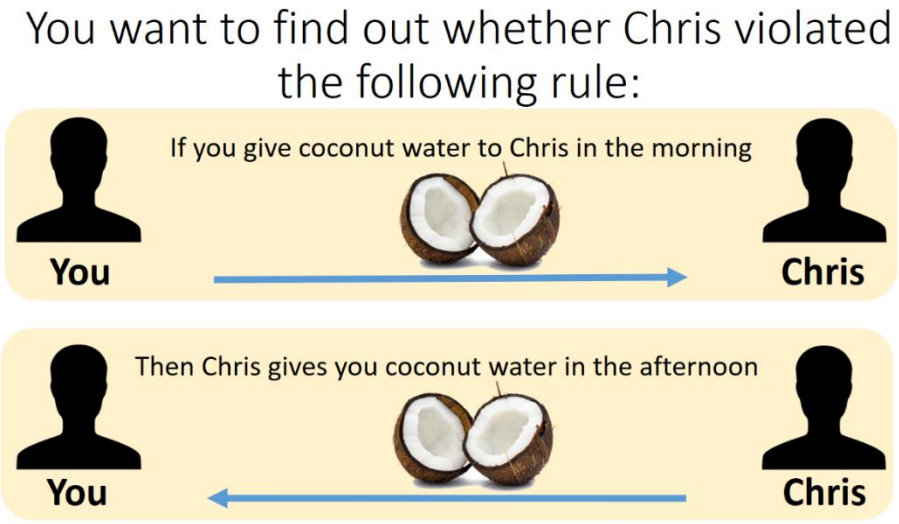

Figure 6. Outstanding debt detection rule. This figure shows the outstanding debt detection rule for the cheater detection task developed by Munoz et al. (in prep).

Figure 7. Descriptive rule. This figure shows the descriptive rule for the cheater detection task developed by Munoz et al. (in prep).

After seeing each rule, participants saw four different questions relating to that rule, one at a time. Each question asked whether the rule could have been violated or not on each of four days. Participants were given one fact for each of the days that provided some information about what happened on that day. The facts given for each day corresponded to one of four possible logical propositions for each rule. That is, one fact stated that the antecedent, or the first part of the logical proposition, was true on that day

(P). Another fact stated that the antecedent was false on that day (Not-P). A third fact stated that the consequent, or the second part of the logical proposition, was true on that day (Q). Finally, a fourth fact stated that the consequent was false on that day (Not-Q). The order in which each fact was presented was randomized for each participant. From a logical standpoint, questions about whether the rule could have been violated that involve P and Not-Q should be answered as "yes, it's possible", while questions involving Not-P and Q should be answered "no, it's not possible".

Figures 8 and 9 are examples of the questions that were presented to participants. *Figure 8* shows the P statement for the stinginess detection rule (true antecedent). The correct answer for this question would be to select "yes, it's possible" since it is possible that Chris could have been stingy by refusing to give even though he had a surplus of coconut water. *Figure 9* shows the Q statement for the greediness detection rule (true consequent). The correct answer for this question would be to select "no, it's not possible" since Chris did not ask for coconut water and therefore, was not greedy.

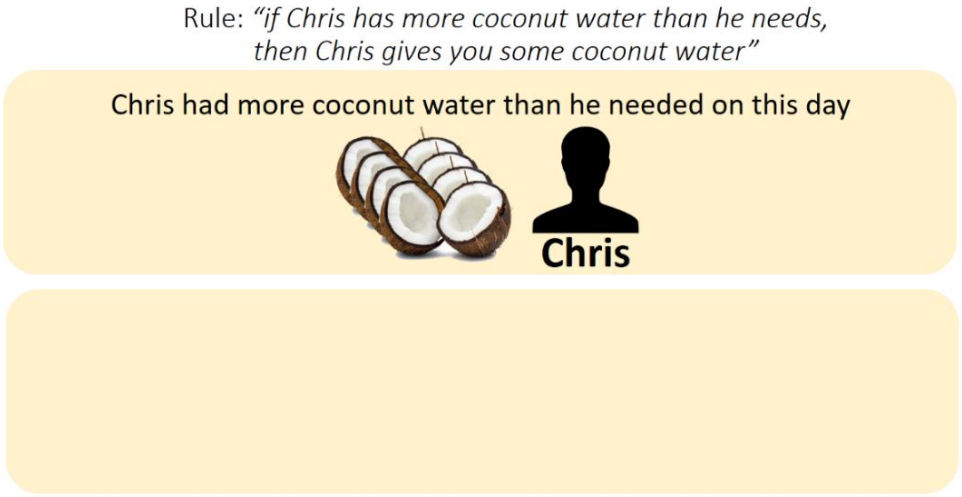

Figure 8. P statement for stinginess detection rule. This figure shows the P statement for the stinginess detection rule for the cheater detection task developed by Munoz et al. (in prep).

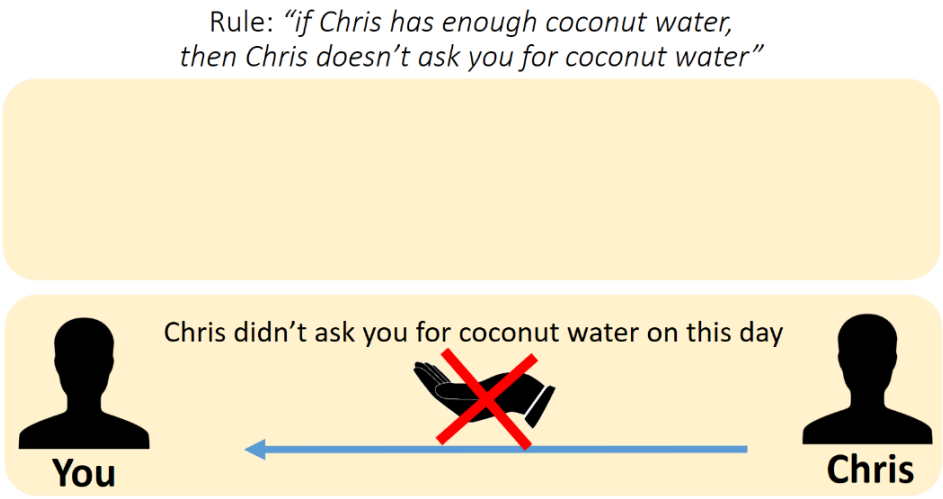

Figure 9. Q statement for the greediness detection rule. This figure shows the Q statement for the greediness detection rule for the cheater detection task developed by Munoz et al. (in prep).

After participants finished answering all four questions for all four types of rules, they were presented with a question that was meant to serve as an attention check. This question told participants that research has shown that participants sometimes skip the instructions and then asked them to choose the last option out of six multiple choice options to show that they were paying attention, regardless of what the question below those instructions was asking. These instructions were followed by a question asking what the study had been about. The last option, which was the correct answer, had nothing to do with the actual study and read: "Issues about geography".

Then, participants saw two questions that were intended as a manipulation check. One of these questions posed a statement about coconut water being necessary for survival on the island and the other posed a statement about the acquisition of coconut water being predictable on the island. The participants were asked whether these statements were true and had to answer by selecting "True" or "False". The correct answer to these questions depended on which type of resource condition the participants

had been randomly assigned to. For example, a participant who was assigned to the high necessity and high unpredictability condition, and who read and understood the instructions, should have answered "true" to the statement about coconut water being necessary and "false" to the statement about coconut water being predictable. After answering the manipulation check questions, participants saw a message thanking them for participating in the study.

This study used the cheater detection task (see *Figures 4-9*) developed by Munoz et al. (in prep). The literature on cheater detection has relied almost entirely on the Wason selection task (Cosmides & Tooby, 1992). The Wason selection task provides strong evidence for the existence of cheater detection mechanisms and it might seem like the obvious choice for conducting research on cheater detection given its monopoly of the cheater detection literature. However, relying on a single task for testing a theoretical framework can make conclusions about that framework vulnerable to biases and idiosyncrasies inherent in the task. The best evidence for a theory can be obtained by using different methodologies and comparing the results to rule out the possibility that any specific feature of the methodology produced the results.

The Wason selection task and the Munoz et al. detection task both present participants with a rule that could be violated, give participants a piece of information that is known, and then ask a question. However, the question asked by each of these tasks is different. The Munoz et al. detection task asks whether the rule could have been violated or not given a piece of information that is known. This is different from the Wason selection task, which asks whether more information is needed to determine whether the rule was violated or not. Therefore, the Munoz et al. detection task more

directly tests if people can determine whether cheating could have occurred given limited information, rather than whether people are good at knowing when to seek or avoid seeking additional information to determine whether cheating could have occurred.

One of the useful features of the Munoz et al. detection task is that, unlike the Wason selection task, it avoids confounding curiosity with incorrect responses. In the Wason selection task, people are presented with four cards, each of which has part of the information on one side of the card and part of the information on the other side of the card. A correct response involves selecting two of the cards and not selecting the other two cards. However, participants might select incorrect cards because they are curious to see what is on the other side of the card. In other words, curiosity and the inability to correctly detect cheating are confounded in the Wason selection task. As opposed to the Wason selection task, the Munoz et al. detection task avoids confounding incorrect responses with curiosity because it doesn't have any hidden information.

Moreover, the new methodology is simpler than the Wason selection task (see *Figures 4-9*). The Wason selection task includes long paragraphs of background information about the rules. These background information paragraphs are often different from one another in many ways between conditions. This could potentially introduce confounds and error. The simpler and more visual nature of the new cheater detection task developed by Munoz et al. also allows for simpler modification to test specific hypotheses, as well as for easier translation for field work (in prep). The task also involves less cognitive load, which would make it easier to implement with non-literate participants. This would allow future work to test whether the same kinds of resources are likely to activate greediness and stinginess detection across different societies.

CHAPTER 3

RESULTS

Variables

There were two independent variables that were considered in this study. The first independent variable of interest was the type of resource. Participants were all asked to imagine that they lived on an island where there is coconut water. However, the information that participants got about coconut water depended on the condition that they were randomly assigned to. Participants were given information about two features of coconut water on the island: necessity, or the degree to which coconut water was essential for survival, and unpredictability, or the degree to which randomness played a role in the acquisition of coconut water. Participants were given information suggesting one of the following combinations of features for coconut water on the island: necessary and unpredictable, necessary and predictable, not necessary and unpredictable, or not necessary and predictable. The second independent variable of interest was the type of rule. Each participant saw four types of rules, each of which presented a different type of potential violation: greediness, stinginess, unpaid debt, and a violation of a descriptive rule.

The dependent variable of interest was people's ability to detect different types of violations. Ability to detect a violation was operationalized in terms of strict correct responses. Each of the four rules that participants saw had four questions associated with it. The questions gave participants information stating that the antecedent was true, that the antecedent was false, that the consequent was true, or that the consequent was false and asked whether a violation of the rule could have occurred. A strict correct response

for each of these rules involved a participant answering all four questions associated with the rule correctly. If a participant answered one or more of these questions incorrectly, it was counted as an incorrect response for that rule.

I expected that resource necessity and unpredictability would increase the likelihood of activation of stinginess and greediness detection mechanisms. Specifically, I predicted a larger proportion of strict correct responses for participants who were told that coconut water was both necessary and unpredictable for questions framed as detecting stinginess and greediness, but not for questions framed as detecting outstanding debts or violations of descriptive rules, than for participants who received any of the other three combinations of features for coconut water (see *Figure 3*).

The dependent variable of strict correct responses is binary: correct or incorrect. The first independent variable, type of resource, was a between-subjects variable. In contrast, the second variable of interest, type of rule, was a within-subject variable because all participants saw all four types of rules. Thus, I chose generalized estimating equations with a binary logistic model to test my hypothesis. Before conducting my analyses, I removed participants who failed an attention-check. The attention-check asked participants to select the last option in a multiple-choice question to demonstrate that they were reading the instructions. Eleven participants failed this attention check, which resulted in a total number of 141 participants included in the analyses. I did not remove from the analyses any participants based on failing either of the two manipulation check questions that were presented at the end of the study.

Goodness of Fit

Before looking at the significance of the main effects and interaction of resource type and rule type on strict correct responses, it's important to consider correlation matrix structures and model fit. One index that gives insight into model fit is the QIC score (Quasi-likelihood under Independence Model Criterion). The QIC score allows for a comparison between different models, such that the smaller the QIC for a model, the better the fit.

First, I ran a model with only a main effect for rule type. The correlation matrix structure for rule type did not influence the QIC score, which was 669.83. Then, I ran a model with only a main effect for resource type. The QIC score was affected by the correlation matrix structure. I tested three potential correlation matrix structures that are common for this type of data: unstructured, which assumes there is no pattern in the correlation matrix, independent, which assumes the correlations are zero, and exchangeable, which estimates a single correlation for all cells in the matrix and assumes that all correlations are equal. The independent and exchangeable correlation matrix structures produced the same QIC score of 725.60, which suggested a better fit than the 729.31 QIC score for the unstructured correlation matrix structure. Interestingly, these QIC scores for the model with only a main effect for resource type are greater than the QIC score for the model with only a main effect for rule type. This suggests that the type of rule is more likely to have an effect on strict correct responses than the type of resource.

Next, I turned my attention to a model with the two main effects, but without the interaction effect. Although an unstructured correlation matrix structure produced the best fit, the QIC score changed less than one point no matter the correlation matrix

structure that was used. Therefore, there is no strong evidence for a matrix structure fitting the data better than the others. Given that the unstructured correlation matrix structure produced the best fit, even if by a narrow margin, and that it is the most general structure, I chose this correlation matrix structure for my analyses. The model with the two main effects, without the interaction effect, and with an unstructured correlation matrix structure, produced a QIC score of 677.82. This QIC score is greater than the QIC score of the model with only a main effect of resource type, but smaller than the QIC score of the model with only a main effect of rule type. These differences in QIC scores suggest that the addition of the main effect of resource type is detrimental to model fit. My hypothesis that resource type would affect the activation of greediness and stinginess detection mechanisms requires that there be an effect of resource type on strict correct scores, so this analysis of goodness of fit does not support my hypothesis.

I also ran a model that included a main effect of resource type, a main effect of rule type, and an interaction effect between these two variables with an unstructured correlation matrix structure. This model produced a QIC score of 685.59. I tested different correlation matrix structures to see whether they would produce better goodness of fit scores, but the goodness of fit did not change with different correlation matrix structures. The QIC for the model with only the two main effects was 677.82, which is smaller than the QIC score of 685.59 produced by the model that included the interaction effect. Smaller QIC scores suggest better fit, which means that the model without the interaction effect seems to fit the data better. My hypothesis that the combination of a resource being necessary and unpredictable would lead to the highest performance on questions for rules framed as detecting stinginess and greediness, but not on questions for

rules framed as detecting outstanding debts or violations of descriptive rules, necessitates that there be an interaction between the type of resource and the type of rule. Thus, the better fit suggested by the smaller QIC score for the model without the interaction effect is not consistent with my hypothesis.

Significance Testing

After looking at the goodness of fit indicators for the models, I ran the model with the main effect for rule type, the main effect for resource type, and the interaction effect to test for significance. The correlation matrix structure did not affect the goodness of fit for this model, so I used an unstructured correlation matrix structure, which is the most general structure. I predicted a larger proportion of strict correct responses for participants who were told that coconut water was both necessary and unpredictable for questions framed as detecting stinginess and greediness, but not for questions framed as detecting outstanding debts or violations of descriptive rules, than for participants who received any of the other three combinations of features for coconut water (see *Figure 3*). The results that would support this hypothesis would be to find a significant interaction between question type and resource type. However, the results suggested that this interaction was not significant; $X^2(9)=10.24$, $p>.05$ (see *Figure 10*). Thus, my hypothesis was not supported.

Aside from my main hypothesis, I expected to find results consistent with my previous findings that participants are better able to detect stinginess, greediness, and outstanding debt violations than violations of descriptive rules (Munoz et al., in prep). The results revealed that the main effect for resource type was not significant; $X^2(3)=0.51$, $p>.05$ (see *Figure 11*). In contrast, the main effect for rule type was

significant; $X^2(3)=34.68$, $p<.001$ (see *Figure 12*). These results suggest that cheater detection ability was not influenced by the features that were assigned to coconut water, but it was influenced by the type of violation that participants were asked to detect.

The significant main effect for rule type provides some evidence to support my previous findings. However, further analyses are needed to determine which rules elicited the best performance. Thus, I ran pairwise comparisons with Bonferroni adjustments and found that the proportion of strict correct responses for greediness detection (47%, $p<.001$), stinginess detection (61%, $p<.001$), and detection of outstanding debts (63%, $p<.001$) were all higher than strict correct responses for detection of descriptive rule violations (15%); $X^2(3)=54.57$, $p<.001$. These results are consistent with my previous research suggesting that people might have cognitive mechanisms for detecting greediness, stinginess, and outstanding debts (Munoz et al., in prep). I also found that there were no significant differences in strict correct responses between greediness, stinginess, and outstanding debt detection rules, which suggests that these three cheater detection mechanisms are similarly effective.

I also looked at the odds ratios while holding resource type constant to gain further insight into the main effect of rule type. The odds of a strict correct response was 6.79 higher given a rule about outstanding debt compared to a descriptive rule. The odds of a strict correct response was 4.20 higher given a rule about greediness compared to a descriptive rule. Finally, the odds of a strict correct response was 6.41 higher given a rule about stinginess compared to a descriptive rule. These odds ratios suggest that participants were better able to detect violations regarding outstanding debts, greediness,

and stinginess than violations about descriptive rules, which is consistent with my previous research (Munoz et al., in prep).

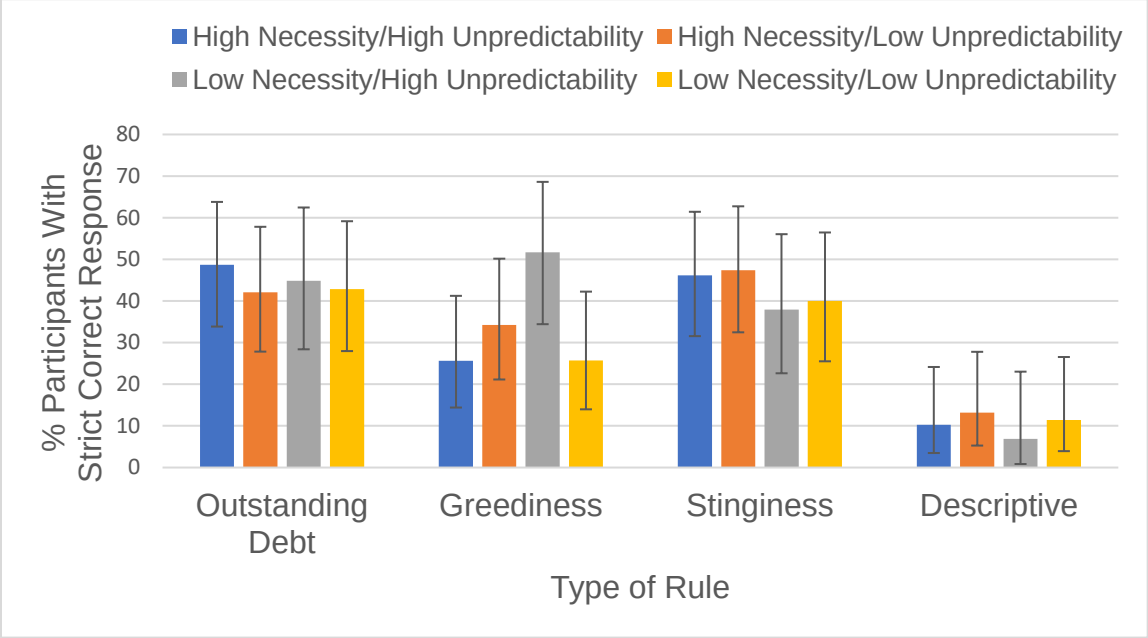

Figure 10. Percent of participants that achieved a strict correct response by type of resource and type of rule. The interaction between type of resource and type of rule was not significant.

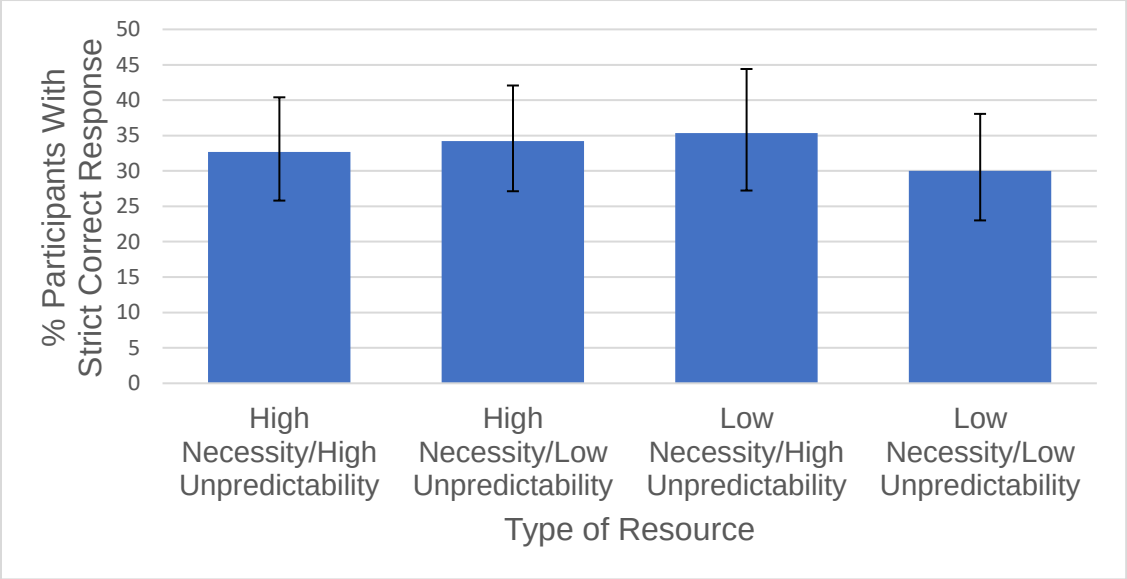

Figure 11. Percent of participants that achieved a strict correct response by type of resource. The main effect of resource type on detection ability was not significant.

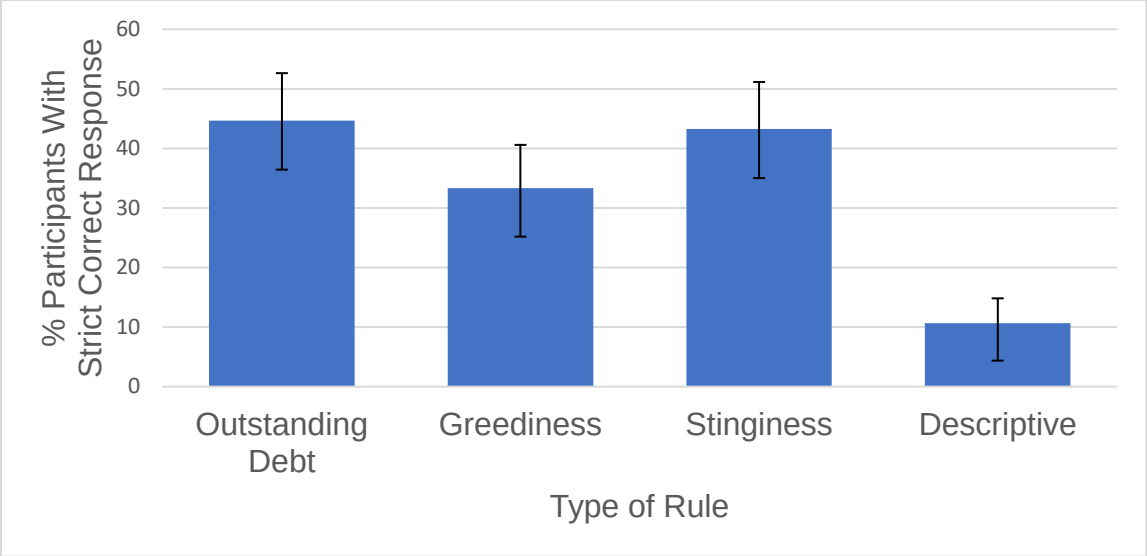

Figure 12. Percent of participants that achieved a strict correct response by type of rule. The main effect of rule type on detection ability was significant.

CHAPTER 4

DISCUSSION

The results showed that people performed better on stinginess, greediness, and outstanding debt detection than on detection of descriptive rule violations. This finding is consistent with my previous studies, which suggest that people might have cognitive mechanisms for detecting stinginess and greediness in addition outstanding debts (Munoz et al., in prep). However, the results did not support my hypothesis that resource type would have an effect on stinginess and greediness detection, such that a combination of high resource necessity and unpredictability would result in the highest performance for detecting greediness and stinginess, but not for detecting outstanding debts and violations of descriptive rules. Below I outline several explanations for why the results were not consistent with my hypothesis and I provide some directions for future research.

Benefits of Detection Even When Resources are Not Necessary and Unpredictable

One possibility for the lack of support of my hypothesis is that the benefits of monitoring for need-based transfer behavior sometimes outweigh the costs even in situations in which need-based transfer rules are unlikely to apply. Detecting violations of need-based transfer rules in situations in which need-based transfers don't apply might provide useful information about the relationship. For example, if you find that an acquaintance is behaving in a need-based transfer manner, even if need-based transfers are not required, it might signal a willingness to become need-based transfer partners and help to build trust before establishing a relationship in which need-based transfers for necessary and unpredictable resources can take place.

Imagine that an acquaintance offers you some of their chocolate while you're at the movies. Chocolate is not a necessary resource and it's not unpredictable; your survival won't be impacted if you don't eat chocolate and you can reliably buy it from the concession stand. Nonetheless, you might still ask for some chocolate from your acquaintance. If your acquaintance agrees to share some of their chocolate, this might serve as a way of building trust. Future studies should test whether the type of relationship that exists (friend, kin, acquaintance, stranger, etc.) with the person involved in a resource transfer — or the potential for a certain type of relationship — impacts the likelihood that different types of cheater detection mechanisms become active given different types of resources.

Future studies should also consider how information about violations of rules can be used in different ways. In some instances, information about violations can serve to punish cheaters. In other instances, information about violations can serve to update our estimation of other people's valuation of us and inform our decisions about how to invest our attention and resources. To tease apart these two functions of cheater detection mechanisms, it would be interesting to look at people's behaviors after detecting a violation of need-based transfer rules. In particular, it would be useful to find out whether people who detect greediness and stinginess for necessary and unpredictable resources are more likely to punish or stop interacting with the rule breaker than people who detect greediness and stinginess for resources that are not necessary and unpredictable. These findings would indicate that people gather information about adherence to need-based transfer rules regardless of the type of resource being transferred, but they act differently towards a target who violated the rules depending on the type of resource involved.

Conscious Control Over Cheater Detection Mechanisms

Another explanation for the results of this study is that people can consciously activate their cheater detection mechanisms, even in situations that would not normally be likely to trigger them. In this study, participants were explicitly asked whether a violation could have occurred, regardless of the type of resource in question. Therefore, if executive control over cheater detection mechanisms is possible, people would have been able to exert executive control to activate their stinginess and greediness detection mechanisms in response to the study's request to detect potential violations. If this is the case, we might expect to see that rates of activation of greediness and stinginess detection mechanisms depend on the type of resource in natural settings in which people are not explicitly asked to detect violations. Natural selection might have shaped these mechanisms to be on auto-pilot most of the time to reduce cognitive load. However, natural selection might have also allowed us to maintain the ability to exert conscious control over cheater detection mechanisms in case an unusual circumstance arises that requires us to do so. This would be comparable to our ability to exert conscious control over our breathing, even though it is usually an automatic process.

Imperfect Manipulation

Something else that should be considered when interpreting these results is that conflicting information might have been inadvertently provided to participants. Specifically, participants in the low necessity conditions were told that coconut water was not essential for survival at the beginning of the experiment. However, when these same participants were presented with stinginess and greediness rules, they were given information that implied that not having enough coconut water was a possibility. The

stinginess rule was: “If Chris has more coconut water than he needs, then Chris gives you some coconut water”. The greediness rule was: “If Chris has enough coconut water, then Chris doesn’t ask you for coconut water”. These rules could have conveyed to participants that coconut water was a resource that was necessary for survival by implying that there is some threshold of need for coconut water (stinginess rule) and that sometimes people don’t have enough coconut water (greediness rule).

The message that the stinginess rule aimed to convey by suggesting that Chris might have had more coconut water than he needed was that people can sometimes obtain a surplus of coconut water, such that they would have more coconut water than they could consume by themselves. The greediness rule was meant to convey that that people can sometimes have less coconut water than they would like. However, the wording for these two rules might not have been optimal for communicating the desired messages and might have instead led to some confusion. Thus, participants in the low necessity conditions might have been forced to choose whether to believe the instructions that specified that coconut water was not necessary for survival or the potential implications of the stinginess and greediness rules that coconut water was necessary for survival.

I included two questions at the end of the study that might shed some light into the effectiveness of the manipulation, which attempted to influence people’s perceptions about the necessity and unpredictability of coconut water. Moreover, these questions can help to examine whether there was confusion about whether coconut water was necessary in the low necessity condition. One of the questions at the end of the study gave a statement about coconut water being necessary on the island and asked participants whether this was true. The other question gave a statement about the acquisition of

coconut water on the island being predictable and asked participants if this was true. Participants had to answer either “true” or “false” to each of these questions. Participants were randomly assigned to four conditions that varied in terms of whether the instructions suggested that coconut water on the island was necessary (or not necessary) and unpredictable (or predictable). Therefore, the correct answer to these two questions at the end of the study varied by condition. For example, a participant who was assigned to the high necessity and high unpredictability condition, and who read and understood the instructions, should have answered “true” to the statement about coconut water being necessary and “false” to the statement about coconut water being predictable.

The results of these two questions at the end of the study show that, out of the 141 participants who passed the attention check, sixteen (11.35%) failed at least one of the manipulation checks. Two failed only the need manipulation check, ten failed only the predictability manipulation check, and four failed both manipulation checks. All six participants who failed the necessity manipulation check were in one of the two low necessity conditions. In other words, 9.4% of participants in the low necessity conditions failed the necessity manipulation check. However, most of the participants (four out of six participants) who failed the necessity condition also failed the predictability condition. This suggests that most of the participants who failed the necessity manipulation check might not have read the instructions carefully, which might have led them to fail both manipulation checks. If we consider the participants who failed the necessity manipulation check, but not the predictability manipulation check, this accounts for only 3.13% of participants (2 participants) in the low necessity conditions. Still, this is higher than the 0 participants who failed the necessity manipulation check in the high

necessity condition. Thus, there does seem to be some evidence for confusion about the necessity of coconut water in the low necessity conditions. Future studies that look at the effect of type of resource on stinginess and greediness detection should use language for the stinginess and greediness rules that does not suggest that a resource might be necessary for survival.

Another explanation for the lack of an effect of resource type on the activation of greediness and stinginess detection is that the manipulation might not have been strong enough. In this study, participants were asked to imagine that they live in an island where there is coconut water. Participants were randomly assigned to one of four different conditions. Each condition gave participants different information about whether coconut water was necessary and/or unpredictable. However, participants might have already had preconceptions about how unpredictable and necessary coconut water is. These preconceptions, which would stem from their own life experiences, might have been stronger than the manipulation. If this was the case, differences in greediness and stinginess detection activation by type of resource might exist, but they might not have been adequately captured by this study.

One way of testing whether the manipulation was inadequate would be to conduct a study in which the resources in each condition are selected to fit different levels of necessity and unpredictability as they are perceived in the participants' real environment. This method would allow less control of extraneous variables between conditions, but it might be more effective than trying to push around participants' perceptions of a resource through story-telling. For example, water could be used as a predictable and necessary resource for participants living in much of the United States. The problem with this

approach might be that many participants in wealthy industrialized nations might not perceive any resource to be both necessary and unpredictable. An alternative approach might be to ask participants to imagine that they live on an island and then presenting them with a resource that they're unfamiliar with. Presenting a resource that is unfamiliar, as opposed to a familiar resource such as coconut water, would avoid triggering preconceptions that participants might have about the resource in question and thus, it might allow for a more effective manipulation of the characteristics of the resource across conditions.

Measuring Level of Suspicion

Future studies should also consider that in the real world, information is often incomplete or inaccurate. In this study, participants were given facts to help them decide whether a violation of a rule could have been possible or not. For example, one of the questions told participants that "Chris had enough coconut water on this day". However, in the real world, Chris could have been hiding his coconut water. Therefore, cognitive mechanisms are likely to be designed to be skeptical even when presented with evidence of something being true. The binary decision given to participants between selecting that "yes, it's possible" that Chris violated the rule and "no, it's not possible" that Chris violated the rule does not adequately measure the level of suspicion that participants might have towards Chris given a specific situation. A more sensitive measure of suspicion level might be to ask participants to rate on a Likert scale the extent to which it is possible that Chris could have violated the rule. It's possible that a study employing this dependent variable might find that the level of suspicion of Chris for stinginess and greediness rules is higher for resources that are necessary and unpredictable than for

other types of resources.

Concluding Remarks

My hypothesis that resource type would have an effect on the activation of greediness and stinginess detection was not supported by the results of this study. There are several potential explanations for this lack of support. One option is that the benefits of monitoring for need-based transfer behavior sometimes outweigh the costs even in situations in which need-based transfer rules are unlikely to apply. Another explanation is that participants were able to exert conscious control over their stinginess and greediness detection mechanisms to activate them, even for resources that would not naturally trigger them in the real world. It's also possible that the results reflect a problem with the manipulation. More research is needed to determine which of these explanations, or what combination of these explanations, best accounts for the results of this study. Further research into what conditions trigger stinginess and greediness detection can help us to understand an important piece of the puzzle that is our ability to cooperate and trust each other. Resource transfers made through need-based transfers make us vulnerable to cheating, but thanks to greediness and stinginess detection, we can reap the benefits of this human behavior while often avoiding exploitation. These cognitive defenses make it possible to both give some of your water to a thirsty camper and to keep your delicious ice cream from a greedy stranger. [Insert your text here]

REFERENCES

Aktipis, C.A. (2011). Is cooperation viable in mobile organisms? Simple Walk Away strategy favors the evolution of cooperation in groups. *Evolution and Human Behavior, 32*(4) 263-276. doi: 10.1016/j.evolhumbehav.2011.01.002.

Aktipis, A., Cronk, L., Alcock, J., Ayers, J. D., Baciu, C., Balliet, D., … Winfrey, P. (2018). Understanding cooperation through fitness interdependence. *Nature Human Behaviour*, *2*(7), 429–431. https://doi.org/10.1038/s41562-018-0378-4

Aktipis, C. A., Cronk, L., De Aguilar, R. (2011). Risk-pooling and herd survival: An agent-based model of a Maasai gift-giving system. *Human Ecology, 39*(2) 131-140.

Aktipis, A., de Aguiar, R., Flaherty, A., Iyer, P., Sonkoi, D., & Cronk, L. (2016). Cooperation in an Uncertain World: For the Maasai of East Africa, Need-Based Transfers Outperform Account-Keeping in Volatile Environments. *Human Ecology, 44*, 353–364. https://doi.org/10.1007/s10745-016-9823-z

Aronow, P. M., Baron, J., & Pinson, L. (2015). A Note on Dropping Experimental Subjects Who Fail a Manipulation Check. *SSRN Electronic Journal.* doi:10.2139/ssrn.2683588

Axelrod, R. (1984). The Evolution of Cooperation. *Basic Books.*

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1610), 749–753. https://doi.org/10.1098/rspb.2006.0209

Brosnan, S. F., & De Waal, F. B. (2002). A proximate perspective on reciprocal altruism. *Human Nature*, *13*(1), 129–152.

Carter, G., & Wilkinson, G. (2013). Does food sharing in vampire bats demonstrate reciprocity? Communicative & Integrative Biology, 6(6). https://doi.org/10.4161/cib.25783

Carter, G. G., & Wilkinson, G. S. (2015). Social benefits of non-kin food sharing by female vampire bats. *Proc. R. Soc. B*, *282*(1819), 20152524. https://doi.org/10.1098/rspb.2015.2524

Clark, M. S. (1984). Record keeping in two types of relationships. *Journal of Personality and Social Psychology*, *47*(3), 549–557. https://doi.org/10.1037/0022-3514.47.3.549

Clark, M. S., & Mills, J. R. (2012). A Theory of Communal (and Exchange) Relationships. In P. Van Lange, A. Kruglanski, & E. Higgins, *Handbook of Theories of Social Psychology* (pp. 232–250). 1 Oliver's Yard, 55 City Road, London EC1Y

1SP United Kingdom: SAGE Publications Ltd.
https://doi.org/10.4135/9781446249222.n38

Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology, 37*(1), 12-24. doi: 10.1037/0022-3514.37.1.12

Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*(3), 187–276. doi: 10.1016/0010-0277(89)90023-1

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences, 107*(Supplement_2), 9007–9014. https://doi.org/10.1073/pnas.0914623107

Cosmides, L., & Tooby, J. (2015). Adaptations for Reasoning About Social Exchange. *The Handbook of Evolutionary Psychology*, 1-44. doi:10.1002/9781119125563.evpsych225

Cronk, L., Berbesque, J., Conte, T., Gervais, M., Iyer, P., McCarthy, B., … Aktipis, C. (2019). *Managing Risk Through Cooperation: Need-Based Transfers and Risk Pooling Among the Societies of the Human Generosity Project*. https://doi.org/10.1007/978-3-030-15800-2_4

Cronk, L., Steklis, D., Steklis, N., van den Akker, O. R., & Aktipis, A. (2019). Kin terms and fitness interdependence. *Evolution and Human Behavior*, *40*(3), 281–291. https://doi.org/10.1016/j.evolhumbehav.2018.12.004

Dahl, G., & Hjort, A. (1976). *Having herds: pastoral herd growth and household economy*. Dept. of Social Anthropology, University of Stockholm.

de Waal, F. B. (1997). The Chimpanzees service economy: Food for grooming. *Evolution and Human Behavior, 18*(6), 375-386. doi:10.1016/s1090-5138(97)00085-8

de Waal, F. B., & Berger, M. L. (2000). Payment for labour in monkeys. *Nature, 404*(6778), 563-563. doi:10.1038/35007138

Essock-Vitale, S. M., & Mcguire, M. T. (1985). Womens lives viewed from an evolutionary perspective. II. patterns of helping. *Ethology and Sociobiology, 6*(3), 155-173. doi:10.1016/0162-3095(85)90028-7

Fehr, E., U. Fischbacher, S. Gächter. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature, 13*, 1-25. https://doi.org/10.1007/s12110-002-1012-7

Fiske A.P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, 99, 689–723. doi: 10.1037/0033-295x.99.4.689

Gurven, M. (2004). To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, 27. https://doi.org/10.1017/S0140525X04000123

Hames, R. (1990) Sharing among the Yanomamö: Part 1, The Effects of Risk. In *Risk and Uncertainty in Ecology in Tribal and Peasant Economies*. E. Cashdan, ed., pp. 89-105. Westview Press: Denver.

Hamilton, W. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology, 7*(1), 1-16. doi:10.1016/0022-5193(64)90038-4

Hao, Y., Armbruster, D., Cronk, L., & Aktipis, C. A. (2015). Need-based transfers on a network: a model of risk-pooling in ecologically volatile environments. *Evolution and Human Behavior, 36*(4), 265–273. https://doi.org/10.1016/j.evolhumbehav.2014.12.003

Hill, Kaplan, Hawkes, & Hurtado. (1987). Foraging decisions among Aché hunter-gatherers: New data and implications for optimal foraging models. *Ethology and Sociobiology, 8*(1), 1-36. https://doi.org/10.1016/0162-3095(87)90055-0

Hruschka, D. J. (2010). *Friendship: Development, Ecology, and Evolution of a Relationship* (1st ed.). University of California Press. Retrieved from http://www.jstor.org/stable/10.1525/j.ctt1ppn90

Hruschka, D. J., & Silk, J. B. (2015). Cooperative Relationships. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, 1-16.

Hruschka, D., Hackman, J., & Macfarlan, S. (2015). Why do humans help their friends? Proximal and ultimate hypotheses from evolutionary theory. In *Evolutionary perspectives on social psychology* (pp. 255-266). Springer, Cham.

Kaplan, H., Hill, K., Cadeliña, R. V., Hayden, B., Hyndman, D. C., Preston, R. J., … Yesner, D. R. (1985). Food Sharing Among Ache Foragers: Tests of Explanatory Hypotheses [and Comments and Reply]. *Current Anthropology*, 26(2), 223–246.

Kruger, D. (2003). Evolution and altruism Combining psychological mediators with naturally selected tendencies. *Evolution and Human Behavior, 24*(2), 118-125. doi:10.1016/s1090-5138(02)00156-3

Munoz, A. E., Sznycer, D., Cronk, L., Aktipis, C.A., Stingy, Greedy, or Bad Lannister? Specialized Cheater Detection Mechanisms in Response to Different Selection Pressures.

Talk presented at the 2017 Human Behavior and Evolution Society Conference, Boise, ID, paper in prep.

Nowak, M.A. (2006). Five rules for the evolution of cooperation. *Science. 314*(5805) 1560–1563. doi: 10.1126/science.1133755

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology, 56*(1), 45–50. doi:10.4103/0301-4738.37595

Park, J. H., & Schaller, M. (2005). Does attitude similarity serve as a heuristic cue for kinship? Evidence of an implicit cognitive association. *Evolution and Human Behavior, 26*(2), 158-170. doi:10.1016/j.evolhumbehav.2004.08.013

Pietras, C., Cherek, D., Lane, S., & Tcheremissine, O. (2006). Risk reduction and resource pooling on a cooperation task. *The Psychological Record, 56*(3), 387-410. doi: 10.1007/BF03395557

Sahlins, M. D. (1972). *Stone age economics*. Chicago: Aldine-Atherton.

Schwinger, T., & Lamm, H. (1981). Justice Norms in Allocation Decisions: Need Consideration As A Function Of Resource Adequacy For Complete Need Satisfaction, Recipients' Contributions, And Recipients' Interpersonal Attraction. *Social Behavior and Personality: An International Journal*, *9*(2), 235–241. https://doi.org/10.2224/sbp.1981.9.2.235

Shackelford, T.K., & Buss, D.M. (1996). Betrayal in mateships, friendships, and coalitions. *Personality and Social Psychology Bulletin, 22*, 1151–1164.

Silk, J. B. (2003). Cooperation without counting. *Genetic and cultural evolution of cooperation*, 37-54.

Silk, J. (2002). Using the 'F'-word in primatology. *Behaviour, 139*(2), 421-446.

Sivan, J., Curry, O. S., & Lissa, C. J. (2018). Excavating the Foundations: Cognitive Adaptations for Multiple Moral Domains. *Evolutionary Psychological Science, 4*(4), 408-419. doi:10.1007/s40806-018-0154-8

Stone, V. E., Cosmides, L., Tooby, J., Kroll, N., & Knight, R. T. (2002). Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proceedings of the National Academy of Sciences, 99*(17), 11531-11536. doi:10.1073/pnas.122352699

Sugiyama, L. S., Tooby, J., & Cosmides, L. (2002). Cross-cultural evidence of cognitive adaptations for social exchange among the Shiwiar of Ecuadorian Amazonia.

Proceedings of the National Academy of Sciences, *99*(17), 11537–11542. https://doi.org/10.1073/pnas.122352999

Tooby, J. & Cosmides, L. (1996). Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism. *Proceedings of the British Academy. 88*, 119-143.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*(3), 273–281. https://doi.org/10.1080/14640746808400161

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, *308*(5955), 181–184. https://doi.org/10.1038/308181a0

Winterhalder, B. (1986). Diet choice, risk, and food sharing in a stochastic environment. *Journal of Anthropological Archaeology, 5*(4), 369-392. https://doi.org/10.1016/0278-4165(86)90017-6

Xue, M., & Silk, J. B. (2012). The role of tracking and tolerance in relationship among friends. *Evolution and Human Behavior*, 33(1), 17-25. doi: 10.1016/j.evolhumbehav.2011.04.004

APPENDIX A

EXPLORATORY TESTS: LESS STRICT CORRECT RESPONSES

To gain further insight, I considered a different measurement for my dependent variable. My planned results used strict correct responses as my dependent variable. A strict correct response is only counted as correct if all four questions, corresponding to the four possible logical propositions for a rule (P, Q, Not-P, and Not-Q) are all answered correctly. This dependent variable provides the best evidence for cheater detection mechanisms because it is incredibly unlikely that someone would be able to answer all four questions correctly by chance alone. I decided to consider a less stringent dependent variable in an exploratory test to determine whether resource type might have an effect on cheater detection ability. To do this, I created a new dependent variable, which I called "cumulative correct". This variable was computed by adding the number of correct responses out of the four possible questions for each given rule. For example, if a participant got the P, Q, Not-P, and Not-Q questions correct for a greediness detection rule, this participant was assigned a cumulative correct score of 4 for that rule. If the same participant got only the P and Q cards correct for the stinginess detection rule, this participant was assigned a cumulative correct score of 2 for that rule. Thus, the scores for the cumulative correct variable ranged from 0 to 4.

Then, I used generalized estimating equations with a model that included a main effect of resource type, a main effect of rule type, and an interaction effect of resource type and rule type with cumulative correct as my dependent variable. Cumulative correct is a count variable, so I used a Poisson loglinear model. The main effect of resource type was not significant; $X^2(3)=0.80$, $p>.05$. The main effect of rule type was significant; $X^2(3)=70.13$, $p<.05$. Thus, the significance of the main effects remained the same regardless of whether a cumulative correct dependent variable or a strict correct

dependent variable was used. The interaction between rule type and resource type was significant for the model with the cumulative correct dependent variable; $X^2(9)=17.17$, $p<.05$. This result contrasts with the model with the strict correct dependent variable, which did not show a significant interaction between rule type and resource type.

I ran pairwise comparisons with Bonferroni adjustments for the interaction effect between rule type and resource type to explore this significant interaction in more detail. A few interesting findings emerged from these pairwise comparisons. First, there were no differences between resource types when comparing within the same rule type. For example, there was no significant difference between cumulative correct responses when comparing a high necessity high unpredictability resource with a low necessity low unpredictability resource for the stinginess detection rule. These results suggest that the resource type did not affect the activation of cheater detection mechanisms, which does not support my hypothesis that greediness and stinginess detection performance would be higher for high necessity/high unpredictability resources than for the other resource types.

The pairwise comparisons also suggested that the stinginess detection rule (High Necessity/High Unpredictability: 78.85%, p<.001; High Necessity/Low Unpredictability: 77.63%, p<.01; Low Necessity/High Unpredictability: 75%, p<.01; Low Necessity/Low Unpredictability: 75.71%, p<.01) had a higher proportion of cumulative correct responses than the descriptive rule (High Necessity/High Unpredictability: 56.41%, High Necessity/Low Unpredictability: 51.32%, Low Necessity/High Unpredictability: 50.86%, Low Necessity/Low Unpredictability: 49.29%) for all resource types; $X^2(15)=120.72$, $p<.001$. This finding suggests that stinginess detection mechanisms can be activated regardless of the type of resource involved, which does not support my hypothesis. A

similar pattern emerged with outstanding debt detection. Outstanding debt detection (High Necessity/High Unpredictability: 80.77%, p<.001; High Necessity/Low Unpredictability: 73.68%, p>.05; Low Necessity/High Unpredictability: 73.28%, p<.05; Low Necessity/Low Unpredictability: 74.29%; p<.05) had a higher proportion of cumulative correct responses than the descriptive rule for all resource types, except for the high necessity/low unpredictability condition.

More relevant to my hypothesis, the proportion of cumulative correct responses for the greediness detection rule (59.62%) was significantly lower than the proportion of cumulative correct responses for both the stinginess detection rule (78.85%, p<.01) and the outstanding debt detection rule (80.77%, p<.01) and was not higher than the proportion of cumulative correct responses for the descriptive rule (56.41%) for the high necessity/high unpredictability resource condition. The proportion of cumulative correct responses for the greediness detection rule was not significantly different from the stinginess and outstanding debt detection rules for any of the other resource conditions. These results are contrary to my hypothesis since I expected that greediness detection performance would be best for resources that are highly necessary and unpredictable. Although not expected, this result could be explained in two different ways.

The first explanation, which I will refer to as the relationship maintenance explanation, assumes that participants believed Chris, the fictitious person in the experiment that was presented in the questions, to be highly interdependent with them. Participants were not given any specific information about who Chris was or what their relationship to Chris might have been. However, participants might have inferred that Chris was another person on the island with whom they were interdependent. Participants

were asked to imagine that they were in a primitive island where resources couldn't be accessed by going to a supermarket, but instead had to be gathered. Our participants might have inferred that being in such an unfamiliar environment required them to be interdependent with others on the island, such as Chris.

People in the real world often have incomplete or incorrect information. Therefore, it is possible to wrongfully accuse someone of being greedy. A person might accuse someone of being greedy only to find out later that the accused didn't have enough resources to sustain themselves. Wrongfully accusing someone of being greedy could damage a relationship that is essential. Moreover, denying someone essential resources when they are in need can be severely detrimental even beyond damaging the relationship when positive fitness interdependence is high. By denying them resources, a person that you depend on could become weak or even perish without your aid. Therefore, greediness detection of essential resources when fitness interdependence is high would likely pose risks that are higher than the protective benefits it can bestow. Given these high risks, people in these situations might suppress greediness detection and instead rely on an estimation of fitness interdependence to decide whether to give resources.

A second explanation for the results, which I will call the suspicion explanation, involves a high activation of greediness detection mechanisms when the resource involved is necessary and unpredictable. This high activation, which would be consistent with my hypothesis, could lead participants to be highly suspicious of Chris and of the information provided about Chris. For example, when told that "Chris didn't have enough coconut water on this day", participants might have been suspicious of that information

and thought that Chris might be hiding some of his coconut water. Similarly, when told that “Chris didn’t ask you for coconut water on this day”, participants might have thought that maybe Chris was not explicitly asking for coconut water, but he was implying that he wanted some in other ways. If participants were highly suspicious of Chris, they might have preferred not to select the option that said that “No, it’s not possible” that Chris violated the rule even if the evidence suggested that this was the case.

APPENDIX B

EXPLORATORY TESTS: DETECTION THEORY

A strict correct response involves responding correctly to all four logical propositions for a rule (P, Q, Not-P, and Not-Q). Thus, strict correct responses evaluate whether people can both correctly identify situations in which a violation could have occurred (P and Not-Q) and correctly avoid suspecting a violation when a violation is impossible (Not-P and Q). However, you could consider these two situations separately to determine whether each detection mechanism performed better in one of these domains than in the other.

This detection theory approach is often used for evaluating medical diagnosis instrumentation, but it can be applied to other domains as well (Parikh, Mathai, Parikh, Chandra Sekhar, & Thomas, 2008). This approach involves classifying responses as true positives, true negatives, false positives, and false negatives. Responses labeled as "true" mean that the participant gave a correct response, while responses labeled as "false" mean that the participant gave an incorrect response. Responses given for situations labeled as "positive" mean that a violation was possible in that situation (P and Not-Q), while the label "negative" means that a violation was not possible in that situation (Q and Not-P).

A true positive is a situation in which a violation of the rule could have occurred, and the participant correctly identifies that a violation is possible. A true negative is a situation in which a violation of the rule is not possible, and the participant correctly identifies that a violation is not possible. In a false positive response, a violation of the rule is not possible, but the participant incorrectly states that a violation is possible. In a false negative response, a violation of the rule is possible, but the participant incorrectly states that a violation is not possible.

In the previous section, I used a cumulative correct dependent variable that was less stringent than the strict correct dependent variable used in the planned analyses described in the results section. An analysis of the model with the cumulative correct dependent variable suggested that, contrary to my hypothesis, the proportion of cumulative correct responses for the greediness detection rule (59.62%) was significantly lower than the proportion of cumulative correct responses for both the stinginess detection rule (78.85%, $p<.01$) and the outstanding debt detection rule (80.77%, $p<.01$) and was not higher than the proportion of cumulative correct responses for the descriptive rule (56.41%) for the high necessity/high unpredictability resource condition. In the previous section, I describe two possible explanations for these results.

The first possible explanation is that people might have assumed a high level of interdependence with their resource transfer partner. If this was the case, they might have estimated the cost of incorrectly accusing their transfer partner of being greedy when the resource was necessary and unpredictable to be high, which could have led them to err on the side of caution. If so, an analysis through a detection theory lens would predict a higher rate of false negatives than false positives for greediness detection when the resource was necessary and unpredictable.

In the previous section, I also discuss a second potential explanation for the greediness detection results. It's possible that people's greediness detection mechanism, consistent with my hypothesis, was highly active in the condition where resources were necessary and unpredictable. If this was the case, people might not have trusted the information provided in the experiment about Chris, leading to several false positives and few false negatives. Therefore, the suspicion explanation would predict a higher rate of

false positives than false negatives, while the relationship maintenance explanation would predict a higher rate of false negatives than false positives for greediness detection when resources are necessary and unpredictable.

To test these competing explanations and analyze the data through a detection theory approach, I created a new independent variable and a new dependent variable. I created a binary independent variable with a label of "1" for situations in which a violation could have been possible (P and Not-Q) and a label of "0" for situations in which a violation was not possible (Not-P, and Q). I labeled this new variable "Positive/Negative" as a reference to the detection theory framework outlined earlier, in which positive is a situation in which a violation is possible and negative is a situation in which a violation is not possible.

I also created a new dependent variable that measured whether each response was correct (labeled "1") or incorrect (labeled "0"). This new dependent variable was different from my prior strict-correct dependent variable because for this new variable, I considered whether each individual question was correct, rather than coding a response as correct only if all four questions for a given rule were correct. I labeled this new dependent variable "Simple Correct". A data point with a "1" for the Positive/Negative variable and a correct Simple Correct response corresponded to a true positive. A true negative was a data point that had a "0" for the Positive/Negative variable and a correct Simple Correct response. A data point for an incorrect Simple Correct response with a "1" for the Positive/Negative variable was a false positive and a data point with an incorrect Simple Correct response with a "0" for the Positive/Negative variable was a false negative.

To test if correct responses depended on whether a violation was possible or not, I used generalized estimating equations with a binary logistic model that included a main effect of resource type, a main effect of rule type, a main effect of Positive/Negative, a two-way interaction between resource type and rule type, a two-way interaction between resource type and Positive/Negative, a two-way interaction between rule type and Positive/Negative, and a three-way interaction between resource type, rule type, and Positive/Negative. I used Simple Correct responses as my dependent variable.

The results of this model showed a significant main effect of rule type ($X^2(3)=52.28$, $p<.001$) and a significant main effect of Positive/Negative ($X^2(1)=4.98$, $p<.05$). The significant main effect of rule type is consistent with the results of the model with the strict correct dependent variable reported in the results section. The main effect of Positive/Negative suggests that there were significantly more false positives (32.54%) than false negatives (31.92%) overall. However, the small percentage difference between false positives and false negatives, in addition to the small Wald Chi-Square value, suggests that this is a small effect. The interaction between rule type and resource type was significant ($X^2(9)=21.15$, $p<.05$), which is consistent with the results of the model with a cumulative correct dependent variable reported in the appendix, but not consistent with the results of the model with a strict correct dependent variable reported in the results section. The interaction between rule type and Positive/Negative was also significant ($X^2(3)=16.30$, $p<.01$). The main effect of resource type and the two-way interaction between resource type and Positive/Negative were not significant.

The three-way interaction between resource type, rule type, and Positive/Negative was not significant. Nonetheless, I ran pairwise comparisons with Bonferroni adjustments

on this three-way interaction to examine whether there was a difference between false positives and false negatives among greediness detection questions for the high necessity/high unpredictability resource condition and I found no significant difference. These results do not support either the relationship maintenance explanation, which would require higher false negatives, or the suspicion explanation, which would require higher false positives. It's possible that the lack of a significant difference in cumulative correct responses between the greediness detection rule and the descriptive rule for highly necessary and highly unpredictable resources is due to chance. Thus, it is important for a follow-up study to find out whether this result replicates.

75

APPENDIX C

EXPLORATORY TESTS: CARD-SPECIFIC ANALYSIS

Each rule in this study is composed of four logical components: true antecedent (P), false antecedent (Not-P), true consequent (Q), and false consequent (Not-Q). Participants in this study saw four questions for each rule, corresponding to the four possible logical components for each rule. Strict correct responses, which are used as the dependent variable in the responses reported in the results section, require that a participant answer all four of these questions correctly to be counted as a correct response. However, it is also possible to conduct a more nuanced analysis of the results by considering each question separately. This type of analysis allows us to examine whether the type of logical component has an influence on participants' ability to provide a correct response. Thus, we can learn more detailed information about the types of mistakes and the cognitive architecture of each detection mechanism.

To conduct this analysis, I created a new independent variable, which I labeled "card type". This variable used the numbers one to four to identify the type of logical component (P, Not-P, Q, or Not-Q) for each question. Then, I used generalized estimating equations with a binary logistic model that included a main effect of resource type, a main effect of rule type, a main effect of card type, a two-way interaction between resource type and rule type, a two-way interaction between resource type and card type, a two-way interaction between rule type and card type, and a three-way interaction between resource type, rule type, and card type. I used Simple Correct responses as my dependent variable.

The results of this model showed a significant main effect of rule type; $X^2(3)=70.60$, $p<.001$. There was also a significant interaction between rule type and card type; $X^2(9)=50.61$, $p<.001$. The main effect of resource type, the main effect of card, the

two-way interaction between rule type and resource type, the two-way interaction between resource type and card type, and the three-way interaction between rule type, resource type, and card type, were not significant.

I also ran pairwise comparisons with Bonferroni adjustments on the two-way interaction between rule type and card type to examine this interaction further. One of the interesting patterns that emerged after running the pairwise comparisons was that when the antecedent was not met (Not-P), the proportion of simple correct responses was not significantly different across different rule types. In contrast, the proportion of simple correct responses for rules about stinginess (80.14%, p<.001), greediness (73.05%, p<.001), and outstanding debt detection (84.40%, p<.001) was significantly higher than the proportion of simple correct responses for the descriptive rule (36.17%) when the consequent was not met (Not-Q); $X^2(15)=154.62$, $p<.001$. The proportion of simple correct responses for rules about stinginess (78.72%, p<.001) and outstanding debt detection (75.89%, p<.01) was also significantly higher than the proportion of simple correct responses for the descriptive rule (52.48%) when the consequent was met (Q). Only the proportion of simple correct responses for the rule about stinginess detection (75.89%, p<.05) was significantly higher than the proportion of simple correct responses for the descriptive rule (57.45%) when the antecedent was met (P).

Also interesting was that the proportion of simple correct responses for the descriptive rule when the consequent was not met (Not-Q: 36.17%) was significantly lower than the proportion of simple correct responses for the descriptive rule when the antecedent was either met (P: 57.45%) or not met (Not-P: 62.41%). In contrast, the proportion of simple correct responses for the stinginess, greediness, and outstanding

debt detection rules did not differ by logical category. These results suggest that the higher proportion of correct responses for stinginess, greediness, and outstanding debt detection rules when compared to the descriptive rule might be primarily driven by cheater detection mechanisms increasing the recognition that a rule could have been violated when the consequent is not met. These results are consistent with the findings reported by Wason, who found the lowest performance for Not-Q cards in the Wason selection card out of all four possible cards (1968). Thus, both the data from this study and from the classic reasoning literature find that people are particularly bad at falsifying the consequent, but this deficiency disappears when cheater detection rules are activated.

APPENDIX D

EXPLORATORY TESTS: HIGH NECESSITY VS LOW NECESSITY

The results suggested that there was no main effect of resource type on cheater detection ability and no significant interaction between resource type and rule type. Nonetheless, I decided to exclude unpredictability and test whether considering need by itself resulted in a significant effect for resource type.

To run this new test, I created a new independent variable and labeled it "Necessity". To create the Necessity variable, I labeled both the High Necessity/High Unpredictability and the High Necessity/Low Unpredictability conditions as a "0" and both the Low Necessity/High Unpredictability and the Low Necessity/Low Unpredictability conditions as a "1". Then, I used generalized estimating equations with a binary logistic model that included a main effect of Necessity, a main effect of rule type, and an interaction between Necessity and rule type. I used strict correct responses as my dependent variable. The main effect of resource type was not significant; $X^2(1)=0.10$, $p>.05$. In contrast, the main effect of rule type was significant; $X^2(3)=34.84$, $p<.05$. Moreover, the interaction between rule type and resource type was not significant; $X^2(3)=2.93$, $p>.05$. Therefore, the significance of the main effects and interaction were the same regardless of whether the resource variable considered only necessity or both necessity and unpredictability.

APPENDIX E

EXPLORATORY TESTS: INCLUSION AND EXCLUSION OF PARTICIPANTS

The tests described in the results section were conducted with 141 participants who passed the attention check. Eleven participants were excluded from these analyses because they did not pass the attention check question. However, it's possible that there was something unique about the participants who didn't pass the attention check that is significant to the interpretation of the results and that by excluding them, the results were skewed. Thus, I decided to examine whether the results changed in any way if I were to include all participants in the analyses.

I used generalized estimating equations with a binary logistic model that included a main effect of resource type, a main effect of rule type, and an interaction effect of resource type and rule type with strict correct responses as my dependent variable. For this analysis, I included all participants, regardless of whether they passed the attention check or not. The main effect of resource type was not significant; $X^2(3)=0.66$, $p>.05$. In contrast, the main effect of rule type was significant; $X^2(3)=32.83$, $p<.05$. Moreover, the interaction between rule type and resource type was not significant; $X^2(9)=11.28$, $p>.05$. Therefore, the inclusion of participants who did not pass the attention check in the analyses made no difference with regards to the significance of the main effects and interaction.

In addition to the attention check, I included two questions at the end of the study that served as a manipulation check. These two true or false questions were intended to measure whether participants understood that coconut water was necessary, not necessary, predictable, or unpredictable based on the condition that they were assigned to. Out of the 141 participants who passed the attention check, sixteen failed at least one of the manipulation checks. Participants who failed the manipulation check were included

in the tests reported in the results section. However, there are points in favor and against excluding these participants.

On the one hand, excluding participants who failed the manipulation check could produce results that are unrepresentative of the general population (Aronow, Baron, & Pinson, 2015). There might be something unique about participants who failed the manipulation check that is relevant to the analyses and thus, excluding them could have biased the results. To avoid this potential bias, I decided to include the participants who failed the manipulation check in the analyses reported in the results section. On the other hand, excluding these participants might provide better evidence for whether the theoretical independent variable of interest, in this case resource type, has an effect on the dependent variable of interest.

If the goal of the study were to assess the effect of the measurement instrument itself on a dependent variable, participants who failed the manipulation check should certainly be included in the analyses. Excluding these participants would misrepresent the results because it would not consider that some participants were not properly affected by the manipulation or treatment. However, if the purpose of the study is to assess the effect of a theoretical construct that is imperfectly induced through a manipulation, as is the case in this study, including participants that failed the manipulation check would obscure the relationship between the construct and the dependent variable.

It could be costly to fail to detect a real effect between two theoretical variables due to the imperfect nature of the manipulation. If a study tested the effects of coffee on productivity and the researchers found that some of the participants did not drink the coffee that was provided to them, it would make sense to exclude the participants who

didn't drink the coffee from the analyses. Much like the participants who didn't drink the coffee in the previous example, some of the MTurk participants in this study might have skipped the instructions or they might not have read them carefully. Including these participants in the analyses would muddy the potential effect of resource type on strict correct responses because these participants didn't receive the manipulation.

I decided to investigate whether the effect of resource type on strict correct responses changed if the participants who failed the manipulation check were excluded from the analyses. To do this, I excluded all sixteen participants who failed the manipulation check and used generalized estimating equations with a binary logistic model that included a main effect of resource type, a main effect of rule type, and an interaction effect of resource type and rule type with strict correct responses as my dependent variable. The main effect of resource type was not significant; $X^2(3)=1.07$, $p>.05$. In contrast, the main effect of rule type was significant; $X^2(3)=36.77$, $p<.05$. Moreover, the interaction between rule type and resource type was not significant; $X^2(9)=11.13$, $p>.05$. Therefore, the exclusion of the participants who did not meet the manipulation check made no difference with regards to the significance of the main effects and interaction.

APPENDIX F

EXPLORATORY TESTS: ORDER EFFECTS

The order in which the four different types of rules (greediness detection, stinginess detection, outstanding debt detection, and detection of violations for a descriptive rule) were presented to participants was randomized through Qualtrics' randomizer tool. Thus, we would expect that the significant effect of rule type on strict correct responses would not be an artifact of the presentation order. Nonetheless, I decided to test whether there were order effects that could account for this effect. To test this, I created a new independent variable that identified the order in which each rule was presented to each participant. For example, if the greediness detection rule was presented first to a participant, a "1" was assigned to the order variable for that participant for that rule. If the same participant saw the stinginess detection rule third, a "3" was assigned to the order variable for that participant for that rule.

Then, I used generalized estimating equations with a binary logistic model that included a main effect of rule type, a main effect of order, and an interaction effect of rule type and order with strict correct responses as my dependent variable. The main effect of order was not significant; $X^2(3)=0.17$, $p>.05$. In contrast, the main effect of rule type was significant; $X^2(3)=35.59$, $p<.05$. Moreover, the interaction between rule type and resource type was not significant; $X^2(9)=5.11$, $p>.05$. Therefore, the order in which the rules were presented to participants did not have an effect on strict correct responses.

APPENDIX F

PERMISSIONS

ASU Knowledge Enterprise Development

EXEMPTION GRANTED

C Athena Aktipis
Psychology
-
aktipis@asu.edu

Dear C Athena Aktipis:

On 3/15/2016 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Cheater Detection for Need-Based Transfer and Account-Keeping Norms of Exchange
Investigator:	C Athena Aktipis
IRB ID:	STUDY00003927
Funding:	Name: Templeton Foundation, Grant Office ID: ID46724
Grant Title:	ID46724;
Grant ID:	ID46724;
Documents Reviewed:	• Questionnaire Cheater Detection, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Consent Cheater Detection, Category: Consent Form; • Templeton Grant Human Generosity Project, Category: Sponsor Attachment; • Recruitment Materials SONA Cheater Detection.pdf, Category: Recruitment Materials; • Cheater Detection NBT and AK IRB Protocol, Category: IRB Protocol;

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 3/15/2016.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).