THE UNIVERSITY OF CHICAGO

WHICH SIDE ARE YOU ON? HOW FRIENDS BALANCE COMPETING OBLIGATIONS OF LOYALTY AND MORALITY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

ANAM KHAN BARAKZAI

CHICAGO, ILLINOIS

JUNE 2019

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Thank you to my Mum for her love, support, and strength. You are truly inspiring.

Thank you to my older brother, Arham, for stepping in and taking on way more responsibilities than he ever should have had to. I owe so much to you.

Thank you to my younger brother, Syem, for the laughs, the belief he held in me, and for being one of my closest friends. I really am so grateful for our friendship.

Thank you to my advisor, Dr. Alex Shaw. You took a chance on me when I was lost and I could not have completed this journey without you. I feel very lucky to have had the opportunity to work with and learn from you. Your brilliance is unparalleled.

Thank you to my dear friends – those from home, those I have met along the way, and those here in Chicago. You all provided much inspiration for the work in this dissertation.

And finally, thank you to Barnaby. You are the purest definition of love and friendship. You embody all that comes with the phrase '(wo)man's best friend'.

ABSTRACT

Friendships are integral to healthy social functioning and successful navigation in our social world. They provide a plethora of benefits, but alongside these benefits, friendships bring certain obligations. Therefore, identifying good friends – and being a good friend – is extremely important. So, what makes for a good friend? Although it may seem intuitive that people would value those who are helpful, loyal, and honest, my dissertation will suggest that in some circumstances these otherwise noble values can be at odds with being a good friend. In Chapter 1, I investigate cases in which people respond to their friend being helpful and generous. I find that people respond negatively to a friend who is more helpful and generous overall, preferring their friend be less helpful and generous when the recipient of the prosociality is another friend (but not a family member or romantic partner). Chapter 2 explores people's inferences about their friend based on how the friend takes sides during a conflict in which a clear transgression was committed. I find that although people do expect loyalty from their friends, they are more accepting of disloyalty after having committed a moral transgression. Moreover, people make inferences about their friend's morality based on the friend's side-taking decision. Finally, Chapter 3 examines whether honesty is indeed the best policy among friends. I find that while benevolent dishonesty is acceptable from both friends and acquaintances, critical honesty is only acceptable from friends. Taken together, this work suggests that although helpfulness, loyalty, and honesty are highly valued and desired characteristics in friends, people sometimes prefer their friend to be less helpful, less loyal, and less kind.

INTRODUCTION

Friendships are integral to healthy social functioning and successful navigation in our social world. Through conversations and experiences with our peers, we learn perspective-taking, moral-reasoning, and social categorization, among other things (Allen, 1981; Gottman, Gonso, & Rasmussen, 1975; Schonert-Reichl, 1999). Friendships are often an indispensable source of pleasure and support, with 36% of unmarried young adults classifying friendship as their closest and most intimate current relationship (Berscheid, Snyder, & Omoto, 1989). Indeed, good friends provide a plethora of benefits such as social and emotional support (Argyle & Henderson, 1984; Fleming & Baum, 1986; Lang & Carstensen, 1994; Tukuno, 1983), job opportunities, (Bayer, Ross, & Topa, 2008; Cappellari, Lorenzo, Tatsiramos, & Konstantinos, 2011; Pistaferri, 1999), practical assistance and cooperation (Peck, 1993), and acceptance, capitalization, honesty, and respect (Gable & Reis, 2010; Frei & Shaver, 2002; Parks & Floyd, 1996). Moreover, friendships have been found to provide a buffer against many health issues such as chronic stress and dementia, and have been associated with lower mortality risks (Cohen, Gottlieb, & Underwood, 2004; Cole, Capitanio, Chun, Arevalo, Ma, & Cacioppo, 2015; Holt-Lunstad, Smith, & Baker, Harris, & Stephenson, 2015; Holwerda, Deeg, Beekman, Tilburg, Stek, Jonker, & Schoevers, 2012; Yang, Boen, Gerken, Li, Schorpp, & Harris, 2016). Therefore, identifying good friends – and being a good friend – is extremely important.

What makes for a good friend? My dissertation tackles this very broad question by examining how people feel about their friends being helpful and generous, taking sides during conflicts, and providing honest but critical feedback. Although it may seem intuitive that people would value those who are helpful (e.g., Barclay, 2016), impartial (e.g., Shaw, 2013), loyal (Keller, 2007), and kind (e.g., Telfer, 1970), my dissertation suggests that in some circumstance

these otherwise noble values can be at odds with being a good friend. Obligations such as loyalty (i.e. one's alliances) and morality (one's belief system about what is right and what is wrong) are powerful motivations that can often clash with one another in everyday interactions among friends (DeScioli & Kurzban, 2013). In this dissertation, I explored numerous cases when this may occur, when people are required to make difficult decisions that will impact not only how they perceive their friendships and their friends, but also how they are perceived by their friends and others. Through vignette-based and behavioral studies, I explored how people felt when their friends show preferential prosociality to another person, responded to their conflicts, and gave honest but critical feedback. I argue that in many of these cases, a good friend is less generous and cooperative, less loyal, and less kind.

In one set of studies, I investigated cases in which people responded negatively to their friend helping another person. To see this intuition, imagine that your friend does not help you and a week later has the opportunity to help another friend. Do you like your friend better when they help the other friend or when they don't help the other friend? Despite the fact that the former friend is more helpful overall (she did not help you but helped another vs. she helped neither of you), this decision is not so easy; when it comes to evaluating our friends, we not only want friends who are helpful and generous generally, we want friends who will be preferentially helpful, generous, and loyal to us (DeScioli & Kurzban, 2009a). In several studies, I predicted and found that people responded negatively to a friend who was more helpful and generous to another person, preferring instead that a friend be less helpful and generous overall. Importantly, this preferential prosociality was viewed as particularly negative when the helped party was another friend versus a relative or a romantic partner. I argue that feeling more upset when a friend helps another friend, as opposed to a relative or significant other, is a functional

manifestation of friendship jealousy; people see the other friend as a potential interloper who may displace their role in their friend's life.

In a second series of studies, I explore people's inferences about their friend's loyalty and morality based on how the friend decides to take sides during a conflict in which a clear transgression has been committed. For example, imagine that your good friend was caught stealing a large amount of money from work. Here, being loyal to your friend can mean ignoring your moral principles and upholding your moral principles can mean being disloyal to your friend (DeScioli & Kurzban, 2013; Shaw, 2013; Waytz, Dungan, & Young, 2013) and damaging your friendship (Shaw, DeScioli, Barakzai, & Kurzban, 2017). On the surface, this appears to be an inextricable dilemma in which you must choose between sacrificing your friendship or your morality. Is there a way out of this dilemma? Will your friend be upset with you if you choose to oppose them? In order to answer this question, my research examines how people involved in moral conflicts feel about their friends taking sides in such situations and suggests that, for friends, this dilemma may not be as difficult as it appears. I argue that although people do expect loyalty from their friends, they also understand that a person's moral commitments might make even a loyal friend blush at taking their side. Thus, although people expect support from their friends all else being equal, they may be substantially less upset at a friend's failure to take their side when they have committed a clear moral transgression (e.g., stealing a large amount of money from work). In a series of studies, I investigated how people felt about their friends when they make such side-taking decisions and the inferences that they make about their friend – the side-taker –'s morality—if the side-taker opposes, supports, or remains neutral on a moral issue, do people make assumptions about the side taker's loyalty or moral commitments?

In the third set of studies, I tested when and where honesty is (or is not) the best policy among friends. Although honesty is said to be an important value in interpersonal relationships (Altman & Taylor, 1973; LaFollette & Graham, 1986; Pearce 1974; Weiselquist, Rusbult, Foster, Agnew, 1999), some have suggested that in many circumstances we only value honesty because being honest is aligned with being prosocial (Van Lange & Kuhlman, 1994). Indeed, when the two are disassociated, people often prefer that others lie in a benevolent way (i.e. tell prosocial lies) rather than to be honest in an harmful way (Levine & Schweitzer, 2014, 2015). Moreover, benevolence is an important trait that we seek in friends. From a young age, people expect greater kindness from closer friends (Clark & Bittle, 1992), and benevolence has been described as the core characteristic of close friends (Kapur, 1991). If benevolence is such an important part of one's reaction to honesty, then one could imagine that people would react more negatively to brutal honesty from a friend because they expect friends to be particularly benevolent toward them. However, it could also be that we expect greater honesty from those closer to us. Indeed, Kant argued that, since friendship is a moral duty, dishonesty in friendships both signals and contributes to mistrust and weakening of the friendship (Kant & Infield, 1963). So do we ever value honesty for honesty's sake? In the third chapter, I used vignette-style experiments to examine one context where people may find honesty to be more valuable from friends: in feedback from another person. In three studies, I found that a prosocial lie (e.g., "I like your haircut") was acceptable from both close friends and acquaintances. But that a brutal truth (e.g., "Your haircut is unflattering") was only acceptable from friends. These results suggest that we tolerate honesty that may be hurtful more from friends than others. I also found that people believe friends to have more pure motivations than acquaintances when conveying brutal honesty.

CHAPTER 1: FRIENDS WITHOUT BENEFITS: WHEN PEOPLE REACT NEGATIVELY TO GENEROUS AND FAIR FRIENDS.

Having friends is key to thriving in the social world, which is why it is critically important that people be able to identify good, reliable friends. In general, people like those who are helpful and generous more than those who are not (for review, see Barclay, 2016). Therefore, it seems obvious that you would like someone better if that person helped a friend rather than if they failed to help a friend. However, imagine that your friend had not helped you the week before. Now, do you like this friend better when this friend helps another friend or if this friend helps neither of you? Despite the fact that the former friend is more helpful, this decision is not so easy because when it comes to evaluating our friends, we not only want friends who are helpful and generous generally, we want friends who will be preferentially helpful and loyal to us (DeScioli & Kurzban, 2009a). If preferential help is taken as a signal that we are less valued by a friend, then we might respond quite negatively to a friend who is more helpful. Here we investigate such cases, but first briefly review two broad accounts of friendship that make different predictions about how someone might respond to cases like the one above.

Many accounts of partner choice and friendship focus on cooperation, holding that people select friends based on the cooperative benefits they provide. People can select partners based on a number of dimensions that track the partner's willingness and ability to confer benefits and impose costs, including formidability, and attractiveness (Barclay, 2016; Lukaszewski, Simmons, Anderson, & Roney, 2016; Noë & Hammerstein, 1995; Sell, 2011; Virgil, 2007). However, a large portion of the literature has focused on one particular dimension: how prosocial or cooperative the partner is. These cooperative accounts argue that direct reciprocity (we like and help those who help us), indirect reciprocity (we like and help those who help others), and

partner choice drive us to select partners and friends based on the cooperative benefits that they provide (Baumard, André, & Sperber, 2013; Kenny, Mohr, & Levesque, 2001; Nelson, 2002; Peck, 1993; Rand & Nowak, 2013). In line with these ideas, we know that people evaluate others positively for being helpful and generous (Barclay, 2013; Delton, Krasnow, Cosmides & Tooby, 2011; Gurven & Winking, 2008, Panchanathan & Boyd, 2004; Shaw, DeScioli, & Olson, 2012) and that people attempt to signal to others that they are cooperative when trying to attract partners (Andreoni & Berheim, 2009; Barclay & Willer, 2007; Reis & Gruzen, 1976). These cooperative accounts suggest that people select partners based on their cooperativeness and generosity, both to the agent specifically and to other agents more generally (for review, see Hess & Hagan, 2006).

In relation to scenarios like the one outlined above, cooperative models argue that an agent should respond more negatively to a friend who helps no one rather than a friend who helps at least one person. Indeed, in this case, helping the other friend inflicts no additional direct costs on the agent and so the friend either effectively "defects" against two agents or only one agent. These cooperative accounts of friendship should therefore predict a more negative response to the person who helps no one.

However, other models of friendship make different predictions because they hold that a good friend is not only helpful in general, but is also more likely to prioritize helping the individual than someone else (DeScioli & Kurzban, 2009a; Tooby & Cosmides, 1996). One such model, the alliance account, argues that people should care about how their friend ranks them relative to others, not only about their friend's overall kindness (DeScioli & Kurzban 2009a). Consistent with this, DeScioli and Kurzban (2009a) found that a person's perceived rank among their friends was the strongest predictor of friendship strength, more than the cooperative

benefits of the relationship, similarity, or other traditional predictors (see also, DeScioli, Kurzban, Koch, & Liben-Nowell, 2011). Specifically, these authors found that people were most likely to rank friends higher if those friends also ranked them higher and use these rankings to decide whose side to take in potential conflicts. Similar to how countries are obligated to take an ally's side over a non-ally in a dispute (Liska, 1962), the alliance account suggests that, all else being equal, people provide support based on how they rank their allies (friends). Because these rankings are necessarily zero-sum (if someone else is ranked higher, then one is ranked lower), the alliance account posits that one's relative standing among other friends is particularly important (a related model by Tooby & Cosmides (1996) based on "irreplaceability" makes similar predictions to the alliance account in the contexts examined here; we return to these two models in the General Discussion).

The importance of friendship rank would prompt people to monitor their place relative to others and respond negatively to the threat of being displaced with some form of friendship jealousy (DeScioli & Kurzban, 2009a, 2011; Shaw, 2013, 2016). There is extensive evidence of friendship jealousy in human friendships (Bevan & Samter, 2004; Kraft & Mayeux, 2016; Nezlek, 1993; Rubin, Bukowski, & Parker, 2006); frantic efforts to avoid losing one's ranking with friends – displacement in friendship ranking, or merely the threat of such, drives people to engage in friend-guarding behaviors when they perceive such threats (Krems, Williams, Kenrick, & Aktipis, 2017). This friendship jealousy appears to be a functional response to information that one may soon be displaced in a friend's ranking and could certainly be triggered by a friend preferentially helping another friend. We know that people infer higher degrees of friendship when they see someone giving someone else special treatment (Kleiman-Weiner, Shaw, & Tenenbaum, 2017; Liberman & Shaw, 2017).

Thus, if the alliance account is correct, it should be possible to find scenarios in which people actually do not like friends who are more prosocial toward others because such prosociality could be taken as a signal that the other person is currently – or soon to be – valued more highly. Such information should make the agent concerned about being displaced in the friend's rankings (DeScioli & Kurzban, 2009a). Of course, this could be consistent with some cooperative accounts based on person-specific generosity (Delton & Robertson, 2016). Such accounts suggest that people's decision-making is guided by an internal regulatory variable the computes the cost a person will pay to give benefits to a specific social partner—also called "welfare-tradeoff ratios" (WTRs). These models might suggest that if an agent's friend helps another person, but not the agent, the agent now knows that their friend will pay cost X for someone else, but not for them. This could easily trigger a negative reaction because it is giving one information that the friend's actual WTRs are lower toward one. Importantly, the alliance account makes a more nuanced prediction that is not made by models based in WTRs; the negativity of the agent's response here should be calibrated to whether the target of the help is someone who is likely to displace the agent (e.g. another friend) rather than someone who fills a very different, yet close role (e.g. a relative or significant other).

In four studies, we investigated the predictions made by cooperative and alliance accounts of friendship and attempt to answer two questions. First, do people respond more negatively to a friend being preferentially helpful (Study 1-2) and generous (Study 3-4) to others (as predicted by the alliance account), or do they respond more negatively to someone who is less helpful and generous overall (as predicted by cooperative accounts)? Second, are these negative reactions particularly strong when the other recipient is a friend rather than someone else (e.g. a relative or romantic partner) as predicted by the alliance account (Study 1, 2, 4)?

Our studies are designed to compare the predictions of the alliance account to different cooperative models of friendship and partner choice. The first two studies test the alliance account against the indirect reciprocity models of cooperation (e.g. Nowak & Sigmund, 1998) which predict that agents should respond more positively to those who are more generous overall (especially if one holds constant negative or positive effects to the agent). Study 3 tests the predictions of the alliance account against models of direct reciprocity (e.g., Trivers, 1971), which predict that people should respond more positively to those who give them more benefits. Finally, Study 4 investigates whether the results from Studies 1-3 can be accounted for by a person-specific model based on WTRs (Delton & Robertson, 2016).

Study 1.1

In Study 1, we investigated people's reaction to a friend helping someone or not. Participants read vignettes in which their friend does not help them and then has a subsequent opportunity to help someone else in similar circumstances. We varied whether the friend helped the other person or not and who else requested help (friend or parent). Cooperative accounts based on indirect reciprocity (e.g., Nowak & Sigmund, 1998) suggest that people track benefits delivered to them and third parties and are positively predisposed toward those who are more helpful. Thus, these accounts predict that, holding constant benefits delivered to the self, people should respond more positively toward someone who helps than someone who does not. Further, these accounts make no specific predictions about whether helping another friend or a parent differs. Conversely, the alliance account predicts that people should feel upset about their friend helping someone else, especially when that person is another friend who could potentially displace them in that friend's friendship rankings.

Method

Participants. Two hundred and two (50% female, *M* = 37.28; *SD* = 12.34) participants completed a 3-minute study for $0.25. Participants in all studies were recruited online through Amazon Mechanical Turk (AMT) and TurkPrime (Litman, Robinson, & Abberbock, 2016). All surveys were presented through Qualtrics. Participation was restricted to AMT workers from the United States with a 95% approval rating or higher as recorded by AMT. Before beginning data collection, we decided not to exclude any participants from any of the studies as the design was fairly simple. In all studies, we attempted to recruit ~50 participants per cell. Each study ended with participants providing basic demographic information.

Design and Procedure. Participants were randomly assigned to one of four conditions in a two-way, between-subjects design in which we varied whether the friend helped (friend's action: helped other, helped neither) and the relationship between one's friend and the other person (relationship: friend or parent). After entering the name of a close friend, participants read a scenario in which they imagined asking to borrow said friend's trailer. In all conditions, their friend (the decision-maker) was unwilling to lend the participant the trailer. Participants were then told that the decision-maker either had (helped other condition) or had not (helped neither condition) agreed to lend the trailer to someone else the following week. We also varied the relationship between the decision-maker and the other person who was either the decision-maker's friend (friend condition) or the decision-maker's dad (parent condition). In the actual study, we used the name participants entered of a real friend, but for simplicity of reading, we will use the name "Pam" and female pronouns throughout the vignettes. The vignette read as follows [within each set of brackets the wording for the helped other condition appears first and

the wording for the helped neither condition appears second, and within the parentheses the wording for the friend condition is followed by the wording for the parent condition]:

> You and Pam are hanging out one evening when you mention to Pam that, this weekend, you want to pick up a dining room table that you'd found online. You know that Pam has a trailer and so you ask if you can borrow her trailer this weekend to pick up the table. Pam pauses, and then tells you that she can't lend you the trailer this weekend because she is trying to sell the trailer in the next couple of weeks and she doesn't want it to get damaged.
>
> The following week, you and Pam are again hanging out when Pam gets a call. After getting off the phone, Pam tells you that (a friend / her Dad) had called asking to borrow her trailer the following weekend to pick up a couch. Pam then tells you that she had [agreed to let her (friend/Dad) borrow the trailer the following weekend / told her (friend/Dad) no because she is trying to sell the trailer].

Participants then rated how upset they felt towards their friend (*1 = Not at all upset, 7 = Extremely upset*).

Results

We ran a 2 (action: helped other, helped neither) x 2 (relationship: friend, parent) ANOVA on how upset participants were at their friend. Although Levene's test for equality of variances was violated, $F(3, 197) = 4.28$, $p = .006$, all other assumptions were met. Please see supplemental materials for a note on this and why this is unlikely to inflate our type 1 errors (Myers & Well, 1995; Richter & Payton, 2003).

The ANOVA revealed a significant main effect of friend's actions, $F(1, 197) = 20.58$, $p < .001$, $\eta^2 = .10$, with participants being more upset when the decision-maker lent the other person the trailer ($M = 4.69$, $SD = 1.57$) than when their friend lent neither person the trailer ($M = 1.82$, $SD = 1.27$). There was also a significant effect of relationship, $F(1, 197) = 197.03$, $p < .001$, $\eta^2 = .50$, with participants being more upset when the recipient was another friend ($M = 3.72$, $SD = 1.44$) than the decision-maker's parent ($M = 2.79$, $SD = 1.42$). Further, there was a significant

friend's action by relationship interaction, $F(1, 197) = 7.88$, $p = .005$, $\eta^2 = .04$, that we followed up on using the Bonferroni correction of $\alpha = 0.008$ for multiple comparisons.

We first examined how upset participants were at the decision-maker when they opted to help the other person and found that participants were more upset when the other person was another friend ($M = 5.44$, $SD = 1.43$) than the decision-maker's parent ($M = 3.94$, $SD = 1.74$), $t(98) = 4.70$, $p < .001$, $d = 0.94$. Further, when the decision-maker helped nobody, there was no significant difference in how upset participants were at the decision-maker regardless of whether the other person was the decision-maker's friend ($M = 2.00$, $SD = 1.46$) or parent ($M = 1.65$, $SD = 1.09$), $t(99) = 1.38$, $p = .171$, $d = 0.27$. See Figure 1.1.

Discussion

After failing to help them, participants were more upset at their friend for helping someone else than for helping no one. These results are not in line with indirect reciprocity cooperative accounts which should predict that people are more upset at the person who failed to help anyone. However, they are consistent with the alliance account in which one's rank is particularly important (DeScioli & Kurzban, 2009a)—if a friend helps someone else after not helping you, this is a sign that you are not as highly valued by that friend as the person who received their help. Importantly, these results cannot be explained simply by the fact that the friend refused to help the participant, because that was constant across conditions. Further, people were more upset when the person helped was another friend rather than the friend's parent. This result is clearly predicted by the alliance account; a friend's friend is more likely to potentially displace one in friendship ranking than a friend's parent. We argue that people feeling more upset when their friend helps another friend is a functional manifestation of this perceived threat of being displaced. Therefore, these results suggest that not only can a friend helping

someone negatively impact a friendship, but also that the relationship between the friend and the person helped influences how upset people get at such help. Thus, these results appear to be best accounted for as reactions of friendship jealousy based on the alliance account.

Study 1.2

We have suggested that people respond more negatively toward preferential help for a friend rather than parent because a parent fills a very different role than a friend. Of course, a parent and friend differ in many ways: one is biologically related to one's parents, does not choose one's parents, and cannot easily leave one's parents. Therefore, in Study 2 we explore a comparison role closer to a friend: a romantic partner. Just like friends, romantic partners (in most cultures) are not related to each other, are chosen, and can be left. However, a friend's romantic partner fills a different role in that friend's life and is unlikely to act as a potential replacement for a friend. Thus, preferential help toward a romantic partner should not be as threatening according to the alliance account.

Here, we again examined how people respond to a friend failing to help them and then subsequently helping or not helping another, and whether the relationship between the friend and other person modulated participants' reaction to their friend helping the other. We compared participants' reactions to their friend preferentially helping or not helping another friend or their romantic partner. We used a new scenario in which the helping behavior involved picking the person up from the airport. Based on the alliance account and our previous results, we predicted that participants would be more upset at their friend preferentially helping another person than helping neither person, especially when that person is another friend as compared to their friend's romantic partner.

Method

Participants. One hundred and ninety-eight participants (50% female, $M = 35.98$; $SD = 11.47$) were recruited in a manner identical to Study 1.1.

Design and Procedure. Participants were randomly assigned to one of four conditions in a two-way, between-subjects design in which we varied who the participant's friend helped (friend's actions: helped other, helped neither) and the relationship between the participant's friend and the other person (relationship: friend, romantic partner). Participants read that their friend (whose actual name appeared in the vignette) denied their request for an airport ride and then did or did not pick up the other person from the airport the following week. Participants read the following [within each set of brackets the wording for the helped other condition appears first and the wording for the helped neither condition appears second, and within the parentheses the wording for the friend condition is followed by the wording for the romantic partner condition]:

> You and Pam are hanging out one evening when you mention to Pam that you have a business trip this weekend and need to arrange transport to the airport. You ask Pam if she can give you a ride. Pam pauses, and then tells you that she can't take you to the airport because she is pretty busy for the next few weekends.
>
> The following week, you and Pam are again hanging out when Pam gets a call. After getting off the phone, Pam tells you that [a friend / her significant other] had called asking to be picked up from the airport the following weekend and that she had [agreed to do so / told her (friend/significant other) no because she is busy].

Participants then rated how upset they felt towards their friend (*1 = Not at all upset, 7 = Extremely upset*).

Results

We ran a 2 (action: helped other, helped neither) x 2 (relationship: friend, romantic partner) ANOVA on how upset participants were at their friend. Although Levene's test for

equality of variances was violated, $F(3, 194) = 3.01$, $p = .032$, all other assumptions were met. The ANOVA revealed a significant effect of friend's actions, $F(1, 194) = 70.63$, $p < .001$, $\eta^2 = .27$, with participants being more upset when the decision-maker helped the other person ($M = 4.14$, $SD = 1.72$) than when their friend helped neither ($M = 2.22$, $SD = 1.48$). There was also a significant main effect of relationship, $F(1, 194) = 8.29$, $p = .004$, $\eta^2 = .04$, with participants being more upset when the recipient was another friend ($M = 3.51$, $SD = 1.62$) than the decision-maker's romantic partner ($M = 2.85$, $SD = 1.58$). Further, there was a marginal friend's action by relationship interaction, $F(1, 194) = 2.85$, $p = .093$, $\eta^2 = .02$, that we followed up on with planned comparisons (based on the fact that this was only marginally significant, we also conducted a direct replication of this study and found a statistically significant interaction; please see supplemental materials, Study S1.1). For these analyses, we used Bonferroni correction of $\alpha = 0.008$ for multiple comparisons.

We first examined how upset participants were with their friend when their friend helped the other person. Participants were more upset when the other person was another friend ($M = 4.67$, $SD = 1.62$) than the friend's partner ($M = 3.62$, $SD = 1.82$), $t(99) = 3.07$, $p = .003$, $d = 0.61$. Further, when the friend helped no one, there was no significant difference in how upset participants were regardless of whether the other person was another friend ($M = 2.35$, $SD = 1.62$) or the friend's romantic partner ($M = 2.08$, $SD = 1.35$), $t(95) = 0.90$, $p = .370$, $d = 0.18$. See Figure 1.1.

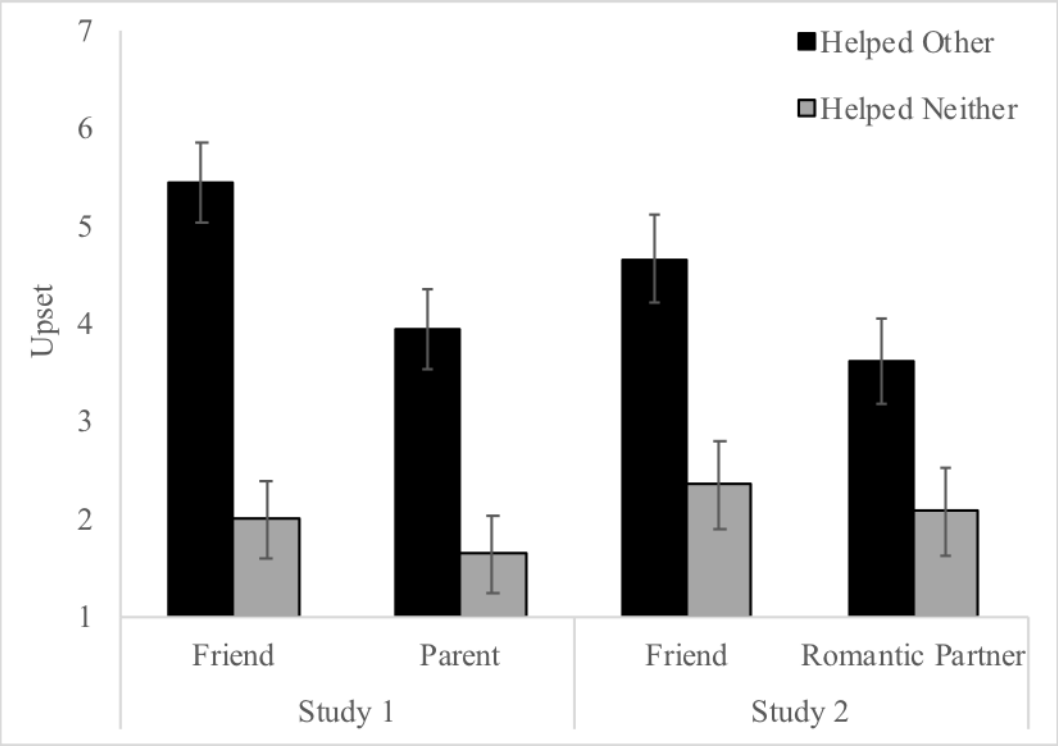

Figure 1.1. Participant reported upset with friend in Studies 1.1 and 1.2 when friend does or does not help either another friend, their parent, or their romantic partner; error bars display 95% confidence intervals.

Discussion

In line with the alliance account, we found that people were upset about their friend helping someone (rather than helping no one) and reacted more negatively when the person helped was another friend than the friend's romantic partner (for a replication of these results along with a comparison between the recipient being a friend, romantic partner, or parent within one study, see supplemental materials, Study S1.2). Along with the results from Study 1.1, these results support the alliance account of friendship and run counter to cooperative accounts based on indirect reciprocity. While this does not mean that people want ungenerous friends (as reviewed above, people do not like those who are overly selfish; also see General Discussion), sometimes helping one friend can lead to negative reactions from other friends.

Study 1.3

So far, we have demonstrated that people are more upset at a friend for not helping them when their friend subsequently helps another friend. We suggest this reveals friendship jealousy in line with the alliance account. However, the most striking evidence in favor of this account and against cooperative accounts of friendship would be to find a case where people respond negatively to a highly generous friend who gives them relatively less than someone else. Our previous results speak against accounts of indirect reciprocity that argue we like people who are broadly more cooperative (for review see, Rand & Nowak, 2013), but do not rule out other cooperative accounts. Models based on direct reciprocity would not make any predictions in the previous studies, but would certainly predict that one should prefer a friend who gives more resources (in absolute terms) to the participant than someone who gives her less. However, the alliance account predicts that one might prefer a friend who is less generous to one in absolute terms (i.e. giving one fewer resources overall) if they prefer one relatively (i.e. giving one a larger share than someone else).

We explore this possibility in Study 1.3 in a scenario that involved a friend distributing raffle tickets. Participants read about a friend who won several raffle tickets and shared them with the participant and another friend. In the less generous condition, the friend distributed fewer tickets (4) overall, but gave relatively more to the participant than the other friend (3 to the participant and 1 to the other friend). In the more generous condition, the friend gave more tickets (10) overall, but gave relatively fewer to the participant than the other friend (4 to the participant and 6 to the other friend). Participants were asked how much they liked their friend (the distributor) and how much their friend liked them. Cooperative accounts based on direct reciprocity predict that people should like the friend better in the more generous condition

because this person gives more overall and more to the participant herself. However, in line with the alliance account, we predicted that participants would like the friend more and predict their friend liked them more in the less generous condition (in which they received relatively more than the other recipient) as compared to the more generous condition (in which they received relatively less). We also added two measure checks, how generous participants rated their friend as (to measure whether the generous friend was indeed more generous) and who their friend felt closer to based on the friend's allocation.

Method

Participants. One hundred and two participants (50% female, $M = 34.01$; $SD = 10.83$) were recruited in a manner identical to previous studies.

Design and Procedure. Participants were randomly assigned to read one of two vignettes. Participants in both conditions read that their friend (this time a hypothetical one named Jamie) had just won raffle tickets and wanted to share them with the participant and another friend. In the less generous condition, the friend gave out fewer tickets but the participant received relatively more tickets than the other friend, while in the more generous condition the friend gave out more tickets overall, but the participant received fewer tickets than the other friend. In both conditions, participants were told: "Imagine you have friend that you recently met named Jamie. Jamie just bought season tickets to your local football team." Participants in the less generous condition read the following:

> As part of the promotion, Jamie got 11 free raffle tickets for an upcoming giveaway. You regularly go to watch the games so Jamie decides to keep 7 of the tickets and gives you 3 of the remaining tickets. When you pick up the tickets, Jamie mentions giving the other ticket to Casey, another of Jamie's friends.
>
> So, Jamie kept 7 tickets, gave you 3 tickets, and gave Casey 1 ticket.

While participants in more generous condition read:

> As part of the promotion, Jamie got 17 free raffle tickets for an upcoming giveaway. You regularly go to watch the games so Jamie decides to keep 7 of the tickets and gives you 4 of the remaining tickets. When you pick up the tickets, Jamie mentions giving the other 6 tickets to Casey, another of Jamie's friends.
>
> So, Jamie kept 7 tickets, gave you 4 tickets, and gave Casey 6 tickets.

Participants then rated their liking for their friend (*1 = Not at all; 7 = A lot*) and closeness to their friend (*1 = Not at all; 7 = Very much so*), which we collapsed into a "Friendship measure" because the two were highly correlated, $r(100) = .68$, $p < .001$. We also measured their friend's predicted liking for them (*1 = Not at all; 7 = A lot*) and included two measure checks. We asked how generous their friend had been (*1 = Not at all generous; 7 = Extremely generous*) and measured the participant's sense of which recipient their friend was closer to (*1 = other friend, Casey; 7 = participant)*. For all analyses, we used a Bonferroni corrected alpha level, $\alpha = .013$.

Results

Friendship measure. An independent samples t-test revealed a significant difference in general friendship; participants liked their friend more when their friend was less generous overall ($M = 5.73$, $SD = 1.12$) than more generous ($M = 5.05$, $SD = 1.19$), $t(100) = 2.95$, $p = .004$, $d = 0.59$. See Figure 1.2.

Friend's Predicted Liking measure. Participants predicted their friend liked them significantly more when their friend was less generous overall ($M = 5.67$, $SD = 1.23$) than more ($M = 4.92$, $SD = 1.28$), $t(100) = 3.00$, $p = .003$, $d = 0.60$.

Generosity check. Participants reported their friend was equally generous whether they received relatively more raffle tickets ($M = 5.73$, $SD = 1.20$) or more overall tickets ($M = 5.67$, $SD = 1.28$), $t(100) = 0.25$, $p = .807$, $d = 0.05$. This result could suggest that participants cannot appreciate the difference in generosity in a between-subjects design or that they see the more generous condition as less generous because they are experiencing jealousy at receiving less. To

differentiate between these accounts, we ran a pre-registered supplemental study (https://aspredicted.org/sj9i3.pdf) in which participants took a third-party role (i.e. they were not asked to imagine being one of the recipients) and were asked to evaluate the generosity of the distributor in one of the two conditions. We found that third parties thought that the friend was more generous in the more generous condition (M = 5.68, SD = 1.16) than the less generous condition (M = 4.45, SD = 1.27), $t(198) = 7.16$, $p < .001$, $d = 1.01$ (see supplemental Study S1.3). This result validates that people not experiencing friendship jealousy can recognize that the more generous person is indeed more generous in a between-participants design.

Relative Closeness check. Participants reported their friend felt closer to them in the less generous condition (M = 6.10, SD = 1.10), than the more generous condition (M = 2.25, SD = 1.31), $t(100) = 16.05$, $p < .001$, $d = 3.18$.

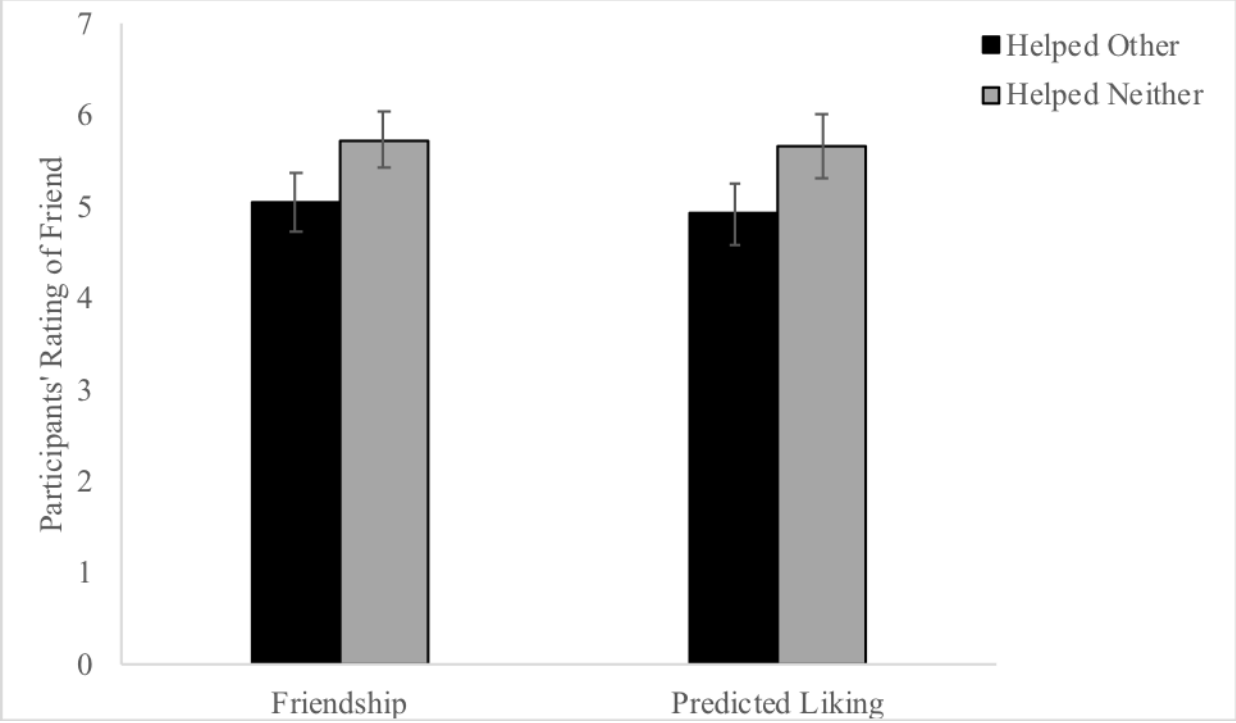

Figure 1.2. Participant rating for friendship and friend's predicted liking in Study 1.3. For the friendship measure, participants are rating how they feel about their friend and the predicted liking measure captures participants' prediction of their friend's liking of them. Error bars display 95% confidence intervals.

Discussion

In line with the alliance account, we found that participants liked a less generous friend who distributed resources in a way that favored the participant, more than a friend who was more generous overall but who distributed resources in a way that favored the other friend (for a direct replication of Study 1.3 with a larger sample size, see supplemental Study S1.5). Note that we obtained this effect in Study 1.3 in a situation where both outcomes were unequal, suggesting that previous negative reactions to preferential treatment of a friend cannot be explained by a simple preference for equality or "fairness" as both outcomes were equally unfair in that they led to similar levels of inequality. These results instead indicate that people's preference for relative advantage in a friendship context can overwhelm their preference for a more generous person, even one who showed more generosity toward them, which is in line with the alliance account of friendship and not cooperative accounts based on direct reciprocity or indirect reciprocity.

Study 1.4

Studies 1.1-1.3 provide support for the notion that people respond particularly negatively to a friend being preferentially prosocial to another friend and argue that this occurs because people are worried about being displaced by another friend, but not a parent or significant other. However, our results thus far could also be accounted for by cooperation theories based on person-specific generosity rooted in WTRs (Delton & Robertson, 2016), which would explain our observed effects in a different way. Such an account might argue that our participants are responding negatively to getting new information about their friend's WTRs with respect to them; namely, that the friend is less willing to pay a cost to benefit them than another friend. Such an account might further argue that people respond less negatively to a friend helping their parent or a romantic partner because they assume that the friend has a higher WTR for their

parent or romantic partner. Therefore, a friend providing preferential help to a parent or romantic partner after failing to help the participant is less indicative of a lack of regard for the participant's welfare. Indeed, there are always ambiguities in how costly a behavior (e.g., picking someone up at the airport or giving away one's raffle tickets) is and therefore people might use their friend's choices and behavior to estimate how costly these behaviors are. In cases where an agent pays cost X to help a very close other (like a parent or romantic partner) but not the participant, this is not particularly informative because the agent likely would pay a very high cost to help these very close others. On the other hand, when an agent pays cost X to help another friend but not the participant, this is now very informative and tells the participant that the agent will pay X cost for another friend, but not to the participant. Therefore, participants may down-regulate their closeness to the agent because they now have information that the agent will pay a certain cost for another friend but not them. According to this account, it is not the relationship type that is driving our previous results, it is the closeness between the friend and the recipients.

In our previous experiments, concerns about a threat of displacement (which would be consistent with the alliance model) and computations about costs based on closeness between donor and recipient (which would be consistent with the model based in WTRs) would lead to similar results. To test between these two plausible models, it is necessary to use someone who is not that close, but also fills a non-friend role—such as a friend's uncle. Here a focus on closeness and potential displacement make different predictions. If people's negative reaction to helping is based on the perceived closeness between the friend and the person helped, then people should respond more negatively to a friend preferentially helping their uncle rather than a close friend because they are probably less close to an uncle than another close friend (an assumption we

confirmed with a measure check). However, if people's negative reaction to preferential helping is based on a fear of displacement, then people should respond less negatively to a friend who preferentially helps their uncle than another close friend.

To test between these accounts, we used the raffle ticket vignette from Study 1.3, but manipulated whether the other recipient was a close friend (a close other) or an uncle (a more distant other) of the participant's friend. Before reading the scenario, participants completed a "closeness check" where they rated how close their friend would be to the other person. After reading one of the two vignettes, participants responded to friendship measures identical to those in Study 3. If people are focused more on closeness between their friend and the recipient, then they should respond more negatively to a friend helping their uncle than another friend because this account assumes that people respond more negatively to a friend giving more to a more distant rather than close other. This could account for our previous findings that people are more upset when someone helps a friend rather than a very close other like a parent or significant other. However, if people are focused more on potential displacement, then they should respond more negatively to a friend giving more to a close friend rather than the friend's uncle because a close friend is more likely to potentially replace one in a friendship rather than an uncle.

Method

Participants. Two hundred and one participants (52% female, $M = 35.60$; $SD = 10.57$) were recruited in an identical procedure to all other studies, but this study was pre-registered (https://aspredicted.org/gp7m8.pdf). Note we ran more participants here because in a pilot study (reported in supplemental materials, see Study S6) we found that the effect size for our liking measure was lower.

Design and Procedure. Participants were randomly assigned to read one of two vignettes about a friend splitting tickets between the participant and another person (another friend or the friend's uncle). Before reading a vignette, participants were asked how close their friend was to the other person (close friend or uncle, *1 = Not at all; 7 = Very much so*). This was included as a closeness check. Participants then read vignettes very similar to those used in the "more generous" condition from Study 1.3. In that condition, the participant receives fewer raffle tickets than the other person. We focused on the "more generous" condition because people are most likely to experience negative reactions in this case. The wording in the friend condition was identical to the "more generous" condition from Study 1.3; the uncle condition was similar except that the friend was replaced with an uncle.

After reading one of the two scenarios, participants answered questions identical to Study 3. Liking and closeness were strongly correlated, $r(199) = .69$, $p < .001$, and combined into a general "friendship" measure. Finally, we used a Bonferroni corrected alpha level, $\alpha = .017$, for all analyses (because we had three relevant comparisons and one measure check that was administered before participants read the target vignette).

Results

Friend's Closeness to Other check. Participants rated their friend as being significantly closer to a close friend ($M = 5.90$, $SD = 1.01$) than an uncle ($M = 4.52$, $SD = 1.11$), $t(195.98) = 9.27$, $p < .001$, $d = 1.32$.

Friendship measure. Participants liked their friend significantly more when the person receiving more was the friend's uncle ($M = 5.81$, $SD = 0.97$) than a close friend ($M = 5.27$, $SD = 1.16$), $t(199) = 3.56$, $p < .001$, $d = 0.51$. See Figure 1.3.

Friend's Predicted Liking measure. Participants predicted their friend liked them significantly more when the person receiving more was the friend's uncle (M = 5.89, SD = 1.04) than close friend (M = 5.16, SD = 1.33), $t(199) = 4.35$, $p < .001$, $d = 0.61$. See Figure 1.3.

Generosity check. Participants reported their friend to be significantly more generous when the person receiving more was the friend's uncle (M = 6.29, SD = 0.90) than a close friend (M = 5.92, SD = 1.13), $t(199) = 2.58$, $p = .011$, $d = 0.36$. This likely reflects their friendship jealousy causing them see the allocation as less generous (see discussion of the generosity check in the results section of Study 1.3).

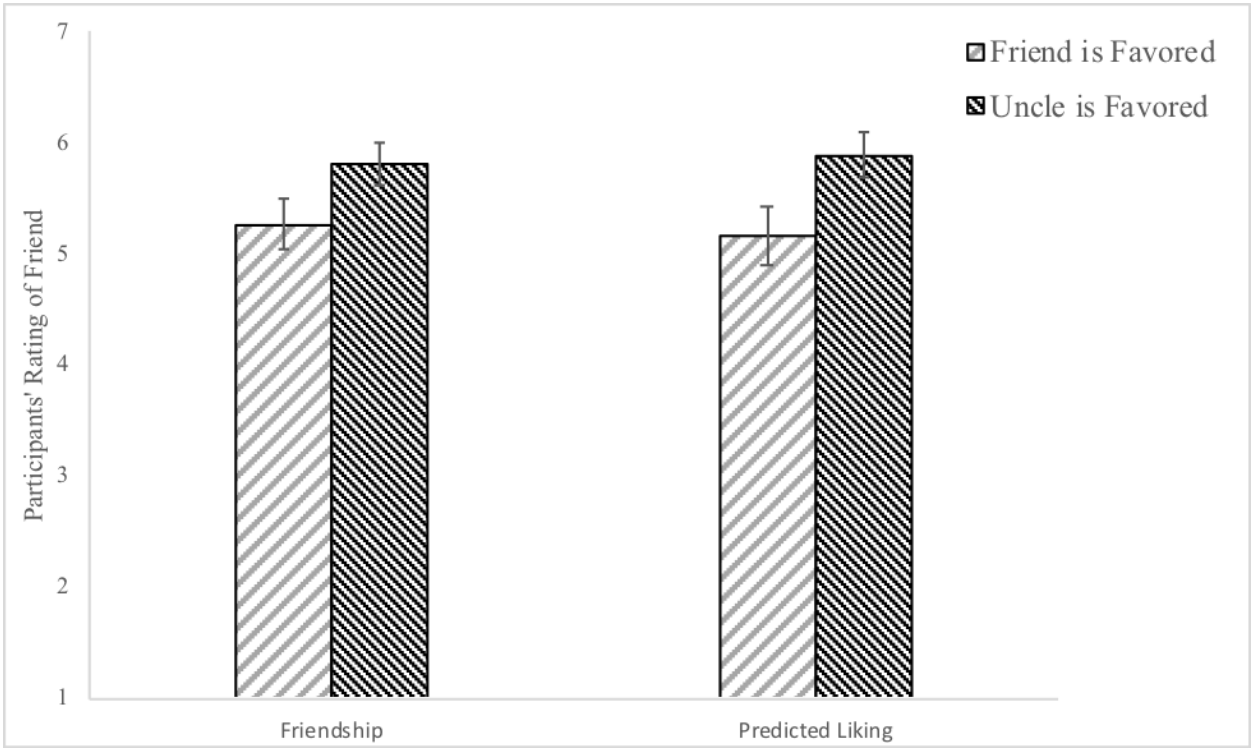

Figure 1.3. Participant rating for friend's friendship and predicted liking in Study 1.4. For the friendship measure, participants are rating how they feel about their friend and the predicted liking measure captures participants' prediction of their friend's liking of them. Error bars display 95% confidence intervals.

Discussion

In line with concerns about displacement driving our results, we found that people were more upset when their friend gave preferentially to a friend rather than an uncle (in two supplemental studies we directly replicate this result and also replicate this effect using our airport vignettes from Study 1.2 with the target being either an aunt or a friend, see supplemental Studies 6a and 6b, respectively). The focus on potential displacement suggests that people should be more upset at a friend's preferential treatment toward another friend because that friend can potentially displace one in one's friend's alliance ranking. These results are not consistent with people focusing on the closeness between the recipient and the decision-maker: People thought their friend was less close to their uncle than a close friend but were still less upset when their friend gave more to their uncle than a friend. Of course, the closeness between who is helped will obviously matter in other contexts (Cosmides & Tooby, 2013; Delton & Robertson, 2016; Sell, 2011). Indeed, if a friend failed to help the participant and then subsequently helped a distant friend rather than a close friend, we imagine that a friend helping the former would be much more upsetting. However, we found that this closeness variable alone could not account for the present results, suggesting that there are concerns beyond closeness driving participants' reactions. These results are consistent with people's negative response to preferential prosociality being based in a potential fear of displacement by another friend.

Nevertheless, it is important to note that while we used closeness as an approximation for WTRs, this single variable cannot capture people's complete WTRs (for discussion, see Delton & Robertson, 2016). While participants agreed that their friend was closer to another friend than to their uncle and were more upset when their friend helped a friend than an uncle, this might be because participants think their friend has higher expected WTRs toward an uncle, if people

expect higher WTRs toward all kin. Indeed, the relationship between closeness and expected WTRs can break down in family relationships; it's certainly possible for one to not be particularly emotionally close to a sibling and yet still be willing to sacrifice a great deal for them (Hackman, Danvers, & Hruschka, 2015; Stewart-Williams, 2007). Therefore, the current results could be incorporated into a person-specific account, but such an account would have to include one's relationship type (kin versus non-kin) as a possible variable and specify if and how interpersonal closeness relates to expected WTRs. Future work should more carefully disentangle these explanations and we return to these issues in the general discussion.

General Discussion

In a series of studies, we find support for the alliance hypothesis; people responded negatively to a friend who preferentially helps or is generous toward someone else, particularly when that other person was another friend. We find that people respond much less negatively to a friend helping a relative or romantic partner (Studies 1.1, 1.2, and 1.4) because a friend is less worried about being displaced by a relative or romantic partner. We argue and provide evidence that these results are better accounted for by a fear of displacement based in friendship jealousy rather than other accounts based on cooperation theories relating to indirect reciprocity (Studies 1.1 and 1.2), direct reciprocity (Study 1.3), and person-specific cooperation based on WTRs (Study 1.4).

These experiments build on previous work arguing that cooperative accounts cannot explain many features of close friendships (for review see, DeScioli & Kurzban, 2011). Reciprocity is clearly a fundamental and pervasive strategy that explains many aspects of social life including trade, bargaining, markets, and cooperative production. However, there are many features of friendship that these models do not predict or explain. First, the clear friendship

jealousy and negative reaction to cooperation that we found here does not fit with cooperative accounts of friendship. Second, the fact that friends do not keep careful accounting of each other's contributions, and moreover, view doing so as offensive or as a signal of a more formal relationship (Clark & Mills, 1979; Fiske, 1992; Xue & Silk, 2012) contradicts cooperative theories which rely on accounting. At the very least, cooperative accounts appear to provide an incomplete picture of people's friendship choices.

It is worth noting that our experimental results and arguments are consistent with both the alliance account and a related account proposed by Tooby and Cosmides (1996). Their account similarly argues that friendships are not focused on cooperation, but instead on irreplaceability—people seek friends who uniquely value them, see them as irreplaceable, and engage in behaviors to demonstrate their irreplaceability to others. This account makes similar predictions about people reacting negatively to the threat of displacement. The current experiments were not designed to differentiate between these two accounts and were instead designed to provide evidence in favor of both accounts and against cooperative accounts of friendship. We hope that future work can more carefully differentiate these accounts (for discussion, see Hruschka, Hackman, & Macfarlan, 2015).

Importantly, we do not deny that people like friends who are cooperative, and when trying to find a good friend, an agent is likely seeking at least two broad qualities: a friend who is not selfish, i.e., unwilling to trade off their own welfare for others (Delton & Robertson, 2016), and values the agent particularly highly (DeScioli & Kurzban, 2009a). The first concern is what many of the cooperative accounts have focused on: it is true that someone who is overly selfish is unlikely to be a good friend. If a friend is too selfish to offer help or aid, their usefulness as a friend will be reduced even if they rank one very highly. The second concern is related to the

alliance account: if someone is very generous generally but ranks one lowly, then they will be less willing to take one's side in potential conflicts and this too will reduce their usefulness as a friend. In some contexts (e.g., where having someone's support is particularly important), we might value good alliance partners who are somewhat selfish, whereas in other contexts (e.g., where receiving resources or help is important), we might prioritize friends who are better cooperative partners. Of course, past work has explored how people balance desires to be cooperative and fair with desires to favor one's allies (DeScioli & Kurzban, 2009b, 2013; Shaw, DeScioli, Barakzai, & Kurzban, 2017; Shaw & Knobe, 2013; Waytz, Dungan, & Young, 2013) and the complex calculus that underlies how people negotiate the tradeoffs between their own welfare and others' welfare in interpersonal relationships (Cosmides & Tooby, 2013; Delton & Krasnow, 2017; Delton & Robertson, 2016; Krems, Kenrick, & Neel, 2017; Pietraszewski, 2016; Sell, 2011). However, there is still considerable work to be done.

It could be argued that social comparison between the self and the potential recipient of help are influencing our findings. Previous research has demonstrated that people experience stronger negative social comparison when the person who receives better outcomes is more similar to them (Festinger, 1954). Certainly, most of our participants would have imagined that their friend's friend was more like them than their friend's father or uncle (who are likely older) or romantic partner (who was likely a different sex), which might have led them to be more upset for similarly based social comparison reasons, rather than friendship jealousy. In a supplemental study, we manipulated whether the helped recipient was similar to the participant or not and found no impact of similarity to the person helped (see supplemental Study S1.7).

If people do respond negatively to such prosociality, one open question is how people avoid these negative reactions in their day-to-day interactions. DeScioli and Kurzban (2009a)

suggest that people commonly engage in behaviors to conceal how they rank their friends in what has been called "friendship masking". While sometimes people cannot avoid making such rankings apparent (e.g., picking a best man or maid of honor), they often feel uncomfortable publicly ranking their friends and displaying behaviors that make it clear where friends are ranked relative to one another (DeScioli & Kurzban, 2011). People likely avoid friendship jealousy by attempting to conceal their preferential help of others and future research should investigate whether and how people attempt to mask their friendship rankings to avoid friendship jealousy and accusations of unfairness (Choshen-Hillel, Shaw, & Caruso, 2015; Shaw, Choshen-Hillel, & Caruso, 2016). Finally, here we focused on people's negative reaction to being displaced in friendships, but it is likely that people respond negatively to a threat of displacement in any significant relationship: romantic partners, kith and kin. Although relationships with kin likely have their own dynamics (e.g., Cosmides & Tooby, 2013; Hackmen et al., 2015), it seems likely that kin too would worry about being displaced as, for example, the favorite uncle or brother. The current work clarifies that these threats of displacement based on preferential treatment are most acute when they occur within rather than between relationship types. Indeed, we found that one's friend may not be concerned with their friend helping an uncle, but an uncle may be concerned if their nephew preferentially helps another uncle. Such a comparison could also help differentiate between the displacement and WTR-based accounts we discussed in Study 4. Such WTR accounts might attempt to accommodate the findings from Study 4 by suggesting that people were more upset about their friend helping another friend than an uncle because the expected WTRs toward the uncle might be higher. If this is so, then this account should predict that one is less jealous when one's nephew preferentially helps another uncle rather than a friend (because the former has a higher expected WTR). Of course, the displacement account would

predict the opposite because one uncle is more likely to displace one and so that should be more threatening. Future work should investigate when and how these concerns with being displaced manifest in varied interpersonal relationships.

Our experiments demonstrate that people respond negatively to the prosociality of their friends toward others, particularly when this prosociality is towards other friends. We have argued that this negative reaction is a functional response to potential displacement in friendship ranking. These studies highlight an important trait that we look for in friends, that goes beyond mere generosity.

CHAPTER 2: PICK THE RIGHT SIDE: PEOPLE INFER ALLIANCES AND MORALITY BASED ON THE SIDE-TAKING DECISIONS OF THEIR FRIENDS.

In Chapter 1, I have explored how people respond to their friends' prosociality and how, from those prosocial acts, they infer and update the friendship and their value to their friend. It appears that people make a variety of inferences about their friendships and how their friends value them from their friends' behaviors. Another aspect where people may make such inferences may be in response to how their friends' take sides when a conflict arises. Previous research suggests that people do make friendship inferences when their friend sides with them, against them, and remains neutral in response to one's conflict with another person. Specifically, people feel negative towards a friend failing to take their side (i.e. siding against one or remaining neutral) and treat neutrality from a friend as similarly negatively as outright opposition (Shaw et al., 2017).

Therefore, in Part 2, I am interested in seeing if people's inferences about their friendships extend to inferences about their friend in such conflict scenarios. Specifically, when one is in the right or wrong of a conflict, and one's friend chooses a side, does one also make inferences about one's friend's morals? For example, imagine that you just found out that your friend was caught stealing from work. When confronted with situations like these, in which we may be stuck needing to choose between loyalty to our friend on one hand and morality on the other, having friends can put us in quite the dilemma. Being loyal to our friend can mean ignoring our moral principles and upholding our moral principles can mean being disloyal to a friend (DeScioli & Kurzban, 2013; Shaw, 2013; Waytz, Dungan, & Young, 2013). It makes sense that our friends want us to take their side in times of strife (Tooby & Cosmides, 1996) and that they may look at our failure to do so as an indication that we did not value them as highly as

they expected. In situations like these, are we in an inextricable dilemma in which we must choose between sacrificing our friendship or our morality?

To answer this question, we will examine how people involved in a moral conflict actually feel towards their friends who are put in such dilemmas and we will argue that the situation is not so dire. We will suggest that although people do expect loyalty from their friends, they also understand that a person's moral commitments might make even a loyal friend hesitate at taking their side. Thus, although people may expect support from their friends all else being equal, they may be substantially less upset at a friend's failure to take their side when they have committed a clear moral transgression. Here we explore such scenarios. Specifically, we examine how people feel about the side-taking decision of their friends in situations in which they are in the moral right versus wrong. However, before getting to these experiments, we briefly review some previous work on friendships and morality.

Background.

Friends expect loyalty and thus expect preferential treatment from their friends. Being a good friend means helping friends in times of need (Beck & Clark, 1988; Tooby & Cosmides, 1996) and taking their side in conflicts (DeScioli & Kurzban, 2011). Indeed, according to the alliance hypothesis of friendship, one particularly crucial benefit that friends provide is support when conflicts arise (e.g., DeScioli & Kurzban, 2011). In the same way that countries are obligated to take an ally's side over a non-ally in a dispute (Liska, 1962), the alliance hypothesis suggests that people rank their friends (allies) and that these ranks are an important determinant of friendship. In line with this hypothesis, DeScioli and Kurzban (2009a) found that one of the best predictors of friendship strength (how close people feel to their friend) is friendship rank, which predicted friendship strength more accurately than other important predictors of friendship

like proximity, similarity, and cues that one is a good exchange partner (for a replication in a large-scale online social network, see DeScioli, Kurzban, Koch, & Liben-Nowell, 2011). People clearly rank their friends and feel closer to those who they rank, and who rank them, more highly.

Importantly, the alliance hypothesis also predicts that people should feel negative toward a close friend who does not favor them over other, presumably lower-ranked friends. In line with this notion, Shaw, DeScioli, Barakzai, & Kurzban (2017) found that people feel negative toward a friend who fails to take their side in a conflict, particularly when the conflict is with a less close ally, i.e., an acquaintance. One might think that a defense against such accusations of disloyalty is to simply sit out of these conflicts, but this research also demonstrated that neutrality from a friend is sometimes quite costly to a friendship and is treated as a mild form of opposition (Shaw et al., 2017). The reason one may feel negative toward a friend in this circumstance is that the friend's failure to provide support may be diagnostic of that friend's general disloyalty or that one is not ranked particularly highly by that friend. However, if alliances were the only strategy by which humans could show support or opposition in a conflict, then people would frequently have to choose between fomenting a costly conflict or upsetting a friend and losing a valuable friendship.

However, as DeScioli and Kurzban (2013) note, people actually take at least two broad sets of considerations into account when making inferences about others: the loyalty considerations highlighted through alliance theory and moral considerations. Because people also often choose sides based on moral rules, it may be possible to oppose a friend when they are morally in the wrong without upsetting that friend. In many circumstances, people condemn and oppose others (even friends) if they violate moral rules and conventions (DeScioli & Kurzban,

2009b; Dungan, Waytz, & Young, 2014; Morgan, 2007; Waytz et al., 2013). Unlike loyalty, which is based on obligations to individuals and relationship strength, moral rules can be applied with a degree of impartiality such that certain actions are wrong regardless of personal ties (DeScioli & Kurzban, 2009b; Kant, 1785; Kohlberg, 1981; Kurzban & DeScioli, 2015). Research suggests that people value impartiality in others (Chen, Chen & Xin, 2004; Choshen-Hillel, Shaw, & Caruso, 2015; Lind, Tyler, & Huo, 1997; Tyler & Blader, 2007; Tyler & Lind, 2002). Thus, morality can provide a counterweight to loyalty: in many situations, people feel the pull of both interpersonal obligations and moral obligations (Everett, Faber, Savulescu, & Crockett, 2018; Graham et al., 2013; Hughes, 2017; Dungan, Waytz, & Young, 2015; Niemi & Young, 2017; Shaw, Choshen-Hillel, & Caruso, 2018; Waytz et al, 2013).

Here we explore two key predictions that follow from the notion that in a conflict, people can take sides based on both loyalty and morality—what we will refer to as the moral side-taking hypothesis (DeScioli, 2017). If one recognizes that one's friend is pulled by both loyalty and morality considerations as suggested by the moral side-taking hypothesis, then two key predictions follow.

The first key prediction of this hypothesis is that people should feel less negative about their friend's failure to support them when they have committed a moral transgression than when they are the victim of a transgression. This prediction is based on the notion that a friend's failure to support one when one is in the wrong should be less diagnostic of that friend's loyalty. As mentioned above, even a good ally may falter if the transgression clashes with their moral commitments. Hence, the friend could still be a valuable ally in future conflicts, even if they are a little less valuable in cases when the person takes wrongful actions. Thus, the moral side-taking hypothesis predicts that people will feel much less negative about a friend's lack of support when

they are clearly in the moral wrong in that situation. However, it could be that people in fact feel more negative about their friend's failure to support them when they are in the wrong as compared to in the right. Indeed, it is likely that, when a friend has committed a moral transgression, they feel especially vulnerable to backlash from others and therefore may wish for their close friends to provide social support. As mentioned above, people expect loyalty from their close friends and expect support in times of need (Beck & Clark, 1988; Tooby & Cosmides, 1996). Therefore, failing to support a friend when they are facing condemnation from others may be worse than failing to show support when a friend has the moral advantage and, hence, may have a negative impact on the friendship.

A second key prediction is that people will make inferences about a person's moral commitments based on their side-taking decision. Indeed, the moral side-taking hypothesis argues that it is critical to be on the right side of moral conflicts to avoid being condemned by others. In line with this, previous work has found that people make inferences about an agent's moral stances based on what they condemn and punish (Barclay, 2006; DeScioli & Kurzban, 2013; Fessler & Haley, 2003; Kennedy & Schweitzer, 2015). Indeed, recent research found that if someone condemned a certain behavior (e.g., using performance-enhancing drugs), participants inferred that they themselves would not use such drugs (Jordan, Sommers, Bloom, & Rand, 2017). This previous research explored broader inferences and, as such, provides strong support for the second prediction of the moral side-taking hypothesis, in relation to one' friends: if a friend opposes a moral transgressor, people may think that the friend opposes that behavior and unlikely to engage in that immoral behavior. And, if a friend supports a moral transgressor, people will make the opposite inference—that the friend supports or might even engage in that behavior themselves.

Relatedly, the moral side-taking hypothesis predicts that people might make interesting inferences about a side-taker's neutrality based on the interaction between their loyalty and moral-based commitments. Neutrality need not be taken as neutral (e.g. Shaw et al., 2017); it can be seen as being tacitly in favor of or against a transgressor's action. One factor that could drastically impact how people interpret one's neutrality is one's pre-existing alliance commitments. If one has no pre-existing alliance commitments, i.e., is not close with either disputant, neutrality when faced with a clear moral transgression may be taken as condoning the transgression – others may even infer that the neutral party would engage in that transgression themselves. Why else would the bystander be reluctant to condemn and oppose something immoral? On the other hand, if the bystander's friend (an obvious pre-existing alliance) commits a moral transgression, people may be less likely to interpret neutrality as an endorsement of that specific moral transgression. Here, the pre-existing alliance provides another, possibly more reasonable, inference about the bystander – that they are just trying to be a good friend. Even if they are opposed to the moral transgression, their loyalty to their friend might make them more reluctant to decry a friend's actions.

To test these predictions, we report five studies in which we examine how people react to their friend's side-taking decisions in hypothetical scenarios. Specifically, we present participants with situations in which the protagonist (the participant) is involved in a dispute over a moral transgression and their friend (the side-taker) must choose which side to support. We vary whether the protagonist is the perpetrator or the victim of the transgression and what decision the side-taker makes: siding with the protagonist, against the protagonist, or remaining neutral. We measure people's inferences about the friendship (Studies 2.1-2.5) and their friend's moral beliefs and behaviors (Studies 2.2-2.5). In line with the moral side-taking hypothesis, we

predict that when people are in the wrong in a conflict, lack of support from their friend will be less damaging to the friendship than when they are in the right (Prediction 1 outlined above). Further, we expect that people will make inferences about their friend's moral beliefs and behaviors based on that friend's siding response (Prediction 2 outlined above). We were also interested in how people perceive neutrality with a desire to de-escalate the conflict (Study 2.6). This study was in relation to previous work (Shaw et al., 2017) and we were curious whether neutrality with a desire to de-escalate the conflict would be perceived more positively than simple neutrality, which previous work has demonstrated as being perceived as similarly negative as outright opposition. Promoting peace is often interpreted as a moral action (Johnson, 2017; Lederach, 2005; Opotow, Gerson, & Woodside, 2005), and peace-making involves attempts to reduce the likelihood of a conflict arising. Given the morality of peace-making, it is likely that people will perceive a friend's attempts to de-escalate a conflict as a positive action. Therefore, we predicted that neutrality with an attempt to de-escalate the conflict would be perceived more positively than simple neutrality.

Study 2.1

In Study 2.1, we investigated how participants feel toward a friend who either takes sides or remains neutral in a conflict, while also varying whether the participant instigated the conflict. We asked participants to imagine they are at a bar with a close friend when they meet a mutual acquaintance. After some time, an argument breaks out between the participant and the acquaintance. Between conditions, we vary two factors: whether the participant's friend remains neutral (remained neutral condition), sides against the participant (sided-against condition), or sides with the participant (sided-with condition). And the participant's moral standing in the argument: the participant was either in the wrong because they instigated the argument (wrong

condition) or they were in the right because the acquaintance started it (right condition). After reading the vignette, participants answered questions about how they would feel toward their friend.

Based on previous research (DeScioli & Kurzban, 2009a; Shaw et al., 2017), we expected participants to feel more negative toward a friend who failed to take their side (remained neutral or sided against them) compared to a friend who sided with them. However, we were most interested in whether the participant's moral standing would influence how they feel about their friend's side-taking decisions. As we noted above, the side-taking hypothesis predicts that a person will feel less negative toward a friend who fails to take their side (whether by remaining neutral or siding against them) when the person is morally in the wrong as compared to morally in the right. The side-taking hypothesis makes this prediction because in such circumstances failure to a take a friend's side is less diagnostic of the friend's allegiance: even if the friend is generally loyal, they may refuse to provide support based on their moral commitments.

Method

*Participants***.** For all studies, participants were recruited online using the Amazon Mechanical Turk (AMT) website and the TurkPrime.com interface (Litman, Robinson, & Abberbock, 2016). Participation was restricted to residents of the United States and subjects were compensated $0.25 each. For each study, we aimed to recruit ~50 participants per cell and did not analyze the data for any study until data collection was completed. No participants were excluded across studies due to the simple nature of the tasks. All measures and manipulations are included in the manuscript. We had no specific hypotheses regarding sex differences and there were no significant correlations between gender and our dependent variables, so we collapse

across gender in all of our studies. Further, we did not collect demographic information beyond age and sex of the participant as we did not make predictions about individual differences.

In Study 2.1, 299 participants (170 females, $M = 33.21$, $SD = 11.68$) were assigned to one of six conditions: when the participant was in the wrong and the friend remained neutral (n = 49), sided against the participant (n = 51), or sided with the participant (n = 51); when the participant was in the right and the friend remained neutral (n = 48), sided against the participant (n = 49), or sided with the participant (n = 51).

Procedure. In a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) design, participants were randomly assigned to read one of six versions of a vignette. The vignette was about three people at a bar: the participant (the protagonist), their close friend (the side-taker), and an acquaintance of both the participant and their friend (the opponent). Participants imagined that they got into a verbal argument with the acquaintance. Across conditions, we varied participants' moral standing in the conflict – they were either in the wrong or in the right. Participants in the wrong condition read the following:

> Imagine that you and your close friend Jamie are out at a bar. The two of you start talking to Casey, a new person you and Jamie recently met.
>
> After sitting in the bar for an hour, you pick a fight with Casey. Unprovoked, you argue with Casey and proceed to insult, yell, and scream at Casey. Casey tries to make peace, but you continue to curse at Casey.
>
> Finally, Casey says, "What's your problem? Why are you being such a jerk?" You say, "Me? You're the one being a jerk." Then Casey looks at Jamie, "Who's the jerk?"

Participants in the right condition read:

> Imagine that you and your close friend Jamie are out at a bar. The two of you start talking to Casey, a new person you and Jamie recently met.
>
> After sitting in the bar for an hour, Casey picks a fight with you. Unprovoked,

> Casey argues with you and proceeds to insult, yell, and scream at you. You try to make peace, but Casey continues to curse at you.
>
> Finally, Casey says, "What's your problem? Why are you being such a jerk?" You say, "Me? You're the one being a jerk." Then Casey looks at Jamie, "Who's the jerk?"

We also manipulated the friend's side-taking decision by varying Jamie's response. In the remained neutral condition, Jamie said, "I'm not getting involved, guys." In the sided-against condition, it was, "You were being the jerk and so you should apologize." In the sided-with condition, Jamie said, "Casey, you were being the jerk and so you should apologize."

After reading the vignette, participants answered four questions. The first question asked how close the participant felt to Jamie (the friend and side-taker). Specifically, participants rated whether they would feel more or less close to Jamie given what had happened, using a scale ranging from "*a lot less close*" (coded as -3) to a "*lot more close*" (coded as 3), with "*neither less close nor more close*" being the midpoint (coded as 0). In the second question, participants answered whether their relationship with Jamie had been damaged or strengthened on a scale from "*damaged a lot*" (coded as a -3) to "*strengthened a lot*" (coded as a 3), with "*neither damaged nor strengthened the friendship*" (coded as a 0) as the midpoint. These two measures were highly correlated, $r(299) = 0.81$, $p < .001$, and therefore combined into a single measure of feelings toward the friend ranging from negative to positive.

Third, participants rated the likelihood that they would side with the friend (Jamie) if the friend needed support in a future conflict; ratings were on a scale from "*not at all*" (coded as 0) to "*extremely*" (coded as 100). We used 0 to 100 because we wanted participants to think of this as a likelihood measure, which more naturally fits on a 0 to

100 scale (similar to percentages, importantly in previous work with similar vignettes we found the same results with a 0 to 100 scale as with a 1-7 scale, Shaw et al., 2017). Finally, as a simple manipulation check, participants rated perceived blame by answering who they thought was in the wrong in the argument on a scale from "*Casey*" (coded as 1) to "*You*" (coded as 7) with "*neither you nor Casey*" (coded as 4) as the midpoint.

Results

Friendship Measure. First, we looked at the overall pattern of means with a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' feelings toward the friend. We found a main effect of moral standing, $F(1, 293) = 12.76$, $p < .001$, $\eta^2 = .04$, where participants generally felt less negative toward their friend when the participant was in the wrong than in the right. We also found a main effect of side-taking decision, $F(2, 293) = 126.19$, $p < .001$, $\eta^2 = .46$; participants felt more negative when the friend sided against them ($M = -1.14$, $SD = 1.18$) than when the friend remained neutral ($M = -0.80$, $SD = 1.04$), $t(195) = 2.09$, $p = .038$, or sided with them, ($M = 1.12$, $SD = 1.34$), $t(200) = 12.70$, $p < .001$, and they felt more negative when the friend remained neutral than sided with them, $t(197) = 11.28$, $p < .001$. Finally, there was also a significant moral standing by decision interaction, $F(2, 293) = 23.85$, $p < .001$, $\eta^2 = .14$, which we followed up with planned comparisons.

Our main interest is how participants felt toward their friend depending on whether the participant was in the wrong or the right for each of the friend's side-taking decisions. When the friend remained neutral, participants felt more negative towards their friend when the participant was in the right ($M = -1.27$, $SD = 0.89$) than in the wrong ($M = -0.35$, $SD = 0.97$), $t(95) = 4.87$, $p < .001$. Similarly, when the friend sided against them, participants felt more negative toward their friend when the participant was in the right ($M = -1.75$, $SD = 1.11$) than in the wrong ($M = -$

0.55, *SD* = 0.92), *t*(98) = 5.88, *p* < .001. Consistent with the side-taking hypothesis, these data suggest that people feel less upset with their friend for failing to take their side when they are in the wrong than when they are in the right. When the friend sided with them, participants felt more negative when they were in the wrong (*M* = 0.74, *SD* = 1.37) than right (*M* = 1.50, *SD* = 1.20), *t*(100) = 3.00, *p* = .003. See Figure 2.1.

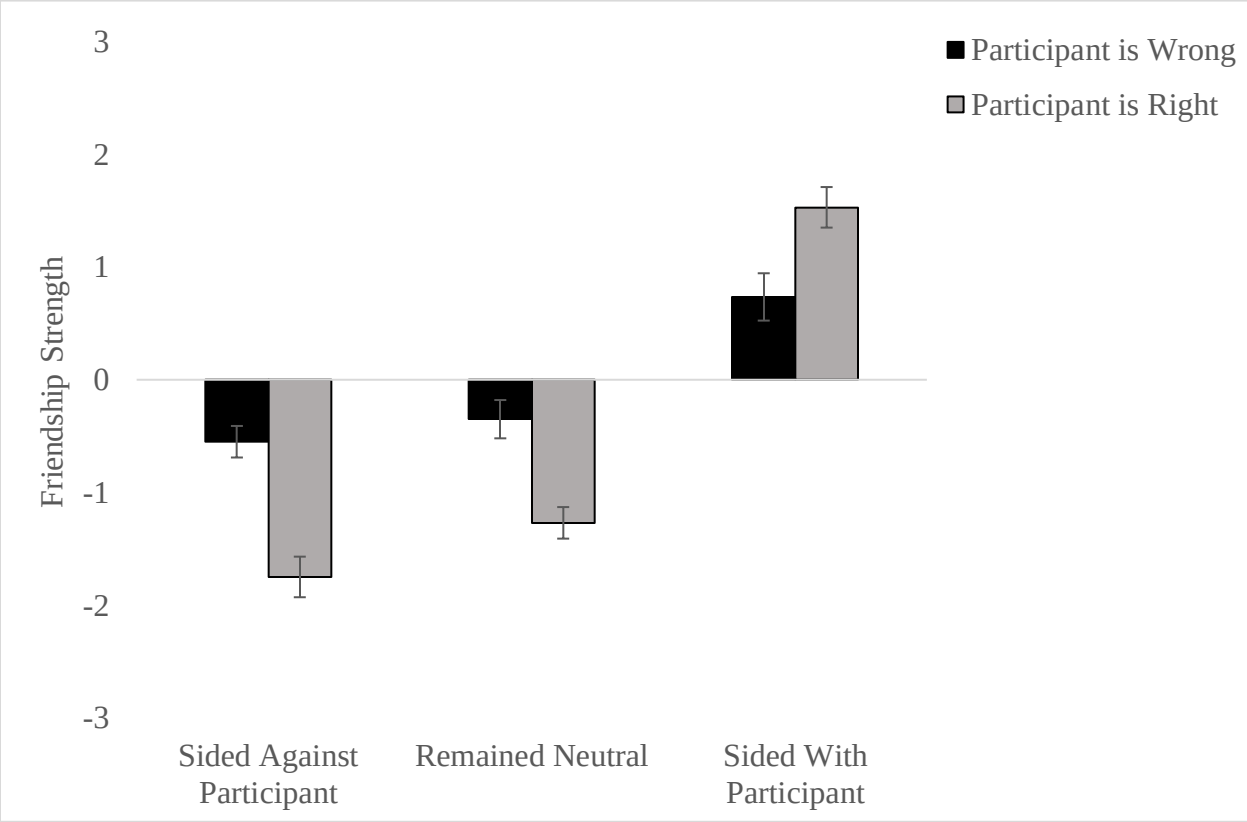

Figure 2.1. Participants' reported friendship strength with the side-taker by the participant's moral standing and the friend's side-taking decision (Study 2.1).

Future Support. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided against, sided with) ANOVA on participants' judgment of their future support for their friend revealed a significant effect of moral standing, $F(1, 293) = 33.07$, $p < .001$, $\eta^2 = .10$, such that participants reported a lower likelihood of siding with their friend in the future when the

participant was in the right than when the participant was in the wrong. The ANOVA also revealed a significant effect of decision, $F(2, 293) = 64.00$, $p < .001$, $\eta^2 = .30$, with participants reporting a lower likelihood of siding with their friend in the future when the friend sided against them ($M = 43.83$, $SD = 23.93$) than when their friend remained neutral ($M = 51.67$, $SD = 23.29$), $t(195) = 2.33$, $p = .021$, or sided with them ($M = 75.81$, $SD = 21.20$), $t(195) = 9.61$, $p < .001$. Further, participants were less likely to provide future support to their friend when their friend remained neutral rather than siding with the participant, $t(195) = 7.52$, $p < .001$. Finally, the analysis also revealed a significant moral standing by decision interaction, $F(2, 293) = 10.42$, $p < .001$, $\eta^2 = .07$, which we followed up on with a series of planned contrasts.

Again, we were primarily interested in comparing participants' response to their friend's side-taking decision in circumstances where participants were in the right vs. in the wrong. When the friend remained neutral, participants reported a lower likelihood of providing the friend future support when the participant had been in the right ($M = 42.79$, $SD = 22.97$) than in the wrong ($M = 60.35$, $SD = 20.30$), $t(95) = 3.99$, $p < .001$. When the friend sided against the participant, participants reported a lower likelihood of providing future support to the friend when the participant had been in the right ($M = 30.84$, $SD = 20.52$) as compared to the wrong ($M = 56.31$, $SD = 20.13$) , $t(98) = 6.27$, $p < .001$. Again, these results suggest that people are more accepting of a friend's lack of support during a conflict when they are responsible for said conflict. Finally, when the friend sided with them, participants were just as likely to provide future support to their friend whether they were in the wrong ($M = 75.35$, $SD = 21.30$) or right ($M = 76.27$, $SD = 21.30$) $t(100) = 0.22$, $p = .827$. All other pairwise comparisons not central to our hypotheses have been reported in Table 1.

Perceived Blame. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided against, sided with) ANOVA on our manipulation check of perceived blame revealed a significant effect of moral standing, $F(1, 293) = 729.95$, $p < .001$, $\eta^2 = .71$; participants in the wrong perceived themselves to blame in the conflict while those in the right perceived the blame fell on their opponent. There was also a marginal effect of decision, $F(2, 293) = 2.73$, $p = .095$, $\eta^2 = .02$, and no moral standing by side-taking decision interaction.

Discussion

The results from Study 2.1 supported the predictions of the side-taking hypothesis: people felt more negative about a friend's failure to take their side when they were in the right as compared to when they were in the wrong. Further, in line with past research, we found that participants felt more negative toward a friend who remained neutral or sided against them than one who sided with them regardless of which moral standing condition they were assigned to. Importantly, the magnitude of this negative response to a friend failing to take their side was substantially reduced when the participant was in the wrong. Taken together, these results suggest that the side-taking decisions of friends influence how people feel about the friendship, but that their negative feeling toward their friend's failure to take their side is tempered by whether the friend had a moral reason to not take their side.

Study 2.2

Study 1 found that participants were more accepting of a friend's failure to support them when they were in the wrong, compared to when they were in the right—that is, when their friend's failure to side with them could be attributed to concerns about the participant's moral behavior. In Study 2, we attempted to replicate our results by asking participants how negative or positive they would feel about their friend's side-taking decision in a scenario involving a

different interpersonal conflict (one in which someone breaks a promise) and to further explore the inferences that people make about the friendship and their friend's morality. To this second end, we added an additional measure to assess participants' inferences about their friend's moral beliefs.

In line with the side-taking hypothesis about alliances and morality that we outlined above (DeScioli & Kurzban, 2013), we hypothesized that a side-taker's decision about whom to support would lead people to infer where she stood on a moral issue. Specifically, the side-taking hypothesis predicts that siding with a promise-breaker should lead people to think that promises are not very important to the side-taker, while siding against the promise-breaker would lead people to think that promises are important to the side-taker. That is, regardless of whether the participant (the protagonist) or her opponent broke the promise, participants would make some inference about the side-taker's morality based on whether the side-taker sided with someone who was in the wrong (broke their promise) or right (someone broke a promise to them).

Further, as we outlined in the introduction, participants will likely make different inferences about the side-taker's morality when they remain neutral depending on whether their friend (the participant) was in the right or in the wrong. Specifically, the side-taking account predicts that participants will be more likely to make a negative inference about the side-taker's morality in situations where the participant was in the right than in the wrong (i.e. participants would think that promise-keeping was not important to the friend). Indeed, if one's friend refuses to take one's side when one is in the right, one might think that their friend has other moral qualms influencing their siding decision.

Based on the side-taking hypothesis, we thus made two predictions. First of all, we predicted that that when the side-taker failed to take their friend (the participant)'s side

(remained neutral or sided against them), participants would again feel closer to the side-taker when in the wrong than when in the right. We made this prediction based on the results of Study 2.1; participants would be more forgiving of the lack of support when in the wrong. Second, we predicted that participants would make inferences about the side-taker's moral beliefs based on the side-taking response.

Method

Participants. 300 AMT participants (173 females, ages ranged from 19 – 84 years, M = 36.10, SD = 11.74) were randomly assigned to one of six conditions: when the participant was in the wrong and the friend either remained neutral (n = 48), sided against the participant (n = 50), or sided with the participant (n = 51); and when the participant was in the right and the friend either remained neutral (n = 48), sided against (n = 49), or sided with the participant (n = 51).

Procedure. We used a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) between-participants design. Participants first thought of a friend and chose the most suitable pronouns to be used throughout the vignette to reference their friend . The vignette was about three people: the participant (the protagonist), their close friend (the side-taker), and an acquaintance (the opponent). Participants read that they shared a story with their close friend at a bar. The story was about the participant getting into a dispute with the acquaintance. Across conditions, we varied two things – whether the participant was in the wrong in the dispute (moral standing) and how the friend responded (decision). Figure 1 shows a schematic of the scenario. Participants read the following scenario (pronouns gender-matched their friend but for ease of reading female pronouns are used throughout):

> Imagine you have a close friend named Jamie. One evening, you and Jamie are at a bar catching up. You just got back from a friend's wedding and are telling Jamie about something that happened with an acquaintance of yours, Casey. Although

> Jamie hasn't actually met Casey, you've talked about Casey in conversations with Jamie many times.

The vignette diverged in line with the two moral standing conditions. Participants in the wrong condition read:

> You tell Jamie that before the wedding, Casey was going to make her own accommodation plans, but that you had, at the time, said that Casey could stay with you at your place. "I'm sure you'll be able to stay here", you'd said. Casey had replied, "No worries, I'd rather make my own plans just to be sure." To which you'd replied, "No, really, I promise you'll be able to stay here."
>
> "So, when Casey got to mine, I told her that I was sorry but that she needed to find another place to stay because I'd said yes to another friend who was in town for a business trip!" You tell Jamie. "Casey was so annoyed, she had all of her bags with her and was pretty tired from traveling. Casey started going off on me and slammed my door as she left."

While participants in the right moral standing condition read:

> You tell Jamie that before the wedding, you were going to make your own accommodation plans, but that Casey had, at the time, said that you could stay with her at her place. "I'm sure you'll be able to stay here", Casey had said. You'd replied, "No worries, I'd rather make my own plans just to be sure." To which Casey had replied, "No, really, I promise you'll be able to stay here."
>
> "So, when I get there, Casey tells me that she's sorry but that I needed to find another place to stay because she had said yes to another friend who was in town for a business trip!" You tell Jamie. "I was so annoyed, I had all of my bags with me and was pretty tired from traveling. I started going off on Casey and slammed her door as I left."

All participants then read:

> You continue, "Now we are both mad at each other. Who do you think was in the wrong?
>
> Jamie looks at you and says,

Across conditions, we then varied friend's side-taking decision by what Jamie said. In the neutral condition, Jamie said "I don't want to take sides in this case." In the sided-against condition, it was "It sounds like you were in the wrong in this case," and in the sided-with condition, Jamie said "It sounds like Casey was in the wrong in this case." After reading one of the vignettes,

participants answered three questions each on a 7-point Likert scale. The first two questions were identical to the questions from Study 2.1, measuring closeness and relationship damage on a scale of 1 (*not at all*) to 7 (*a lot*), and again showed a strong correlation, $r(300) = .71$, $p < .001$. Therefore, we again combined them into a single "friendship" measure. We also measured participants' inferences about the side-taker's moral beliefs about promise-keeping; they rated their agreement with the statement: "*Keeping promises is important to Jamie*", on a scale of -3 (*completely disagree*) to 3 (*completely agree*).

Results

Friendship Measure. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' evaluation of the friendship revealed a main effect of moral standing, $F(1, 294) = 10.04$, $p = .002$, $\eta^2 = .03$, such that participants felt more negative toward their friend when the participant was in the right than in the wrong. The ANOVA also revealed a main effect of decision, $F(2, 294) = 53.96$, $p < .001$, $\eta^2 = .27$, such that participants felt as negative toward their friend for siding against them as remaining neutral, $t(197) = 0.92$, $p = .356$. Further, participants felt more negative toward their friend when their friend sided against them ($M = -0.39$, $SD = 1.24$) than with them ($M = 0.90$, $SD = 1.28$), $t(200) = 7.27$, $p < .001$, and more negative when their friend remained neutral ($M = -0.53$, $SD = 0.94$) than sided with them, $t(197) = 8.98$, $p < .001$. The analysis also revealed a significant moral standing x decision interaction, $F(2, 290) = 22.92$, $p < .001$, $\eta^2 = .14$, which we followed up on with planned comparisons.

We were again primarily interested in comparing participants' response to their friend's side-taking decisions in circumstances where the participant was in the right vs. in the wrong. When their friend remained neutral, participants felt more negative toward their friend when they

were in the right (*M* = -0.79, *SD* = 0.99) than when they were in the wrong (*M* = -0.26, *SD* = 0.82), *t*(95) = 2.89, *p* = .005. Similarly, when their friend sided-against them, participants felt more negative toward their friend when they were in the right (*M* = -1.05, *SD* = 1.16) than when in the wrong (*M* = 0.29, *SD* = 0.92) , *t*(99) = 6.44, *p* < .001. Consistent with our hypotheses and previous study, these results indicate that a friend's failure to side with the participant has a less negative impact on the friendship when the participant was in the wrong. When their friend sided with them, participants felt more negative toward their friend when they were in the wrong (*M* = 0.56, *SD* = 1.38) than when they were in the right (*M* = 1.25, *SD* = 1.06), *t*(99) = 2.81, *p* = .006. See Figure 2.2a.

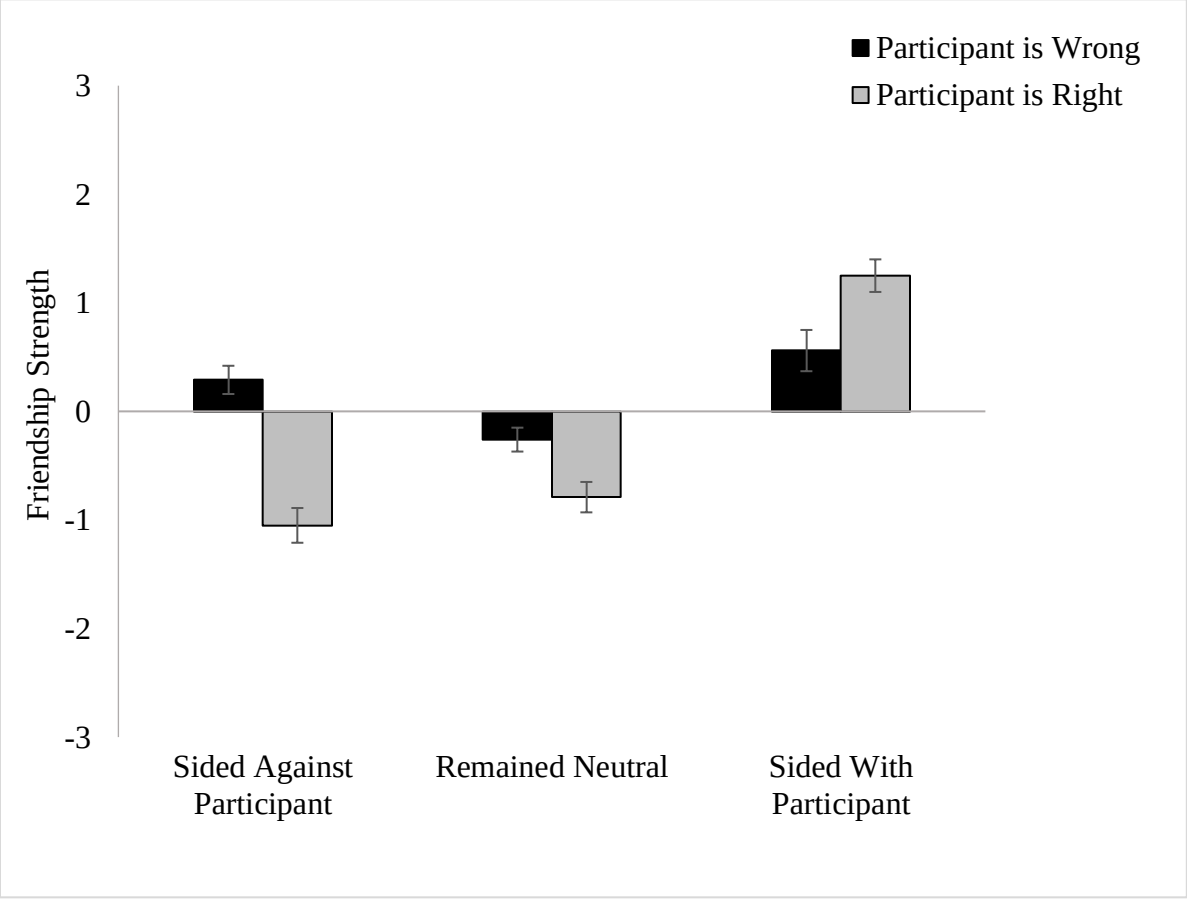

Figure 2.2a. Participants' reported friendship strength with the side-taker by the participant's moral standing and the friend's side-taking decision (Study 2.2).

Moral Inference. We conducted a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' judgments about their friend's (i.e. the side-taker's) beliefs about keeping promises. We found a main effect of moral standing, $F(1, 294) = 9.74$, $p = .002$, $\eta^2 = .03$, such that participants inferred promise-keeping to be less important to their friend when the participant was in the right than in the wrong. See Figure 2b. There was no significant effect of decision, $F(2, 294) = 0.66$, $p = .516$, $\eta^2 = .04$, and there was a significant interaction, $F(2, 290) = 119.99$, $p < .001$, $\eta^2 = .45$, which we followed up with planned comparisons.

Our primary interest was in comparing participants' inferences about their friend's moral beliefs in cases where their friend remained neutral when the participant was in the right vs. in the wrong. When their friend remained neutral, participants inferred their friend cared less about keeping promises when the promise breaker was a non-friend (remained neutral, right condition; $M = -0.42$, $SD = 1.34$) as compared to when the participant (who was close friend's with the side-taker's) was the promise-breaker (remain neutral, wrong condition; $M = 1.19$, $SD = 1.32$), $t(96) = 5.99$, $p < .001$. That is, when loyalty and morality made differing suggestions about which side the friend should take (when the side-taker's friend was the promise-breaker) participants did not make a negative inference about the side-taker's moral beliefs. However, when both morality and friendship made the same suggestion (remaining neutral when someone broke a promise against the side-taker's friend, the participant) participants did make a negative evaluation about the side-taker's beliefs: she must not care very much about the moral rule, in this case keeping promises, or she would have sided against the promise-breaker (in this case, the acquaintance). See Figure 2.2b.

We also examined participants' inferences when the side-taker chose a side depending on whether the participant or her opponent was the promise-breaker. When the participant was the promise-breaker (i.e. in the wrong), participants inferred that keeping promises was less important to their friend (the side-taker) when the side-taker sided with the promise-breaker (sided-with participant, wrong condition; M = -1.04, SD = 1.47) than against the promise-breaker (sided-against participant, wrong condition; M = 2.06, SD = 1.25), $t(99) = 11.40$, $p < .001$. Similarly, when the opponent was the promise-breaker (i.e. right condition), participants inferred their friend (the side-taker) cared less about keeping promises when the side-taker sided with the promise-breaker (sided-against participant, right condition; M = -0.84, SD = 1.82) than against the promise-breaker (sided-with participant, right condition; M = 1.94, SD = 1.20), $t(99) = 9.07$, $p < .001$. That is, participants judged that keeping promises was important to someone who sided against a promise-breaker, but not someone who sided with a promise-breaker. This general pattern occurred regardless of whether the participant or her opponent broke the promise. Figure 2b shows these results; note that to look at moral inferences, we re-sorted the x-axis according to whether the side-taker sided with or against the promise-breaker (rather than whether the side-taker sided with or against their friend, the participant).

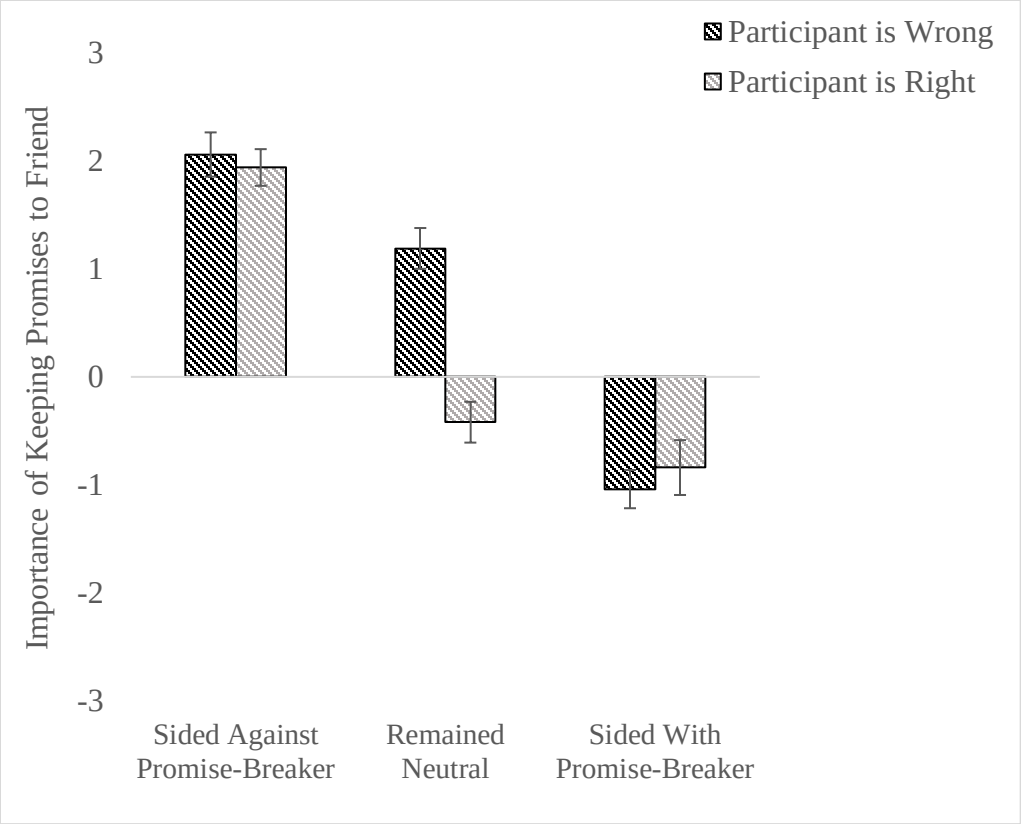

Figure 2.2b. Mean (*SE*) for participants' moral inferences in Study 2.2. The dependent measure is how important participants believe keeping promises is to their friend, where negative values indicate participants inferring lower morality (promise-keeping to be less important to their friend) and positive values indicate inferences of higher morality (promise-keeping to be more important to their friend) based on their friend's side-taking decision.

Discussion

The results from Study 2.2 replicated and extended our findings from Study 1: participants felt less negative about their friend's failure to take their side (siding against them or remaining neutral) when they were in the wrong than when they were in the right. These results again suggest that morality can mitigate some of the negative costs associated with failing to side with a friend. We argue that this occurs because in such cases a friend's refusal to take one's side is no longer diagnostic of whether they are a loyal friend or not—the friend might base their side-taking behavior on their own moral beliefs, a defensible justification for lack of support.

We also found that people make moral inferences about their friend's beliefs based on that friend's side-taking response and whether the participant was morally in the right or wrong. Specifically, when their friend sided against the promise-breaker, participants inferred that keeping promises was important to their friend, and when their friend sided with the promise-breaker participants inferred that keeping promises was not important to their friend. When a friend took sides on these moral issues, participants made these inferences about promise-breaking regardless of whether the participant or her opponent broke the promise. Interestingly, participants made different inferences about their friend's moral beliefs based on her decision to remain neutral in situations where the participant was in the wrong or right. When the participant was in the wrong (i.e. they were the promise-breaker) and their friend remained neutral, participants inferred that keeping promises was important to their friend, presumably because in this circumstance morality and loyalty would pull in opposite directions and they likely thought their friend remained neutral due to their friendship. In contrast, neutrality from a friend when the participant was in the right (i.e. the acquaintance broke a promise made to the participant) led participants to infer that keeping promises was not very important to their friend (the side-taker), because here morality and loyalty would both push the friend to take the participant's side and so their refusal to do so might mean that they may hold looser morals in regards to promise-breaking, i.e. that they do not think of promise-breaking as being that bad of a moral transgression. Taken together, these results suggest that people make complex inferences about personal loyalties and friends' moral beliefs based on those friends' side-taking decisions and whether the participant was in the moral right or wrong.

Study 2.3

Our results thus far suggest that people use information about a side-taker's decision and whether they were in the wrong or right to calibrate how they feel about their friendship with the side-taker and also to make inferences about the side-taker's morality. In Study 3 we attempted to replicate these results in a new situation involving someone doing cocaine at someone else's house (without first gaining permission to do so) as the moral transgression. Further, and to extend our findings, in Study 2.3 our moral inference question asks about the side-taker's moral behavior, i.e. – given the side-taker's siding response, the likelihood of the side-taker engaging in the immoral act in question.

To investigate these possibilities, participants were again told to imagine that they were either in the wrong or in the right and their friend either sided with them, sided against them, or remained neutral in a conflict before rating the impact of their friend's response on their friendship. We made the same predictions as our previous studies for this measure: people would respond less negative to a friend siding against them or remaining neutral when they were in the wrong than when they were in the right.

Participants were also asked a morality measure about the side-taker, but this time they were asked to make a judgment about the side-taker's likelihood of engaging in the moralized behavior: how likely the side-taker would be to engage in drug-taking, based on their siding response. The side-taking account does not require that people make the inference that the side-taker will actually engage in the transgression they fail to condemn, but it is a possible inference that one could make based on their friend's side-taking decision if one thinks that people are more likely to refrain from condemning behaviors that they themselves engage in (i.e. that they realize that others may engage in strategic moral condemnation, DeScioli & Kurzban, 2013).

Therefore, we expected that participants would infer not only that a side-taker condones cocaine use if they fail to condemn it, but also that the side-taker may engage in that behavior. In line with our previous results, we expected that siding with the cocaine-user would lead people to think that the side-taker uses cocaine and siding against the cocaine-user would make people think that the side-taker does not use cocaine. Further, in line with the side-taking hypothesis, we predicted that people would think that the side-taker was more likely to use cocaine when she remained neutral in the situation where her friend was in the right than when her friend was in the wrong.

Method

Participants. We recruited 303 AMT participants (99 females, ages ranged from 18 – 73 years, $M = 33.46$, $SD = 10.56$). Participants were assigned to one of six conditions: when participant was in the wrong and the friend remained neutral (n = 48), sided against the participant (n = 50), or sided with the participant (n = 51), when the participant was in the right and the friend remained neutral (n = 51), sided against the participant (n = 51), or sided with the participant (n = 52).

Procedure. We again used a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) between-participants design. Here, as with Study 2, participants were asked to first imagine a friend and select appropriate pronouns, for ease of reading we again use female pronouns throughout the vignette. The vignette was again about three people: the participant, their close friend, and another friend of the participant with whom the participant got into a conflict. Participants read that they were at a bar with their close friend and were recalling a story about the participant and the other friend recently getting into a dispute at a wedding. Across conditions, we again varied two things – who was responsible for

the dispute: the participant (wrong condition) or the other friend (right condition), and friend's side-taking decision, i.e. how the participant's close friend responded to the fight: remained neutral, sided against, or sided with the participant. Participants in all conditions first read:

> Imagine you have a close friend named Jamie. One evening, you and Jamie are at a bar catching up. You just got back from a friend's wedding and are telling Jamie about something that happened with your other friend, Casey.

The vignettes then diverged depending on moral standing. Participants in the wrong condition read the following:

> You tell Jamie that you were staying at Casey's place during the wedding. You two are pretty good friends, and so you were excited to spend some time with Casey. You tell Jamie, "Everything was going great until Casey and I got into a fight. That Saturday night, before we went out, I did some cocaine in Casey's bathroom. Casey walked in halfway through and completely lost it. I told Casey I didn't realize it would be a problem, which was why I didn't ask. But Casey was super mad, asked me to leave early, and told me never to ask to stay over again. We haven't really spoken since."
>
> You continue, "Can you believe that happened? Do you think Casey was out of line?"

While participants in the right condition read:

> You tell Jamie that Casey was staying at your place during the wedding. You two are pretty good friends, and so you were excited to spend some time with Casey. You tell Jamie, "Everything was going great until Casey and I got into a fight. That Saturday night, before we went out, Casey did some cocaine in my bathroom. I walked in halfway through and completely lost it. Casey didn't realize it'd be a problem, which is why Casey didn't ask me apparently. But I was super mad, I told Casey to leave early, and to never ask to stay at my place again. We haven't really spoken since."
>
> You continue, "Can you believe that happened? Do you think Casey was out of line?"

Across conditions, we then varied friend's side-taking decision in a manner identical to previous studies. In the neutral condition, Jamie refused to take sides, in the sided against condition, Jamie sided against the participant, and in the sided with condition, Jamie supported the participant.

After reading the vignette, participants answered three questions each on a 7-point Likert scale. The first two questions were again identical to the friendship measures from previous studies. Answers to these questions again showed a strong correlation, $r(203) = .73$, $p < .001$; so we again combined them into a single friendship measure. The third question was intended as a moral inference question, this time about participants' inferences about their friend's own moral behavior, and asked participants the extent to which they agreed with the statement: "*Jamie is likely to do cocaine*.", on a scale of 1 (*completely disagree*) to 7 (*completely agree*). Note, to make the moral inference measure more comparable to our measure from Study 2.2, we first reverse-coded this question such that lower numbers signified that participants thought their friend would more likely do drugs, and then we re-scaled it to range from -3 to 3, as Study 2.2 also. Therefore, when examining the corresponding graph and results below, a more negative score would mean that participants would rate their friend as more likely to commit the transgression – i.e. take drugs.

Results

Friendship Measure. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' evaluation of the friendship revealed a significant effect of moral standing, $F(1, 297) = 5.18$, $p = .024$, $\eta^2 = .02$, such that participants felt more negative toward their friend when the participant was in the right than in the wrong. The ANOVA also revealed a significant effect of decision, $F(2, 297) = 79.43$, $p < .001$, $\eta^2 = .35$, such that participants felt as negative toward their friend for siding against them ($M = -0.74$, $SD = 1.45$) or remaining neutral ($M = -0.83$, $SD = 1.08$), $t(198) = 0.53$, $p = .598$; participants felt more negative toward their friend when their friend sided against them than with them ($M = 1.08$, $SD = 1.36$), $t(202) = 9.23$, $p < .001$, and more negative when their friend remained neutral than

sided with them, $t(200) = 11.04$, $p < .001$. Finally, there was also a significant moral standing x decision interaction, $F(2, 297) = 22.57$, $p < .001$, $\eta^2 = .13$, which we followed up on with planned comparisons.

We were again primarily interested in comparing participants' response to their friend's side-taking decisions in circumstances where participants were in the right vs. in the wrong. When their friend remained neutral, participants felt similarly negative about the friendship regardless of whether they were in the right ($M = -0.79$, $SD = 1.21$) or in the wrong ($M = -0.89$, $SD = 0.94$), $t(97) = 0.42$, $p = .677$. When their friend sided against them, participants felt more negative toward their friend when they were in the right ($M = -1.54$, $SD = 1.15$) than when they were in the wrong ($M = 0.07$, $SD = 1.27$), $t(99) = 6.67$, $p < .001$. Somewhat consistent with our hypotheses and previous results, a friend siding against the participant has less negative impact on the friendship when the participant was in the wrong. Further, when the friend sided with them, participants felt more negative toward their friend when they were in the wrong ($M = 0.79$, $SD = 1.61$) rather than in the right ($M = 1.36$, $SD = 1.01$), $t(101) = 2.12$, $p = .036$. See Figure 2.3a.

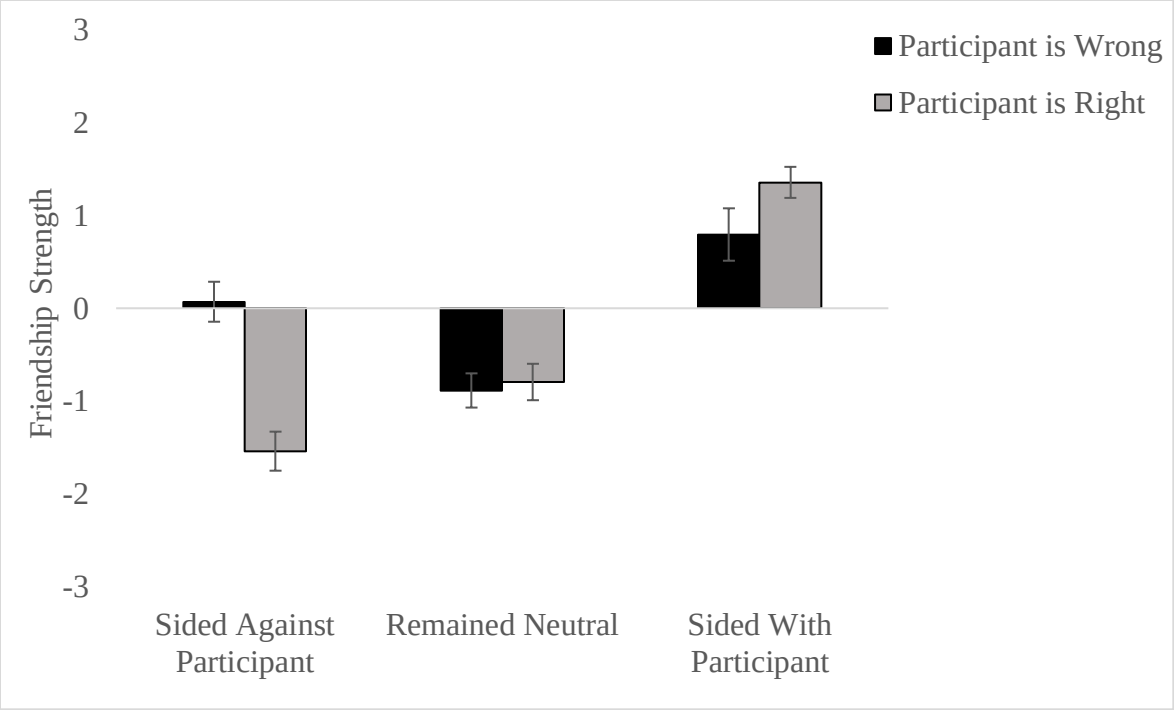

Figure 2.3a. Participants' reported friendship strength with the side-taker by the participant's moral standing and the friend's side-taking decision (Study 2.3).

Moral Inference. We conducted a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' judgments about their friend's (i.e. the side-taker's) beliefs about drug use. We found a significant effect of moral standing, $F(1, 297) = 9.27$, $p = .003$, $\eta^2 = .03$, such that participants inferred their friend to be more likely to engage in drug use when the participant was in the right than in the wrong. See Figure 3b. The ANOVA also revealed a significant effect of decision, $F(2, 297) = 12.18$, $p < .001$, $\eta^2 = .10$. As in our previous studies, these were qualified by a large moral standing by decision interaction, $F(2, 297) = 111.25$, $p < .001$, $\eta^2 = .43$, which we followed up on with planned comparisons.

Our primary interest was in comparing participants' inferences about their friend's moral behavior in cases where their friend remained neutral when the participant was in the right vs. in

the wrong. When their friend remained neutral, participants inferred their friend more likely to do cocaine when the cocaine-user (i.e. the transgressor) was the non-friend (remained neutral, right condition; M = -0.33, SD = 1.34) than when the cocaine-user was the participant (remained neutral, wrong condition; M = 1.75, SD = 1.39), $t(97) = 7.60$, $p < .001$. That is, when loyalty and morality made differing suggestions about which side the friend should take (when one's friend was the cocaine-user) participants did not make a negative inference about the friend's moral behavior regarding cocaine use. However, when both morality and friendship made the same suggestion (remaining neutral when someone committed a transgression against one's friend) participants did make a negative evaluation about the friend's behavior: she must not care very much about the moral rule and, therefore, she may be inclined to use cocaine herself, or she would have sided against the cocaine-user (in this case, the acquaintance). Hence, participants made a more negative evaluation about their friend's moral behavior when their friend remained neutral in a situation where the participant was in the right rather than in the wrong. See Figure 2.3b.

We also examined participants' inference when the friend chose a side depending on whether the participant or her opponent was the drug-user. When the participant was the drug-user (i.e. in the wrong), participants inferred their friend would be more likely to do cocaine when the friend sided with the drug-user (sided-with participant, wrong condition; M = -1.84, SD = 1.30) than against the drug-user (sided-against participant, wrong condition; M = 1.08, SD = 1.65), $t(99) = 9.89$, $p < .001$. Similarly, when the opponent was the drug-user (i.e. right condition), participants inferred their friend more likely to do cocaine when the friend sided with the drug-user (sided-against participant, right condition; M = -1.24, SD = 1.21) than against the drug-user (sided-with participant, right condition; M = 1.08, SD = 1.56), $t(101) = 8.40$, $p < .001$.

That is, participants judged that drug-use would be more likely from their friend if the friend sided with a drug-user, but not when they sided against a drug-user. This general pattern occurred regardless of whether the participant or her opponent was the person taking cocaine. Figure 3b shows these results; note that similar to our moral inference graph for Study 2.2, to look at moral inferences, we re-sorted the x-axis according to whether the friend sided with or against the cocaine-user (rather than whether the friend sided with or against the participant or whether the participant was or was not the cocaine-user.

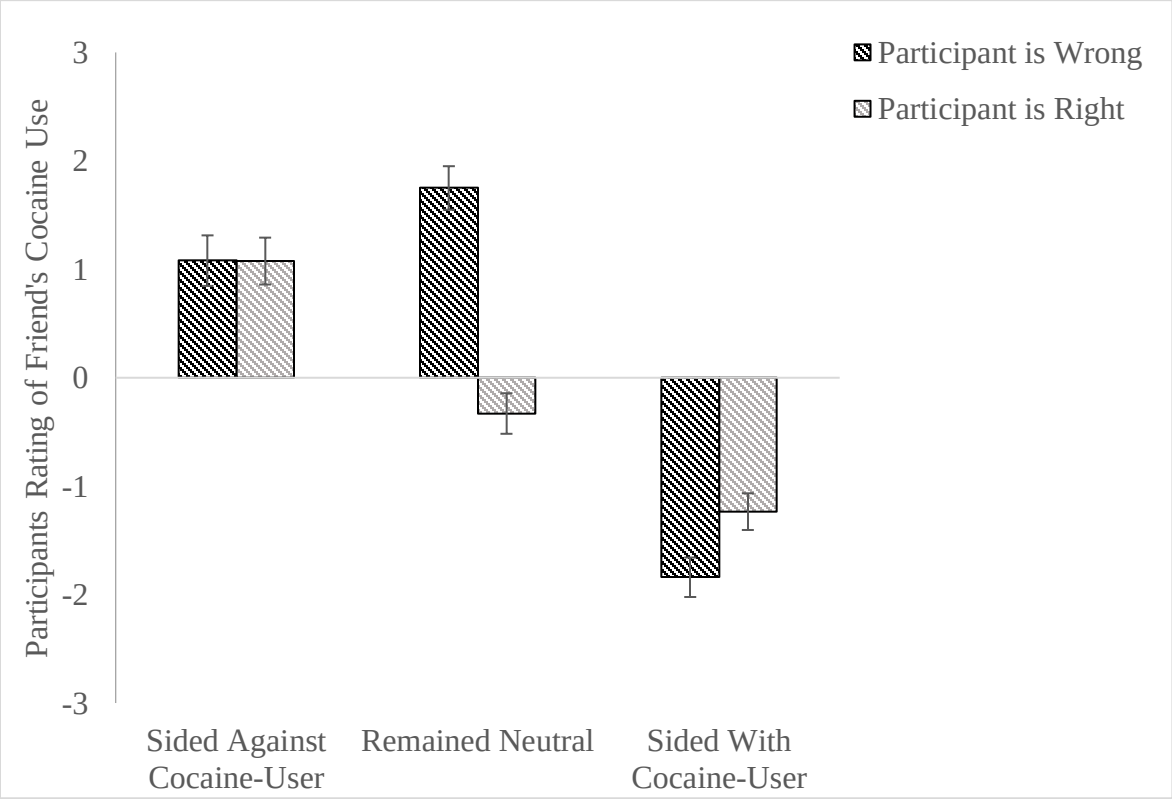

Figure 2.3b. Mean (*SE*) for participants' moral inferences in Study 2.3. The dependent measure is how likely participants believe their friend is to use cocaine, where negative values indicate participants inferring lower morality (higher likelihood of cocaine use by friend) and positive values indicate inferences of higher morality (lower likelihood of cocaine use by friend) based on their friend's side-taking decision.

Discussion

The results of Study 2.3 replicate and extend our findings from previous studies: people feel less negative about their friend failing to take their side when they are in the wrong than in the right and people also make moral inferences about their friend based on that friend's side-taking behavior. Moreover, we extended our moral inference findings from Study 2, finding that people not only make inferences about others' moral beliefs based on their side-taking behavior, but also make inferences about their moral behavior. Neutrality on a moral issue is interpreted differently when one is in the right as opposed to in the wrong on the moral issue. That is, when one is responsible for a conflict, neutrality from a friend is not taken as opposition to the friendship and is instead taken as silent condemnation about the moral transgression. However, when one is the victim in the conflict, neutrality from a friend leads one to believe the friend would also commit the moral transgression because both friendship and morality should have led the friend to take one's side, meaning that a refusal to do just that may mean that the friend does not agree that the transgression was morally wrong.

Study 2.4

Our results suggest that people make different inferences about their friendship and their friend's morality based on their friend's side-taking decisions and whether they themselves are in the moral wrong or right. However, we have thus far only examined the inferences that people make when they are directly involved in a conflict. While it is important to understand the inferences that people make about friendships and their friend's morality, it is also important to understand how third parties interpret such decisions (DeScioli & Kurzban, 2013). Indeed, people may be more tempted to take sides based on loyalty in a world where no one outside of the conflict would ever know about their decision. If third parties make inferences about a

person's morality based on their side-taking decisions, then people might be less likely to take their friend's side when their friend is in the wrong. Thus, it seems important to understand whether third parties make such inferences based on the factors we found to be relevant in our previous studies.

In Study 2.4, we investigated the inferences that third parties make by having participants read a scenario in which they were third-party evaluators of a scenario about three people: a protagonist (formerly the participant in previous studies), her friend (the side-taker), and her opponent in the conflict (an acquaintance of the protagonist). We used a scenario almost identical to Study 2 in which one person broke a promise they'd made to another person (the protagonist or acquaintance). We again varied whether the protagonist was in the wrong or in the right in a conflict with the opponent. The friend of the protagonist (the side-taker) again either remained neutral, sided against, or sided with the protagonist. Participants were asked how the protagonist would feel about the friendship based on the side-taker's decisions and also their inferences about the side-taker's likelihood of engaging in the bad action (would they break a promise?)

Thus, in Study 2.4 we explore if our main findings from the previous studies replicate in this third-party context. If we did replicate our previous findings, we would find that participants expect the protagonist to feel less negative about neutrality and being sided against when the protagonist was in the wrong than when she was in the right. Further, we would find that participants make different moral inferences about the side-taker's moral behavior when the side-taker decides to remain neutral depending on whether the protagonist in the wrong or in the right. We expected that participant inferences about the friend's moral behaviors based on their neutral response would be in line with our previous findings. Specifically, when the protagonist was in

the wrong, participants should infer that the friend believes the action is wrong, but refuses to take sides because of her loyalty to her friend. In contrast, a friend's neutrality when the protagonist is in the right should lead people to infer that the friend's own moral code (in regards to the transgression) may be lower than the protagonist's. Additionally, we wanted to run third-party scenarios in case people were having trouble imaging themselves committing a moral transgression. Given that our measures are all self-report questions, our previous results may have been subject to certain demand characteristics. We discuss this further in the discussion, and therefore, similar results from a third-party scenario would provide stronger evidence in support of our hypotheses.

Method

Participants. We recruited 306 participants (151 females, ages ranged from 18 – 69, M = 35.16, SD = 10.99). Participants were randomly assigned to one of six conditions: when the protagonist is the wrong and their friend remained neutral (n = 45), sided against (n = 55), or sided with the protagonist (n = 54); when the protagonist was in the right in the conflict and the friend remained neutral (n = 45), sided against (n = 52), or sided with the protagonist (n = 55).

Procedure. We again used a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) between-participants design. Here, participants were asked to first imagine two friends at a bar and the protagonist is telling their friend about something that happened with an acquaintance of theirs (the opponent). Therefore, the vignette was again about three people: the protagonist, their close friend, and the protagonist's opponent. Across conditions, we varied two things – whether the protagonist was responsible for the conflict (wrong condition) or not (right condition); and how the protagonist's friend responded. Participants in all conditions first read:

> Imagine two close friends named Adrian and Jamie. One evening, Adrian and Jamie are at a bar catching up. Adrian just got back from a friend's wedding and is telling Jamie about something that happened with an acquaintance, Taylor. Although Jamie hasn't actually met Taylor, Adrian has mentioned Taylor before.

The vignette then varied by moral standing condition. Participants in the wrong condition read the following:

> Adrian says, "Before the wedding, Taylor was going to make plans to stay at a hotel and I had, at the time, said that Taylor could stay at mine during the wedding. Taylor appreciated the offer but wanted to make other plans just in case, to which I had replied, "No, really, I promise you'll be able to stay here."
>
> "So, when Taylor got to mine, I apologized and said that Taylor needed to find another place to stay because I had said yes to another friend who was in town for a business trip!" Adrian tells Jamie. "Taylor was so annoyed and was pretty tired from traveling. Taylor started going off on me and slammed my door on the way out."
>
> Adrian continues, "Now we are both mad at each other. Who do you think was in the wrong?"
>
> Jamie looks at Adrian and says,

While participants in the right condition read:

> Adrian says, "Before the wedding, I was going to make plans to stay at a hotel and Taylor had, at the time, said that I could stay over during the wedding. I'd told Taylor that I appreciated the offer but that I'd rather make my own plans just in case, to which Taylor had replied, "No, really, I promise you'll be able to stay here."
>
> "So, when I arrived, Taylor apologized and told me that I needed to find another place to stay because Taylor had said yes to another friend who was in town for a business trip!" Adrian tells Jamie. "I was so annoyed and was pretty tired from traveling. I started going off on Taylor and slammed the door as I left."
>
> Adrian continues, "Now we are both mad at each other. Who do you think was in the wrong?"
>
> Jamie looks at Adrian and says,

Across conditions, we then varied the friend's side-taking decision. In the neutral condition, Jamie refused to take sides, in the sided-against condition, Jamie sided against the protagonist,

Adrian, and in the sided-with condition, Jamie supported Adrian. After reading one of the six versions of the vignette, participants answered three questions each on a 7-point Likert scale. The two friendship questions were similar to those of previous studies in asking about closeness and relationship damage, but were worded in line with the third-party format to inquire about the relationship between the protagonist and their friend. Answers to the two friendship questions (closeness to friend and relationship damage) again showed a strong correlation, $r(306) = .65$, $p < .001$; so we again combined them into a single friendship measure and, as with previous studies coded the friendship measure as -3 (*strongly negative*) and 3 (*strongly positive*).The moral inference question asked participants to make inferences about the friend's moral beliefs. However, this time we asked two moral inference questions. The first was identical to the moral inference question from Study 2.2 except was presented in a third-party context. Further, we presented the first moral inference question before the friendship questions in order to rule out any impact of question order on our results. The second asked participants' inferences about the friend's moral beliefs about the transgression in general. Given that all of our inference questions that far were specific to the scenario (prefaced with "Given what happened...") we added the second moral inference question in order to examine whether participants thought the friend's moral beliefs about the transgression in general differed and was worded as: "*In general, Jamie thinks it is wrong to break a promise*" coded on a 7-point Likert scale from 1 (*completely disagree*) to 7 (*completely agree*). Participant response to the two moral inference questions was strongly correlated, $r(306) = .81$, $p < .001$.

Results

Friendship Measure. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided against, sided with) ANOVA on participants' evaluation of the friendship revealed a

significant main effect of moral standing, $F(1, 300) = 21.38$, $p < .001$, $\eta^2 = .07$, such that participants rated that the protagonist felt more negative toward their friend when the protagonist was in the right than in the wrong. The ANOVA also revealed a significant effect of decision, $F(2, 300) = 76.62$, $p < .001$, $\eta^2 = .39$, such that participants rated that the protagonist felt as negative toward their friend for siding against them ($M = -0.49$, $SD = 1.09$) or remaining neutral ($M = -0.54$, $SD = 1.08$), $t(195) = 0.34$, $p = .734$; participants, however, rated the protagonist felt more negative toward their friend when their friend sided against them than with ($M = 1.07$, $SD = 1.18$) them, $t(214) = 10.09$, $p < .001$, and more negative when their friend remained neutral than sided with them, $t(197) = 9.98$, $p < .001$. Finally, there was also a significant moral standing x decision interaction, $F(2, 300) = 3.92$, $p = .021$, $\eta^2 = .03$, which we followed up on with planned comparisons.

We were primarily interested in comparing participants' response to the friend's side-taking decisions in circumstances where the protagonist was in the right vs. in the wrong. In line with our hypotheses and previous studies, when the friend remained neutral, participants rated the protagonist felt more negative toward their friend when in the right ($M = -0.93$, $SD = 1.01$) than when in the wrong ($M = -0.14$, $SD = 1.01$), $t(88) = 3.70$, $p < .001$. When the friend sided against the protagonist, participants rated the protagonist felt more negative toward their friend when in the right ($M = -0.91$, $SD = 1.04$) than when in the wrong ($M = -0.08$, $SD = 0.99$), $t(105) = 4.24$, $p < .001$. These data suggest that even in third-party contexts, people predict someone to be less upset with their friend for failing to take their side when that person was in the wrong than when they are in the right. When the friend sided with the protagonist, participants judged that the protagonist felt similarly toward their friend whether the protagonist was in the wrong ($M = 1.12$, $SD = 1.11$) or right ($M = 1.03$, $SD = 1.25$), $t(107) = 0.41$, $p = .681$. See Figure 2.4a.

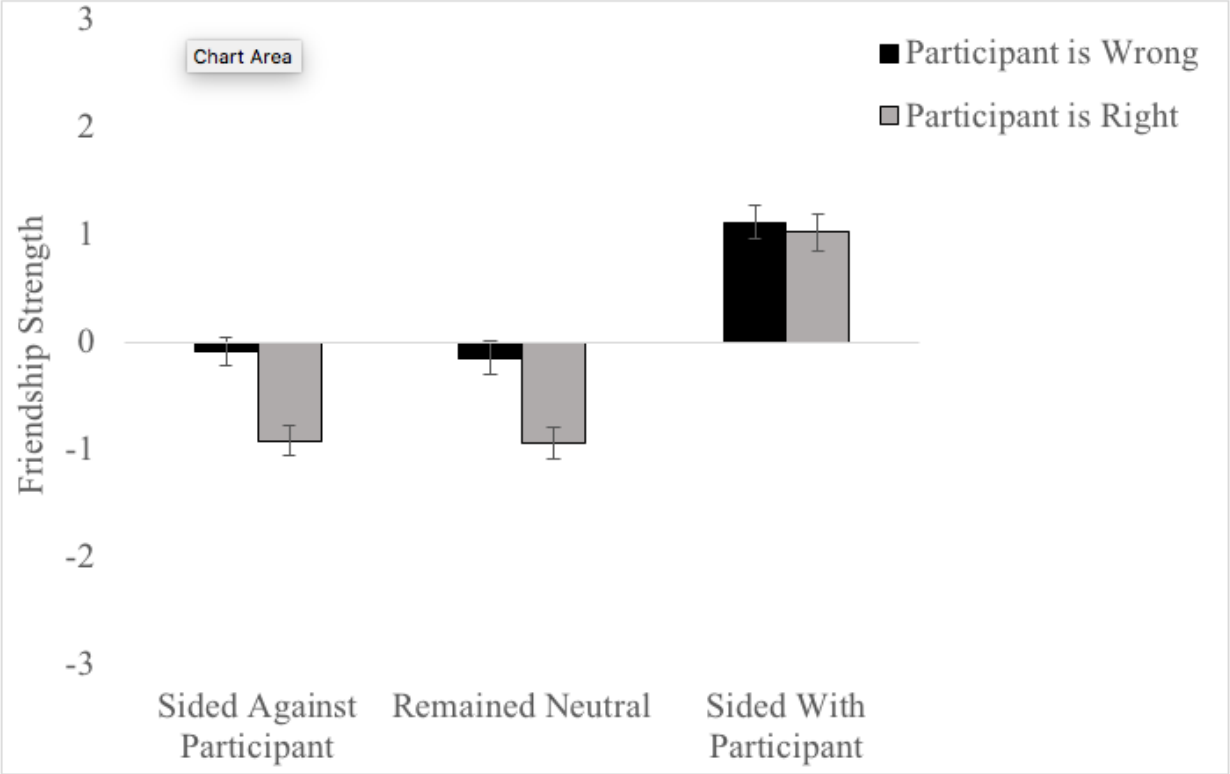

Figure 2.4a. Participants' reported friendship strength with the side-taker by the participant's moral standing and the friend's side-taking decision (Study 2.4).

Moral Inference. We conducted a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided against, sided with) ANOVA on participants' judgments about the friend's (i.e. the side-taker's) beliefs about keeping promises. We found no effect of moral standing, $F(1, 300) = 1.99$, $p = .159$, $\eta^2 = .01$, and no effect of decision, $F(2, 300) = 2.22$, $p = .111$, $\eta^2 = .02$. However, there was a large moral standing x decision interaction, $F(2, 300) = 207.51$, $p < .001$, $\eta^2 = .58$, which we followed up on with planned comparisons.

Our primary interest was in comparing participants' inferences about the friend's moral beliefs in cases where their friend remained neutral when the participant was in the right vs. in the wrong. In line with our previous results, when the friend remained neutral, participants inferred the friend cared less about keeping promises when the promise-breaker was a non-friend

(remained neutral, right condition; M = - 0.20, SD = 1.32) as compared to when the protagonist (who was the side-taker's friend) was the promise-breaker (remained neutral, wrong condition; M = 0.80, SD = 1.46), $t(88) = 3.56$, $p = .001$. That is, even in third-party contexts, when loyalty and morality made differing suggestions about which side the friend should take (when the protagonist was the promise-breaker) participants did not make a negative inference about the friend's moral beliefs. However, when both morality and friendship made the same suggestion (remaining neutral when someone broke a promise against one's friend) participants did make a negative evaluation about the friend's beliefs: she must not care very much about the moral rule, in this case keeping promises, or she would have sided against the promise-breaker (in this case, the acquaintance). Hence, even in third-party situations, participants made a more negative evaluation about the friend's moral behavior when the friend remained neutral in a situation where their friend – the protagonist – was in the right rather than in the wrong.

We also examined participants' inferences when the side-taker chose a side depending on whether the protagonist or her opponent was the promise-breaker. When the protagonist was the promise-breaker (i.e. in the wrong), participants inferred that keeping promises was less important to the friend when the friend sided with the promise-breaker (sided-with protagonist, wrong condition; M = -1.57, SD = 1.59) than against the promise-breaker (sided-against protagonist, wrong condition; M = 2.38, SD = 0.93), $t(107) = 15.91$, $p < .001$. Similarly, when the opponent was the promise-breaker (i.e. right condition), participants inferred the friend cared less about keeping promises when the friend sided with the promise-breaker (sided-against protagonist, right condition; M = -1.08, SD = 1.64) than against the promise-breaker (sided-with protagonist, right condition; M = 2.24, SD = 0.90), $t(105) = 13.03$, $p < .001$. That is, participants judged that keeping promises was important to someone who sided against a promise-breaker,

but not someone who sided with a promise-breaker. This general pattern occurred regardless of whether the participant read that the protagonist or her opponent broke the promise. Figure 2.4b shows these results; note that to look at moral inferences, we re-sorted the x-axis according to whether the friend sided with or against the promise-breaker as with previous studies.

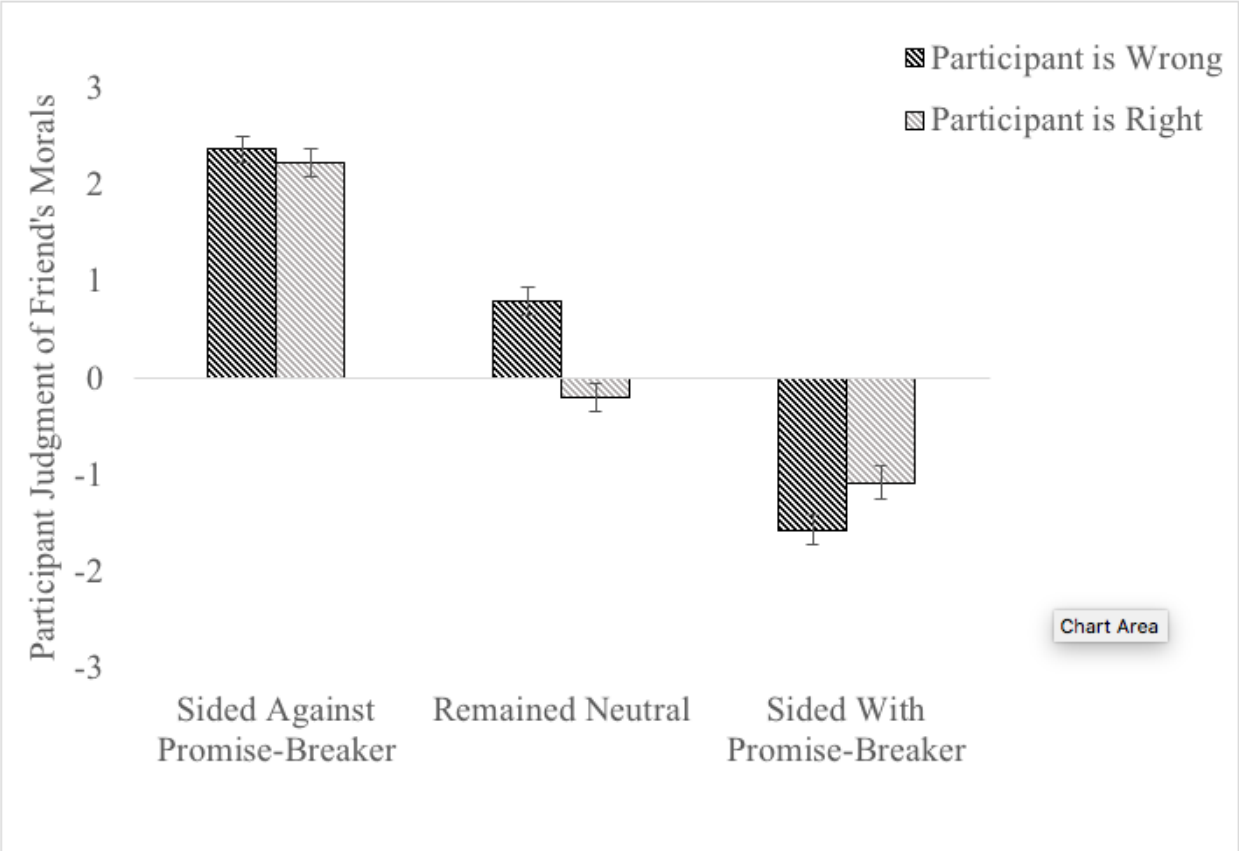

Figure 2.4b: Mean (SE) for participants' moral inferences in Study 2.4. The dependent measure is how important participants believe promise-keeping is to the friend of the protagonist, where negative values indicate participants inferring lower morality (lower importance placed on keeping promises by friend) and positive values indicate inferences of higher morality (higher importance placed on promise-keeping by friend) based on the friend's side-taking decision and the protagonist's moral standing.

Discussion

The results from Study 2.4 replicate our previous findings in a third-party context; demonstrating that third parties also make inferences about friendship and a side-taker's morality

based on her side-taking decisions. In Study 2.5 we attempt to replicate the results from Study 2.4 in a different moral domain involving plagiarism. We expected to replicate our results from Study 2.4 and provide further support for the idea that people make inferences about others' friendships and morality based on their side-taking decisions even from a third-party perspective.

Study 2.5

Method

Participants. We recruited 302 participants (122 females, ages ranged from 18 – 85, M = 32.58, SD = 10.33). Participants were assigned to one of six conditions: when the protagonist is the wrong and their friend remained neutral (n = 50), sided against (n = 50), or sided with the protagonist (n = 50); when the protagonist was in the right in the conflict and the friend remained neutral (n = 51), sided against (n = 50), or sided with the protagonist (n = 50).

Procedure. We again used a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) between-participants design. Here, participants were asked to first imagine two friends at a bar (the protagonist and their friend). Participants read that the protagonist is telling their friend about something that happened with an acquaintance of theirs (the opponent). Therefore, the vignette was again about three people: the protagonist, their close friend, and the protagonist's opponent. Across conditions, we varied two things – whether the protagonist or the opponent was responsible for the conflict; and how the protagonist's friend responded to the conflict. All participants first read:

> Imagine two close friends named Jonathan and David. One evening, Jonathan and David are catching up on the phone. Jonathan is telling David about something that happened with a work friend, Billy. Although David has heard of Billy, he does not really know him.

The vignette then varied depending on moral standing. Participants in the wrong condition read the following:

> Jonathan tells David, "The other week at work, Billy and I were having lunch and he told me this really great idea he had for a project. He hadn't told our boss the idea because he was not sure it was his best idea. Well, later that day, our boss asked to see me. As you know, I have been struggling at work and Billy has been doing pretty well. Our boss then asked if I had anything new and I accidentally told the boss Billy's idea and took credit for it."
>
> Jonathan continues, "Our boss absolutely loved the idea and for the first time in a while told me I'm doing a good job and that if I kept it up, I might get a raise soon! I know I could have told him it was Billy's idea, but I didn't. When Billy found out, he was really upset and confronted me.
>
> Anyway, Billy was super angry and now he isn't talking to me anymore. I mean it was just one idea, I didn't think it was a big deal, but Billy thinks it is."
>
> Jonathan continues, "Can you believe that happened? Do you think Billy was in the wrong?"
>
> David looks at Jonathan and says,

While participants in the right condition read:

> Jonathan tells David, "The other week at work, Billy and I were having lunch and I told him this really great idea I had for a project. I hadn't told our boss the idea because I was not sure it was my best idea. Well, later that day, our boss asked to see Billy. As you know, Billy has been struggling at work and I have been doing pretty well. Our boss then apparently asked Billy if he had anything new and Billy accidentally told our boss my idea and took credit for it."
>
> Jonathan continues, "Our boss absolutely loved the idea and for the first time in a while told Billy he's doing a good job and that if he kept it up, he might get a raise soon! Billy could have told him it was my idea, but he didn't. When I found out, I was really upset and confronted him.
>
> Anyway, I was super angry and now I'm not talking to him anymore. I mean I know it was just one idea, Billy didn't think it was a big deal, but I think it is."
>
> Jonathan continues, "Can you believe that happened? Do you think Billy was in the wrong?"
>
> David looks at Jonathan and says,

Across conditions, we then varied decision by how the friend, David, responded. In the neutral condition, David refused to take sides, in the sided-against condition, David sided against the protagonist, Jonathan, and in the sided-with condition, David supported the protagonist. Participants then answered three questions each on a 7-point Likert scale. As in Study 2.4, we wanted to ask the moral inference question regarding the friend's own behavior first and so participants in Study 2.5 were first asked to rate how much they agreed with the statement that the friend was likely to take credit for someone else's work/ideas, ranging from 1 (*completely disagree*) to 7 (*completely agree*). To again make these results comparable with all moral inference questions across studies, we reverse-coded this question such that lower numbers signified that participants thought the friend would plagiarize, and then re-scaled it for consistency on a 7-point scale from -3 to 3. Then we asked participants to rate how the protagonist would feel about their friend using our two friendship measures of the protagonist's closeness (1 = *a lot less close*; 7 = *a lot more close*) to the friend and how damaged/strengthened the friendship had been, given the friend's response (1 = *damaged a lot*; 7 = *strengthened a lot*). Answers to the two friendship questions (closeness to friend and relationship damage) again showed a strong correlation, $r(302) = .83$, $p < .001$, so we again combined them into a single friendship measure and, as with previous studies coded the friendship measure as -3 (*strongly negative*) and 3 (*strongly positive*).

Results

Friendship Measure. A 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' evaluation of the friendship revealed no effect of moral standing, $F(1, 296) = 2.70$, $p = .101$, $\eta^2 = .01$, but a significant effect of decision, $F(2, 296) = 116.57$, $p < .001$, $\eta^2 = .44$, such that participants rated that the protagonist felt more

negative toward their friend for siding against them (M = -1.25, SD = 0.95) than remaining neutral (M = -0.89, SD = 1.04), $t(199) = 2.52$, $p = .013$, or siding with them (M = 1.03, SD = 1.39), $t(199) = 13.60$, $p < .001$. Further, participants rated the protagonist felt more negative about the friendship when their friend remained neutral as compared to siding with them, $t(200) = 11.17$, $p < .001$. Finally, there was no moral standing x decision interaction, $F(2, 296) = 0.27$, $p = .762$, $\eta^2 = .002$. We believe it is worth noting that although we did not see a significant interaction here, we do still observe the expected reduction in perceived friendship strength, in line with our previous results, when the protagonist is in the wrong vs. right. See Figure 2.5a.

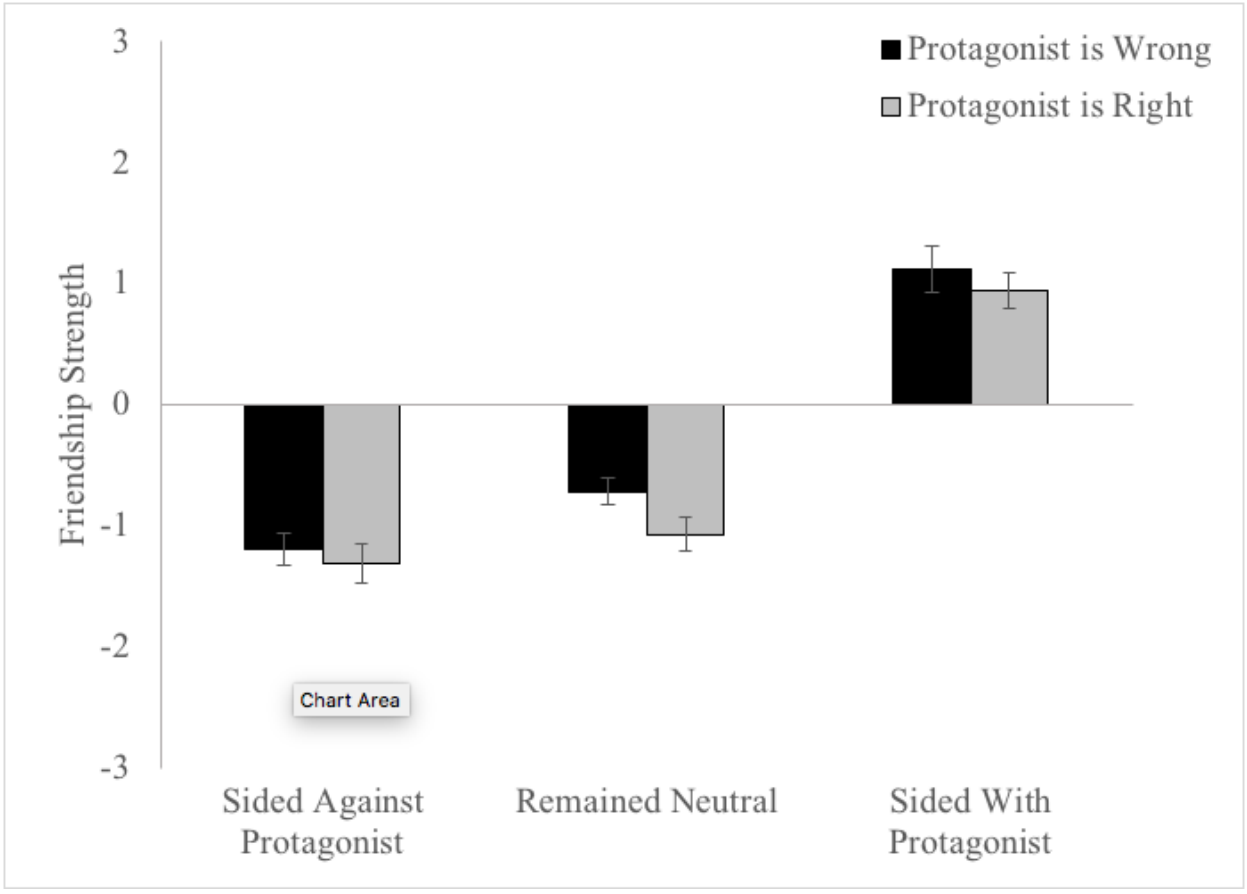

Figure 2.5a. Mean (SE) participants' response to friendship measure in Study 2.5 by protagonist's moral standing and friend's siding behavior conditions.

Moral Inference. We conducted a 2 (moral standing: wrong, right) x 3 (decision: remained neutral, sided-against, sided-with) ANOVA on participants' judgments of the friend's likelihood to plagiarize. The ANOVA revealed no main effect of moral standing, $F(1, 296) = 2.03$, $p = .155$, $\eta^2 = .007$, and a marginal effect of decision, $F(2, 296) = 2.70$, $p = .069$, $\eta^2 = .02$. Further, the ANOVA revealed a large moral standing by decision interaction, $F(2, 296) = 66.96$, $p < .001$, $\eta^2 = .31$, which we followed up on with planned comparisons.

Our primary interest was in comparing participants' inferences about the friend's moral behavior in cases where the friend remained neutral when the protagonist was in the right vs. in the wrong. When the friend remained neutral, participants inferred that the friend would be more likely to plagiarize when the non-friend plagiarized (remained neutral, right condition; $M = -0.31$, $SD = 1.52$) compared to when the protagonist (who was the side-taker's friend) plagiarized (remained neutral, wrong condition; $M = 0.86$, $SD = 1.75$), $t(99) = 3.61$, $p < .001$. Again, when loyalty and morality made differing suggestions about which side the friend should take (when one's friend – the protagonist – was the plagiarizer) participants did not make a negative inference about the friend's moral beliefs. However, when both morality and friendship made the same suggestion (remaining neutral when someone plagiarized one's friend) participants did make a negative evaluation about the friend's beliefs: she must not care very much about the moral rule, in this case stealing credit for another's idea, or she would have sided against the promise breaker (in this case, the acquaintance). Hence, in a new third-party scenario, participants made a more negative evaluation about the side-taker's moral behavior when the side-taker remained neutral in a situation where the protagonist (the side-taker's friend) was in the right rather than in the wrong. See Figure 2.5b.

We also examined participant inferences when the friend chose a side depending on whether the protagonist or her opponent was the plagiarizer. When the protagonist was the plagiarizer (i.e. in the wrong), participants inferred that the friend was more likely to plagiarize when the friend sided with the plagiarizer (sided-with protagonist, wrong condition; M = -1.66, SD = 1.45) than against (sided-against protagonist, wrong condition; M = 1.58, SD = 2.04), $t(98) = 9.15$, $p < .001$. Similarly, when the opponent was the plagiarizer (i.e. right condition), participants inferred the friend was more likely to plagiarize when the friend sided with the plagiarizer (sided-against protagonist, right condition; M = -1.00, SD = 1.71) than against the plagiarizer (sided-with protagonist, right condition; M = 1.24, SD = 1.91), $t(99) = 6.18$, $p < .001$. That is, participants judged that plagiarizing would be less likely from someone who sided against a plagiarizer, but not someone who sided with a plagiarizer. Again, this general pattern occurred regardless of whether the participant read that the protagonist or her opponent plagiarized. Figure 5b shows these results; the x-axis is again re-sorted according to whether the friend sided with or against the plagiarizer, as with previous studies.

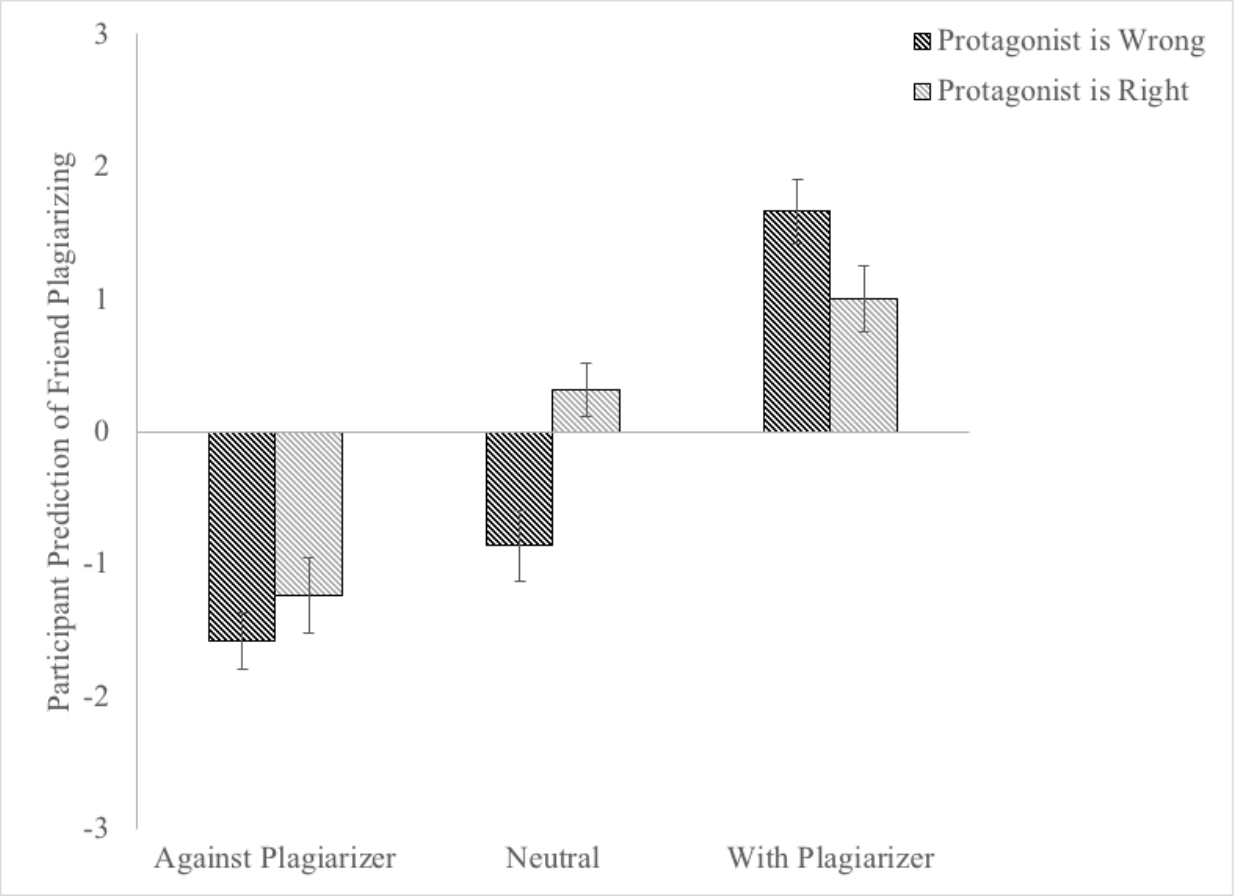

Figure 2.5b. Mean (SE) participants' response to the moral inference measure from Study 2.5 by protagonist's moral standing and friend's siding behavior conditions. Here positive numbers mean the participant thought the side-taker was unlikely to plagiarize and negative numbers mean the participants thought the side-taker was likely to plagiarize.

Discussion

The results of Studies 2.4 and 2.5 replicate and extend our findings from previous studies demonstrating that third parties too make moral inferences about a friend's moral behavior based on both their side-taking decision and on which party was in the wrong. Of course (as in our previous studies) the latter only mattered in the neutrality conditions; when the friend remained neutral third parties inferred that the person would engage in the immoral behavior if their friend was in the right (i.e. the victim of the immoral behavior), but not when their friend was in the wrong (i.e. the perpetrator of the immoral behavior). Indeed, we extended our moral inference

findings from previous studies across two scenarios in which people were evaluating others' inferences from a third-party perspective, finding that people make inferences about others' moral beliefs based on their side-taking behavior in situations where they are not personally invested.

Although we were primarily interested in the third-party ratings for the moral inferences of third parties, we also investigated how third parties predicted that a friend would react to their friend's side-taking decisions. These measures mostly accorded with our previous studies: people thought that a friend would feel less negative about neutrality or being sided against when in the wrong rather than in the right. However, we found a different pattern of results than our previous studies in participants' responses to being sided with and these results may provide some clarity to our previous first party studies and why participants may have responded the way that they did. In those studies, we found that people felt more positive about a friend siding with them when they were in the right rather than when they were in the wrong, which is not entailed by the side-taking hypothesis. One reason that participants may have responded this way in our first party vignettes (in which participants imagined being the wrongdoer) is that they may have experienced some dissonance and did not think of themselves as being the kind of person who would be a wrongdoer and, therefore, thought better of their friend for not siding with them. In line with this possibility, in the third-party vignettes here, we did not find that people thought that one would feel more positive about a friend taking one's side when one was in the right rather than in the wrong (indeed, it was, if anything directionally the opposite). That is, when participants did not need to imagine themselves as being the wrongdoer, we did not find the same results. Future work will be needed to address this question, but these third-party results demonstrate that third parties do indeed make moral inferences based on side-taking behavior

and also are able to make predictions about how friends will feel when others make these decisions.

Study 2.6

One reason the neutrality condition gives such varying results in the previous studies could be that it isn't clear what the friend's intentions are when remaining neutral. Given the lack of information conveyed when the friend remains neutral it could be that people are interpreting this in a variety of way. For example, people may interpret the neutrality to signal condemnation (as these results suggest), or people may interpret neutrality as a desire to avoid the conflict and possibly even reduce any rising tensions. Therefore, in Study 2.6, we used a scenario from previous studies (Shaw et al., 2017) and included an explicit de-escalation condition in which the side-taker conveys that they are not taking sides and would like to contain the escalating conflict. De-escalation would convey a desire to exceed impartiality by reducing the chances of a costly conflict occurring. Therefore, we predict that people will respond more positively to a close friend attempting to de-escalate a conflict than to their failure to show support.

Method

Participants. We recruited 149 participants (66 females, ages ranged from 20 – 70, M = 34.99, SD = 11.02). Participants were assigned to one of three conditions: when the side-taker sided against the participant (n = 47), remained neutral (n = 50), or remained neutral and tried to de-escalate the conflict (n = 52).

Procedure. In a between-participants design, participants read a vignette similar to one used in previous experiments (Shaw et al., 2017). In the vignette, participants were asked to first imagine that they were at a bar with their friend and that they shortly begin talking to an acquaintance. Participants then read that after some time they got into a fight with the

acquaintance. Across conditions, we varied how their friend responded – by either siding against the participant, remaining neutral, or trying to de-escalate the situation. All participants read the following, in brackets are the different conditions with sided against followed by neutrality, followed by neutral de-escalation:

> Imagine that you and your close friend Jamie are out at a bar. The two of you start talking to Casey, a new person you and Jamie recently met.
>
> After sitting in the bar for an hour, you and Casey get into a big argument and eventually start yelling, screaming, and cursing at each other. Finally, Casey says, "What's your problem? Why are you being such a jerk?" You say, "Me? You're the one being a jerk." Then Casey looks at Jamie, "Who's the jerk?"
>
> Jamie looks at both of you, ["You were being the jerk and so you should apologize to Casey/I'm not getting involved guys/Guys, it sounds like you're both pretty upset right now. Maybe we should just forget about it for now and chat about it when you're calmer and not insulting each other."]

Participants then completed measures identical to those in Shaw et al. (2017). Namely, participants rated how close they felt to their friend and whether the relationship was damaged or strengthened on 7-point Likert scales (*1= not at all; 7 = extremely*). These two measures were highly correlated, $r(149) = 0.78$, $p < .001$, and therefore combined into a composite friendship measure (coded from -3 to 3). In line with Shaw et al. (2017), participants also rated the likelihood that they would side with the friend (Jamie) if the friend needed support in a future conflict; ratings were on a scale from "*not at all*" (coded as 0) to "*extremely*" (coded as 100). Finally, participants rated perceived blame by answering who they thought was in the wrong in the argument on a scale from "*Casey*" (coded as 1) to "*You*" (coded as 7) with "*neither you nor Casey*" (coded as 4) as the midpoint. Finally, participants provided brief demographic information.

Results

Friendship Measure. A one-way ANOVA revealed that participants' rating of the friendship differed by condition, $F(2, 148) = 8.36$, $p < .001$, $\eta^2 = .10$. We conducted a series of planned comparisons to test of specific hypotheses. Participants reported a stronger friendship when the side-taker tried to de-escalate the conflict ($M = 0.27$, $SD = 1.03$) than when the side-taker remained neutral ($M = -0.37$, $SD = 0.94$), $t(100) = 3.28$, $p = .001$, or sided against them, ($M = -0.60$, $SD = 1.31$), $t(97) = 3.66$, $p < .001$. Finally, there was no significant difference in participants' friendship ratings when the friend remained neutral or sided against the participant, $t(95) = 0.98$, $p = .330$. See Figure 2.6a for graph including only the remained neutral and neutral de-escalation conditions.

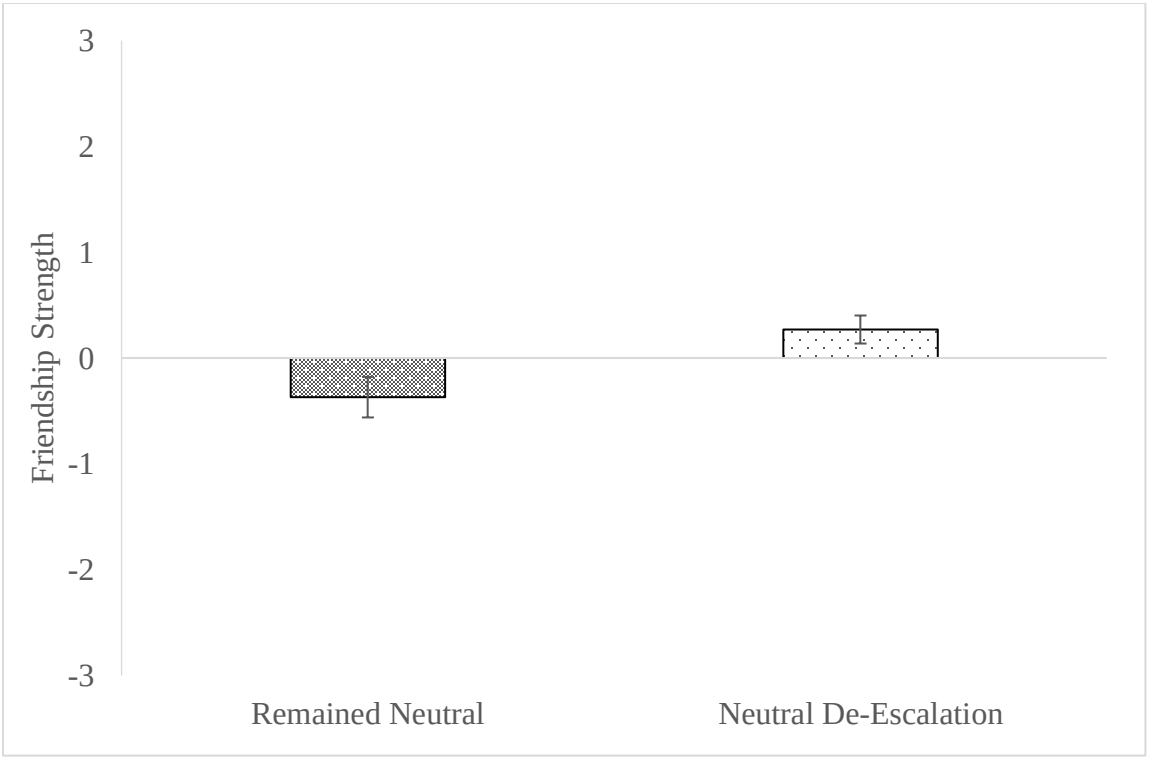

Figure 2.6a. Mean (SE) participants' response on the friendship measure in the remained neutral and neutral de-escalation conditions in Study 2.6. Participants rated the friendship as significantly stronger when the side-taker attempted to de-escalate the conflict, than when the side-taker remained neutral.

Future Support. A one-way ANOVA revealed that participants' rating of their likelihood to support their friend in the future differed by condition, $F(2, 148) = 5.91$, $p = .003$, $\eta^2 = .08$. We again followed up on this with a series of planned comparisons to test of specific hypotheses. Participants were significantly more likely to show support for their friend in the future when their friend had tried to de-escalate the conflict ($M = 59.63$, $SD = 15.25$) than when the side-taker remained neutral ($M = 50.32$, $SD = 24.15$), $t(100) = 2.34$, $p = .021$, or sided against them, ($M = 45.66$, $SD = 21.92$), $t(97) = 3.71$, $p < .001$. Finally, there was no significant difference in participants' likelihood of future support when the friend remained neutral or sided against the participant, $t(95) = 0.99$, $p = .323$.

Perceived Blame. A one-way ANOVA revealed that participants' rating of who was to blame for the conflict did not differ significantly by condition, $F(2, 148) = 2.40$, $p = .094$, $\eta^2 = .03$.

Discussion

Study 2.6 demonstrates that people respond more positively to a friend who attempts to de-escalate a conflict than one that remains neutral or shows opposition. Further, people's reported likelihood of showing future support for their friend was in line with their friendship ratings. That is, people were more likely to report that they'd support their friend in the future if their friend had tried to de-escalate the conflict as compared to if their friend had remained neutral or sided against them. Moreover, in-line with previous research and the findings from Studies 2.1 to 2.5, participants reacted similarly to a friend who remained neutral as to one who sided against them. These data suggest that people do indeed perceive neutrality from a friend as similarly negative as outright opposition, and that attempts at de-escalation from a friend – which isn't a form of side-taking – will lead to a more positive reaction.

General Discussion

In five studies, we found that people make inferences about their friend's loyalty and morality based on how that friend takes sides during a moral conflict and based on who was in the right or wrong during said conflict. Specifically, participants felt less negative toward a friend who failed to provide support when the participant was morally wrong, compared to when they were morally right. Indeed, whereas participants disliked a friend who abandoned them when they were in the right (opposing or staying neutral), participants did not feel this way when they were in the wrong. Further, we demonstrate that people make inferences about their friend's morality based on how the friend reacts to the conflict. If the friend backs the transgressor, people believe that their friend both condones and would be more likely to commit that transgression regardless of who committed it. Interestingly, we find that people also make such moral inferences when their friend remains neutral. However, in response to their friend's neutrality, we find that people's inferences about their friend's morals critically depend on who committed the transgression. Specifically, if the participant (i.e., the friend's friend) committed the moral transgression, then neutrality did not lead to people believe their friend condoned the behavior. However, if the participant was the victim in the conflict, then neutrality from their friend led people to infer that their friend condoned the transgression in question and would be likely to commit that transgression as well. Finally, we also found that participants respond more positively to neutrality with a desire to de-escalate the conflict than when the side-taker only chooses not to show loyalty by remaining neutral.

We demonstrated these results in vignettes involving a bar fight (Study 2.1), breaking promises (Study 2.2 and 2.4), drug use (Study 2.3), and plagiarism (Study 2.5). Further, we found this pattern of results for inferences about the friendship and the friend's morals from both

first- (Study 2.1, 2.2, and 2.3) and third- (Study 2.4 and 2.5) party perspectives, suggesting that, even when people are not personally involved in a conflict, a friend's siding behavior and the moral standing of those involved in the conflict influence people's perceptions about the side-taker's loyalty and morality. These studies demonstrate the intricate ways in which our inferences about loyalty and morality inform the way we interpret the side-taking decisions of allies in morally charged conflicts (for discussion, see DeScioli & Kurzban, 2018). These results may naturally lead one to ask: why are people sometimes more forgiving of a friend's failure to show support in these morally-laden dilemmas?

We argue that people are less upset at their friend for failing to take their side when they are in the wrong because, in these cases, a friend's siding behavior is less diagnostic of that friend's overall loyalty (or lack thereof). When there is no moral conflict, a friend should provide support, and a failure to do so should be fairly diagnostic that this friend is either a bad ally overall (i.e. he or she will not provide support in times of strife) or that the friend does not regard one as a highly-ranked ally. Indeed, previous work has found that people feel quite negative about a friend's neutrality when both sides are equally in the wrong (Shaw et al., 2017). However, when there is a moral conflict in which one party clearly committed a transgression, friends must choose between loyalty and morality; here, a friend's refusal to provide support is less diagnostic of their loyalty. Indeed, people recognize that their friend can be an excellent ally, valuing them very highly and wanting to avoid being on the wrong side of a moral issue. Therefore, a friend showing support when one is in the wrong would of course be highly valuable. However, our data seem to suggest that that such actions may be considered supererogatory based on people's tepid response to the lack of support in such cases—they likely do not expect a friend to take their side in such cases and therefore they are not upset at the lack

of support. Further, we would argue that anything that reduces the extent to which people can make inferences about their friends' loyalty based on a lack of support (e.g., someone remaining neutral due to a motivation to reduce the conflict, a fear for their own or others' safety or livelihood, etc.), will reduce people's negative response to that lack of support (for extended discussion see, Shaw et al., 2017). However, the upshot of these results is that morality appears to provide a particularly effective counterweight to loyalty and people appear to understand the differential pull of both obligations (Everett et al., 2018; Graham et al., 2013; Hughes, 2017; Dungan, Waytz, & Young, 2015; Niemi & Young, 2017; Shaw, Choshen-Hillel, & Caruso, 2018).

The fact that people take sides based on both loyalty and morality can also potentially explain the different moral inferences that people drew about their friends' neutrality. When participants imagined committing a moral transgression and read that their friend remained neutral, we found that people did not infer that their friend endorsed the transgression. It is reasonable that people would not be able to make a strong inference about their friend's morality in this case—the friend is likely conflicted between siding based on morality (which would prompt them to oppose their friend) and loyalty (which would prompt them to support their friend). Therefore, neutrality in this case is not particularly diagnostic of that friend's morality or loyalty. Conversely, when participants imagined being the victim of a moral transgression and their friend remained neutral, the friend's neutrality here could potentially be revelatory about that friend's loyalty and morality. In this case, obligations of both morality and loyalty should lead the friend to provide support and thus, failing to do so may lead to the reasonable inference that the friend may condone and themselves commit that transgression— the friend must not

think the transgression is so bad and may also commit such a transgression, otherwise there would be no reason for the friend to not to show support.

Taken together, the current results provide suggestive evidence for the side-taking function of moral judgment proposed by DeScioli & Kurzban (2013). As noted above, being a good friend means providing help and preferential support to one's allies (Beck & Clark, 1988; Tooby & Cosmides, 1996; DeScioli & Kurzban, 2011). While siding based on alliances is a very sensible strategy, doing so can present a potential problem: getting embroiled in other people's conflicts. Every potential conflict between two parties could blow up into a larger conflict if each of the disputants' friends joined in (for an extended discussion, see DeScioli & Kurzban, 2013). One might think that a defense against this problem is to merely sit out of these potential conflicts, but previous research has demonstrated that neutrality from a friend is sometimes quite costly to the friendship and is treated as a mild form of opposition (Shaw et al., 2017), particularly when the fight is between a close friend and a more distant acquaintance. If alliances were the only strategy by which human beings could take sides in a conflict, then people would be constantly forced into situations where they had to choose between engaging in costly conflict or losing a potential friendship. Morality provides a potential solution to this dilemma; DeScioli and Kurzban (2013) suggest that morality functions as a means for individuals to overcome the costly dilemma of having to become involved in every one of their friends' conflicts. If human beings generally understand that certain kinds of violations (moral rules) may even lead to a lack of support from a deeply loyal friend, then morality may provide a sensible justification for that friend's failure to provide support in some circumstances: "I am still your friend, but I'm siding based on your actions and my morality". Our current results provide some support to this notion.

Understanding the dynamics of how people draw inferences about loyalty and morality based on their own actions and others' side-taking decisions is critical not only in interpersonal relations, but also in politics and international relations (DeScioli & Kurzban, 2009a; Shaw et al., 2017). Since the alliance hypothesis posits that friendships function like alliances – especially in the face of a conflict in the same way that countries are obligated to take an ally's side over a non-ally in a dispute (Liska, 1962), it may be possible to extend the notion of these results to the political arena. Politicians, and nations as a whole, have their own allegiances and are often forced to make decisions about divisive moral issues. Therefore, it seems possible that the dynamics and resulting inferences in play at an interpersonal level can be applied to broader-level relations. On a global level, conflicts are frequently morally valenced and occur between nations with pre-existing alliance commitments. For example, imagine that one country annexed land from another sovereign nation. The response from the allies of the aggressor would likely lead the aggressor – and other nations – to make inferences about these allies' level of support for the aggressing nation and the allies' own policies surrounding sovereignty. If the allies failed to show support to the aggressor, the aggressor may not believe this lack of support to be diagnostic of the strength of their alliances. While neutrality from these allies in response to an aggressive action would likely lead to other nations to infer that the allies condone – and may even commit – such a transgression. How nations respond to immoral acts committed by other nations – allies and otherwise – would not only lead to inferences about alliance strength but also what those taking sides in the conflict are themselves capable of. Relatedly, other researchers have also noted the interesting side-taking dilemmas that people face when they are in a position of being a whistleblower (Waytz et al., 2013) Whistleblowers are caught between their loyalty to their employer or groups (who is committing a transgression that the whistleblower believes the

public should have knowledge of) and their morality (in doing the "right thing"). Given the ubiquity of such conflicts in all realms of social life, this research has the potential to inform a plethora of spheres examining conflict mediation and how different parties perceive not only their opponents' actions, but also the actions of their allies.

In sum, disputants' moral standing in a conflict (i.e., who is in the right and who is in the wrong) and the response from their friends (i.e., to show support or not) lead to varying inferences about the relationships between the involved parties and the morals of those taking sides. When obligations of loyalty and morality collide, people in the wrong are more forgiving of a lack of support from their allies. Imagine that your good friend was caught stealing from work. Our results reveal that the side-taker's dilemma in this situation is not so dire. In this situation, if you fail to support your friend, they are unlikely to think of you as a disloyal companion. Instead, your friend most likely understands that you, like most people, think stealing is wrong regardless of who does it. Your friend will also likely understand that if you do take their side, people may indict your own morality, which should further mitigate the damage done by a lack of support. In short, in some cases, doing what you think is right can also be the right decision for navigating dilemmas involving interpersonal conflict.

CHAPTER 3: TRUE ALLIES: WHEN FRIENDS DO AND DO NOT WANT HONESTY

Honesty is a foundational aspect of moral character and is often considered an important trait in evaluating others' virtue, integrity, and trustworthiness (Anderson, 1968; Goodwin, Piazza, & Rozin, 2014). In line with these ideas, there is a wealth of research demonstrating that people react negatively to being deceived by others (Bok, 1978; Boles, Croson, & Murnighan, 2000; Carr, 1968; Croson, Boles, & Murnighan, 2003; Ford, King, & Hollender, 1988; Greenberg & Wagner 2016; Lewis & Saanri, 1993; O'Connor & Carnavale, 1997; Pillutla & Murnighan, 1996; Robinson, 1996; Santoro & Paine, 1993; Schweitzer & Croson, 1999; Schweitzer, Hershey, & Bradlow, 2006) and that they distrust and dislike liars (Brandts & Charness, 2003; Toris & DePaulo, 1984; Tyler, Feldman, & Reichert, 2006). Thus, in line with the old adage it might seem like honesty is the best policy. However, much of the research that has found strong negative reactions to lying tends to conflate lying with being selfish; when people lie for non-selfish or benevolent reasons, others often think that lying is not only acceptable but sometimes even preferred (Levine & Schweitzer, 2014, 2015). For example, imagine that you just got a new haircut and ran into a casual acquaintance named Mary. Unbeknownst to you, Mary dislikes your haircut; would you feel more negatively toward Mary if she were honest or dishonest in this situation? Even though honesty is often touted as the more moral option, in this case, you would probably appreciate it more if Mary lied to you and said your new haircut looked nice, even if she didn't think so. Indeed, research has found that people often lie out of politeness (Broomfield, Robinson, & Robinson, 2002; Brown & Levenson, 1987; Talwar, Murphy, & Lee, 2007) or with the intention of benefiting others (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Guerrero & Afifi, 1995; Levine & Schweitzer, 2014, 2015; Knobloch & Carpenter-Theune, 2004; Metts, 1989).

Taken together, this previous research suggests a revision to the old adage: *benevolent honesty* is the best policy, but avoid brutal truths that cause unnecessary harm (Levine, 2015). This seems reasonable but, when honesty and benevolence clash, do people have a preference for one over the other? Below I review past research on people's reactions to honesty and benevolent dishonesty.

Preferences for honesty and benevolence

As noted above, people's reactions to honesty is strongly calibrated based on how benevolent or harsh they find the honest statement to be. Therefore, if friends as opposed to strangers or acquaintances give us a piece of harsh criticism, is it likely that we will construe it as more harsh or more benevolent? Deception rooted in benevolence, or with the desire to shield others from unnecessary harm and has found that, in some cases, deception not only increase trust, but to also be better received than honesty (Levine, 2015; Levine & Schweitzer, 2014, 2015) and omission (Levine, Hart, Moore, Rubin, Yadav, & Halpern, 2018). In a series of studies utilizing various economic games, Levine & Schweitzer (2015) found that certain types of lies, broadly those of a prosocial nature, can in fact increase one's trust in the liar. In these experiments, participants who told prosocial lies (i.e. lied about the outcome of a coin-flip in order to earn money for their partner) were perceived to be more ethical and were trusted more than those who were honest and harmed their partner. Hence, prosocial lies involve the transmission of information that misleads and benefits the target (Levine & Schweitzer, 2014). These results suggest that people place more importance on the intention to help or harm than they do on following moral rules such as "never lie". Consistent with this finding, a large body of research suggests that people place great importance on motive attributions and intentions when judging other people (Weber, Malhotra, & Murnighan, 2004; Malhotra & Murnighan,

2002; Pillutla, Malhotra, & Murnighan, 2003; Dunn, Ruedy, & Schweitzer, 2012; Lount & Pettit, 2012). Given the importance people place on others' intentions and motivations, dishonesty intended to be helpful or rooted in benevolence may be more acceptable (than critical honesty that may be interpreted to hold malevolent intentions) in countless everyday interactions.

In fact, adults lie in approximately 20% of their daily interactions (DePaulo & Bell, 1996), and most of these lies are of a prosocial nature (DePaulo & Kashy, 1998). Even in the workplace, research finds that the desire to be benevolent and compassionate leads individuals to be more dishonest with colleagues (Lupoli, Jampol, & Oveis, 2017). Similarly, in interpersonal relationships people will avoid being honest at times to prevent hurt feelings. For example, if they dislike their friend's romantic partner, people are less likely to convey such information as they believe it may harm the friendship (Mayer, 1957). Therefore, it appears that prosocial lies occur frequently in all types of relationships and scenarios and may be less damaging to the relationship than a harsh truth. However, are there instances when we might accept and even expect those around us to be completely honest with us, even though it may hurt our feelings? And is there a difference in how we interpret harsh truths from friends vs. non-close others (i.e. acquaintances)? Here we attempt to answer this question by exploring how people respond to brutal honesty in the context of close relationships. In the haircut example above, had Mary been a close friend, your preference for honesty versus benevolent dishonesty may have been different. Might critical truths be more acceptable, and therefore less harmful, from close others? A series of studies investigating how people react to critical truths (e.g., hearing that their new haircut is bad) from acquaintances versus friends suggests this is indeed the case. Past work on friendship and lying provides the logic for this prediction.

Friendships

Friendships are a crucial aspect of life and are integral to healthy social functioning and successful navigation in our social world. Friendships are often indispensable and provide countless benefits including social and emotional support (Argyle & Henderson, 1984; Fleming & Baum, 1986; Lang & Carstensen, 1994; Tukuno, 1983). This social and emotional support can range from taking one's side during a conflict (DeScioli & Kurzban, 2009a; Shaw et al., 2017) to listening to one's problems and providing feedback in crucial times (Cohen & Hoberman, 1983; Cohen & Willis, 1985; Jones, 1973). We befriend others based on a variety of reasons including core characteristics that we value in interpersonal relationships. One such characteristic is honesty (Altman & Taylor, 1973; LaFollette & Graham, 1986; Pearce 1974; Weiselquist, Rusbult, Foster, Agnew, 1999). Honesty is often a highly valued trait that we not only seek to uphold but also desire from others. Indeed, Kant argued that, since friendship is a moral duty, dishonesty in friendships both signals and contributes to mistrust, weakening, and possibly even destruction of the friendship (Kant & Infield, 1963). Moreover, honesty often signals greater intimacy. Indeed, honesty is a component of self-disclosure and self-disclosure among interaction partners has been shown to lead to feelings of greater closeness (Derlega, Metts, Petronio, & Margulis, 1993; Greene, Derlega, & Mathews, 2006; Sedikides, Campbell, Reeder, & Elliot, 1999; Sprecher & Treger, 2013).

However, as reviewed above, sometimes people avoid being honest in order to behave more benevolently and avoid causing harm (Levine, 2015; Lupoli, Jampol, & Oveis, 2017). Benevolence is another core characteristic that we seek in friends. From a young age, people expect greater kindness from closer friends (Clark & Bittle, 1992), and our friends are often the chosen recipients of our benevolence (Telfer, 1970). Moreover, benevolence has been described as the core characteristic of close friends (Kapur, 1991). We are undoubtedly kinder to and about

our friends than strangers and an awareness of this bias – for greater benevolence towards one's friends – is even used to evaluate and discount information that people convey about their friends (Brandt, Vonk, & Knippenberg, 2011; Johnston, Mills, & Landrum, 2015; Sedikides, Hoorens, & Dufner, 2015). Given that people expect benevolence from their friends, and if benevolence is such an important part of people's reaction to honesty, then one could imagine that people would react more negatively to brutal honesty from a friend because they expect friends to be particularly benevolent toward them. However, since honesty signals a greater closeness among friends (versus acquaintances), it could be that a critical truth may be interpreted harshly from an acquaintance, while being perceived as less harsh from a friend. If we expect kindness from friends, we may interpret their intentions to be more benevolent than an acquaintance's in the case of a critical truth. Therefore, it is important to know if people prefer their friends to provide honest feedback that may be critical or if, even from close friends, people prefer dishonesty rooted in benevolence, i.e. prosocial lies.

Honesty vs. benevolence within friendships

Although honesty from one's friends is crucial (Altman & Taylor, 1973; LaFollette & Graham, 1986; Pearce 1974; Weiselquist, Rusbult, Foster, Agnew, 1999), previous research has also demonstrated the benefits of prosocial lying in friendships (Levine & Schweitzer, 2015; Lupoli, Jampol, & Oveis, 2017; Mayer, 1957). This previous work has also demonstrated that once benevolence is established, honest and dishonest statements are evaluated more equally in terms of the relationship (Levine & Schweitzer, 2015; Studies 3a and 3b). In these studies, the authors found that deception itself has no effect on benevolence-based trust and, instead, it was the person's intentions that influenced people's evaluation of the relationship: selfish intentions harmed the relationship while altruistic intentions, whether they were associated with honesty or

deception, strengthened the relationship. Moreover, previous work has found that people are more likely to ascribe benevolent intentions to their friends' behavior (Berndt & McCandless, 2009), which may lead to people being more tolerant of brutal honesty from friends than acquaintances. That is, people may not judge critical truths from friends as negatively because they do not attribute bad intent to their friends. Given the increased familiarity that one has with their friends, people may feel more certain about being able to more easily infer their friends' (good) intentions than those of an acquaintance, who would not be assumed to have benevolent intentions.

Indeed, behaviors of those who have not yet established a friendship (i.e. acquaintances) can often be ambiguous and uncertain (Kelley & Thiabaut, 1978). Such feelings of uncertainty and ambiguity are typically experienced as aversive (Brickman, 1987; Holmes, 1991) and this negative feeling may lead people to attribute malevolence to critical honesty from distant others while making the same critical statement from a close friend appear more benevolent. Relatedly, the criticalness of an honest statement is therefore likely to be construed as less damaging to a close friendship than an acquaintanceship. Moreover, people have different expectations of how interactions and social exchanges will play out with friends versus acquaintances (Clark, 1984a). Given that people often lie in everyday situations, e.g., to colleagues, it is likely that an acquaintance providing feedback considered a critical truth may be seen as a violation of the expectations of the acquaintanceship, which people also find aversive and damaging to the relationship (Argyle & Henderson, 1984; Clark & Mills, 1979). Further, given the increased intimacy among close friends versus acquaintances, people are more likely to share honest statements with close friends (LaFollette & Graham, 1986). Therefore, if honesty is more common among close friends than acquaintances, it is likely that critical truths occur more often

in friendships as well and are less damaging to that relationship as they are more expected from close friends than acquaintances (Clark, 1984a). Indeed, close relationship partners have been found to tell everyday lies at a lower rate than strangers (DePaulo & Kashy, 1998).

Therefore, we predict that while prosocial lies are preferred to critical truths in distant relationships, this is less likely to be the case in friendships. If people seek benevolence regardless of the relationship, then one might expect that people will respond more positively to a prosocial lie from both close friends and acquaintances. However, if people ascribe more benevolent intentions to their friends, we would expect that people are more accepting of critical truths from close friends than acquaintances. It is also possible that both could be the case; that while prosocial lies are preferred, people find critical truths less damaging from a close friend. Therefore, we predicted that people would feel more positive towards acquaintances who told prosocial lies rather than critical truths, consistent with existing research (Levine & Schweitzer, 2014, 2015), but that this difference would be attenuated or eliminated when the feedback came from a close friend: prosocial lies may be preferred from close friends as well but a critical truth coming from a close friend would be more acceptable than a critical truth coming from an acquaintance. That is, we predicted an interaction between friendship and truth-telling. I test this hypothesis across three vignette-style studies in which people imagine receiving feedback from a friend or an acquaintance. The feedback is in regards to a low-stakes decision (e.g., a new haircut) and is either a prosocial lie or a critical truth.

Study 3.1

In Study 1, we investigated how participants felt toward either a close friend or an acquaintance who told a prosocial lie or a critical truth. Given that people focus on others' motives and intentions, if benevolence is more important when receiving feedback, then we

would expect that prosocial lies are more acceptable from both friends and acquaintances. However, if people ascribe greater benevolence to their friends than their acquaintances (Brandt, Vonk, & Knippenberg, 2011; Johnston, Mills, & Landrum, 2015; Sedikides, Hoorens, & Dufner, 2015), then critical honesty may be more acceptable, and therefore less damaging to the relationship, when coming from a friend versus an acquaintance.

Method

Participants. For all studies, we aimed to recruit ~50 subjects per cell and did not analyze the data for any study until data collection was completed. No participants were excluded across studies due to the simple nature of the tasks. Further, we did not collect demographic information beyond age and sex of the participant as we did not make a priori predictions about individual differences.

In Study 1, we recruited two-hundred participants (108 females, M = 37.41; SD = 13.13) via Amazon Mechanical Turk (AMT) using the TurkPrime.com interface (Litman, Robinson, & Abberbock, 2016). Participation was restricted to residents of the United States and subjects were compensated $0.50 each for completing the 3-minute study.

Procedure. In a 2 (relationship: close friend, acquaintance) x 2 (communication type: critical truth, prosocial lie) between-subjects design, participants were randomly assigned to read one of four vignettes that were programmed through Qualtrics survey software. In the vignette, participants were asked to first imagine a close friend or an acquaintance. Participants in the acquaintance condition read:

> Please think of someone you consider to be an acquaintance. This person should be someone you do not spend a lot of time with and do not think of as a close friend.
>
> Please enter the NAME of the person you are thinking about

While participants in the close friend condition read the following:

> Please think of someone you consider to be a really good friend. This person should be someone you spend a lot of time with and think of as a close friend.
>
> Please enter the NAME of the person you are thinking about:

All participants then entered a person's name in accordance with one of the two above prompts and selected the appropriate pronouns. We then used the piped-text function in Qualtrics to insert the person's name and associated pronouns throughout the vignette. All participants read:

> Imagine that you are having lunch one day with [Name]. The two of you are chatting about your respective weekends. During the conversation, you mention to [Name] that you got a haircut over the weekend.

Immediately following this, participants in the critical truth condition read:

> [Name] doesn't particularly like your haircut and tells you this. "It's just not very flattering", [s/he] says.

While participants in the prosocial lie condition read:

> [Name] tells you that your new hairstyle looks great, even though [s/he] doesn't particularly like it and think it's unflattering.

After reading the vignette, participants answered nine questions intended to examine how they felt about the person based on the scenario. All questions were presented on a 7-point Likert scale with the mid-point 4 always labelled as "*neutral*". The exact wording of the measures can be found in the supplemental materials.

We measured how close participants felt to the person (1 = *extremely distant*; 7 = *extremely close*); and the extent to which participants desired the friendship (1 = *not at all*; 7 = *a great deal*). These two measures were highly correlated, (200) = .79, $p < .001$, and thus combined into a composite friendship measure.

We also measured perceptions of benevolence using four items (report alpha here): the extent to which participants believed that the person had the participant's best interests in mind (1 = *not at all*; 7 = *a great deal*), was trying to protect them (1 = *not at all*; 7 = *a great deal*), was trying to hurt them (1 = *not at all*; 7 = *a great deal)*, reverse-scored, and was trying to help them (1 = *not at all*; 7 = *a great deal*).

We also measured the extent to which participants valued the person's response (1 = *not at all*; 7 = *a great deal*); and participants' belief of the appropriateness of the person's statement (1 = *extremely inappropriate*; 7 = *extremely appropriate*). As a manipulation check, we measured the extent to which participants believed the person was being honest with them (1 = *not at all*; 7 = *a great deal*); our manipulation check worked such that participants rated the person as being significantly more honest in the honest condition (M = 5.72, SD = 1.30) than the prosocial lie condition (M = 3.79, SD = 2.03), $t(198) = 8.02$, $p < .001$. Finally, we collected basic demographic information.

Results

First, we looked at the overall pattern of means with 2 (relationship: acquaintance, close friend) x 2 (feedback: prosocial lie, critical truth) between-subjects ANOVAs on all dependent measures.

Friendship. In line with our predictions, the ANOVA revealed a significant relationship x feedback interaction, $F(1, 196) = 4.91$, $p = .028$, $\eta^2 = .03$: when the feedback came from an acquaintance, participants reported a stronger friendship when the acquaintance told a prosocial lie (M = 3.99, SD = 1.39) than a critical truth (M = 3.32, SD = 1.44), $t(98) = 2.38$, $p = .020$, but when the feedback came from a close friend, there was no statistical difference in friendship strength whether the close friend told a prosocial lie (M = 5.53, SD = 1.12) or a critical truth (M

= 5.68, *SD* = 1.27), *t*(98) = 0.61, *p* = .542. That is, friendship remained strong when a close friend shared a critical truth as compared to a prosocial lie, but not when an acquaintance did the same. We also found a main effect of relationship, *F*(1, 196) = 111.16, *p* < .001, η^2 = .36, such that after receiving feedback, participants reported a stronger relationship with a close friend (*M* = 5.60, *SD* = 1.19) than an acquaintance (*M* = 3.65, *SD* = 1.42). We did not find a main effect of feedback type, *F*(1, 196) = 2.04, *p* = .155, η^2 = .01. See Figure 3.1a.

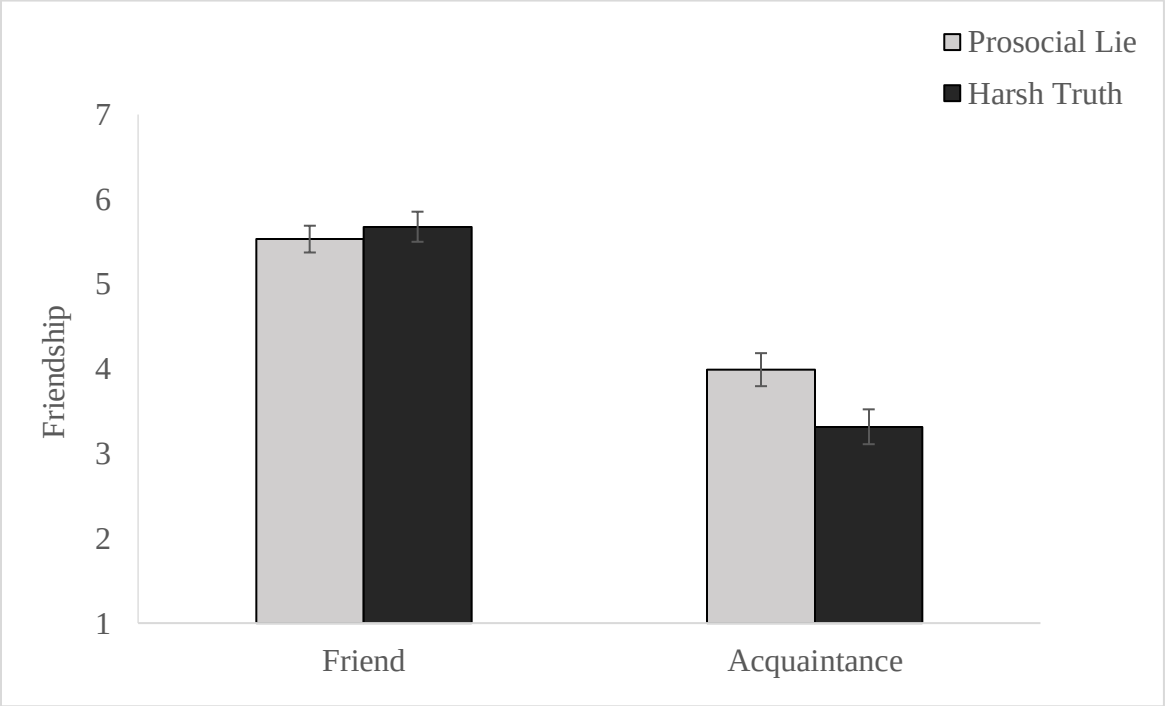

Figure 3.1a. Mean (SE) participants' rating of the friendship by condition in Study 3.1. Participants reported a stronger friendship with a close friend than an acquaintance and with an acquaintance when they told a prosocial lie rather than a critical truth.

Perceptions of Benevolence. The ANOVA revealed no relationship x feedback interaction, *F*(1, 196) = 0.89, *p* = .347, η^2 = .00. However, in line with our predictions, there was a significant main effect of relationship, *F*(1, 196) = 75.98, *p* < .001, η^2 = .22, such that participants perceived greater benevolent intentions from a close friend (*M* = 5.59, *SD* = 0.62) than an acquaintance (*M* = 4.35, *SD* = 1.24). Finally, we also found a main effect of feedback,

$F(1, 196) = 8.27$, $p = .004$, $\eta^2 = .04$, with participants perceiving greater benevolent intentions when told a prosocial lie ($M = 5.21$, $SD = 1.00$) than a critical truth ($M = 4.73$, $SD = 1.30$). See Figure 3.1b.

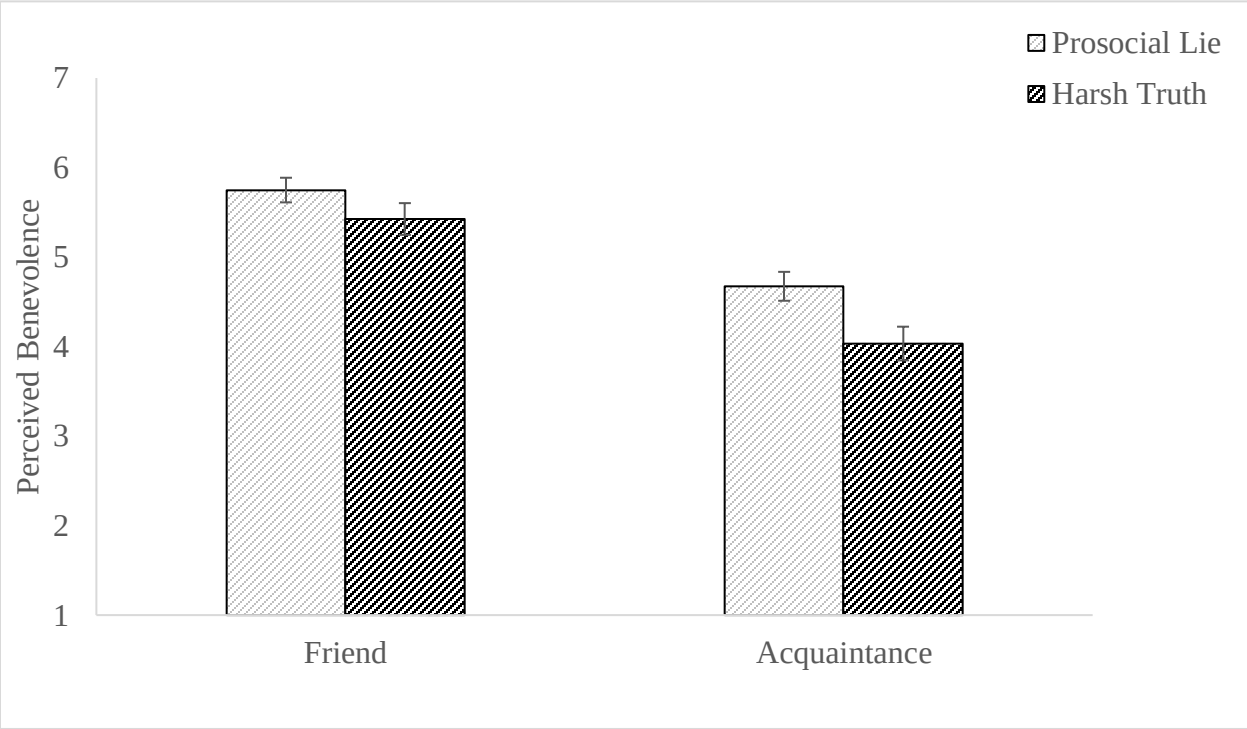

Figure 3.1b. Mean (SE) participants' rating of perceived benevolence of feedback provider. Participants perceived greater benevolence from a close friend than an acquaintance, and greater benevolence when told a prosocial lie versus a critical truth.

Help. The ANOVA revealed no relationship x feedback interaction, $F(1, 196) = 0.23$, $p = .632$, $\eta^2 = .00$. There was a significant main effect of relationship, $F(1, 196) = 36.34$, $p < .001$, $\eta^2 = .16$, such that participants believed a close friend was trying to help them ($M = 5.34$, $SD = 1.48$) more than an acquaintance ($M = 4.05$, $SD = 1.53$). We also found a significant effect of feedback, $F(1, 196) = 1.78$, $p = .184$, $\eta^2 = .01$, with participants finding a critical truth ($M = 4.84$, $SD = 1.45$) more helpful than a prosocial lie ($M = 4.55$, $SD = 1.57$).

Valued Response. The ANOVA revealed no relationship x feedback interaction, $F(1, 196) = 0.28$, $p = .598$, $\eta^2 = .00$. There was a significant main effect of relationship, $F(1, 196) = 46.96$, $p < .001$, $\eta^2 = .19$, such that participants valued the response more when it came from a close friend ($M = 5.39$, $SD = 1.46$) than an acquaintance ($M = 3.91$, $SD = 1.60$). Finally, there was no effect of feedback, $F(1, 196) = 0.02$, $p = .891$, $\eta^2 = .00$.

Appropriateness. The ANOVA revealed no relationship x feedback interaction, $F(1, 196) = 0.32$, $p = .571$, $\eta^2 = .00$. There was a significant main effect of relationship, $F(1, 196) = 38.34$, $p < .001$, $\eta^2 = .16$, such that participants rated the response as more appropriate when it came from a close friend ($M = 5.03$, $SD = 1.43$) than an acquaintance ($M = 3.78$, $SD = 1.40$). There was also a significant main effect of feedback, $F(1, 196) = 11.29$, $p = .001$, $\eta^2 = .05$, such that participants rated a prosocial lie as more appropriate ($M = 4.74$, $SD = 1.27$) than a critical truth ($M = 4.07$, $SD = 1.56$).

Discussion

Consistent with our hypothesis, when the feedback came from a close friend, participants reported feeling as strong of a friendship with a close friend regardless of whether their close friend relayed a critical truth or a prosocial lie. However, when the feedback came from an acquaintance, participants reported a stronger friendship when the acquaintance told a prosocial lie versus a critical truth. That is, critical truth was more acceptable from a close friend than an acquaintance. Further, participants perceived greater benevolent intentions when receiving any type of feedback from a close friend than an acquaintance. Study 3.1 provides initial evidence that people perceive greater benevolent intentions from close friends than acquaintances and that they are more accepting of critical truths from close friends than acquaintances. Moreover, people have a greater knowledge of their close friend's intentions than they do of an

acquaintance's intentions. As such, people are more likely to attribute benevolent intentions to a close friend than an acquaintance. Although people generally have greater knowledge of their friends than their acquaintances, it could be that knowledge about the subject topic of the feedback influences how people perceive the feedback they receive. For example, if an acquaintance had expert knowledge about the topic of the feedback – e.g., if the acquaintance was an experienced hair stylist, would a critical truth still be more damaging to the friendship and perceived as less benevolent? We explored this in Study 3.2.

Study 3.2

In Study 3.2, we examined how participants feel about critical feedback from a close friend or an acquaintance who is either an expert or a novice about the topic area of the feedback. Here we predicted that people would still perceive more benevolent intentions from close friends than acquaintances, but that they would also perceive more benevolent intentions from an expert than from a novice. Research shows that, in a consumer setting, people favor the advice of experts, over those offering benevolent advice, when it concerns a decision with low emotional difficulty (Barnett White, 2005). Although advice is arguably different than feedback, and a consumer setting may elicit different responses than an interpersonal one, receiving feedback on a new haircut is rather low-stakes, and therefore, it could be argued that critical feedback is more acceptable when a person has expertise knowledge on the topic than when they are a novice. That is, if the expertise of the feedback provider is important, then it is likely that a critical truth is less damaging, especially from an acquaintance, since their feedback may reflect their knowledge rather than their intentions. However, if people do not consider expertise as a valid reason for critical but honest feedback, then they are likely to respond negatively to critical truth.

In order to examine this, we only look at how people react to critical but honest feedback in Study 3.2.

Method

Participants. In Study 2, we recruited two-hundred and two participants (85 females, M = 34.73; SD = 11.86) via AMT using the TurkPrime.com interface (Litman, Robinson, & Abberbock, 2016). Participants were assigned to one of four conditions: when the person was a close friend and an expert (n = 51), or a novice (n = 51); when the person was an acquaintance and an expert (n = 49) or a novice (n = 51). Participation was restricted to residents of the United States and subjects were compensated $0.50 each for completing the 3-minute study. For Study 2 onwards, we preregistered all aspects of the studies on aspredicted.org; the pre-registration of Study 2 can be found at: https://aspredicted.org/qz6nd.pdf.

Procedure. In a 2 (relationship: close friend, acquaintance) x 2 (expertise: expert, novice) between-subjects design, participants were randomly assigned to read one of four vignettes presented via Qualtrics. In Study 2, instead of participants imagining a real friend or acquaintance a in Study 1, they were asked to imagine being friends or acquaintances with a fictitious person named Jordan. Participants read the scenarios described below. The relationship conditions are bracketed with close friend followed by acquaintance and the expertise conditions in parentheses with only the expert condition included as the novice condition simply omitted that sentence:

> Imagine that you have a(/n) [good friend/acquaintance] named Jordan. You and Jordan [have been good friends for quite some time and often hang out in your free time/met recently through mutual friends and hang out rather infrequently. You don't dislike each other, you just have not known each other for very long].
>
> One evening, you bump into Jordan while out and the two of you start catching up about things that happened during the week. During the conversation Jordan asks you, "Did you get a new haircut?" Happy that it was noticed, you reply, "Yes! I

> really needed a new look and the stylist said this style is really popular right now. What do you think?"
>
> (You know Jordan went to cosmetology school and has years of experience as a stylist, so you expect that Jordan will recognize the cut).
>
> Jordan responds, "Hmm, it looks kind of old fashioned. That cut might have been trendy a few years ago, but it doesn't seem all that stylish."

After reading the vignette, participants answered eight questions intended to examine how they felt about the person. Most of the questions were identical to those in Study 1. Closeness and desire for continuing the friendship were highly correlated again, $r(202) = .83$, $p < .001$, and therefore combined into a general friendship measure. We calculated perceived benevolence with a composite score of: perceived best interests, extent to which the participant believed the person was trying to protect them; and the hurt measure, reverse-scored. All three measures were again highly correlated (see supplemental materials). We removed the help item from our benevolence scale in Study 2 as it was highly correlated with other measures examining the perceived benevolence of the feedback provider.

We included a measure examining how useful participants perceived the feedback to be (*1 = not at all useful; 7 = extremely useful*). We also measured how honest people believed the feedback provider was being (*1 = not at all honest; 7 = extremely honest).* Finally, we included a manipulation check at the end of the study which asked participants "*how knowledgeable Jordan is about hair"* on a 7-point Likert scale (*1 = not at all knowledgeable; 7 = extremely knowledgeable*. Consistent with the intent of the manipulation, participants rated an expert as having greater knowledge ($M = 5.46$, $SD = 1.31$) than a novice ($M = 3.91$, $SD = 1.72$), $t(200) = 7.20$, $p < .001$. Participants also provided brief demographic information.

Results

To analyze our data, we conducted 2 (relationship: acquaintance, close friend) x 2 (expertise: expert, novice) between-subjects ANOVAs on all dependent measures.

Friendship. The ANOVA revealed a significant main effect of relationship, $F(1, 198) = 65.39$, $p < .001$, $\eta^2 = .25$: after receiving critical feedback, participants reported a stronger relationship with a close friend ($M = 5.15$, $SD = 1.13$) than an acquaintance ($M = 3.55$, $SD = 1.54$). We also found a main effect of expertise, $F(1, 198) = 6.86$, $p = .010$, $\eta^2 = .03$: participants reporting a stronger friendship after receiving critical feedback from an expert ($M = 4.63$, $SD = 1.26$) than a novice ($M = 4.10$, $SD = 1.58$). The ANOVA revealed no relationship x expertise interaction, $F(1, 198) = 0.83$, $p = .361$, $\eta^2 = .00$. See Figure 3.2a.

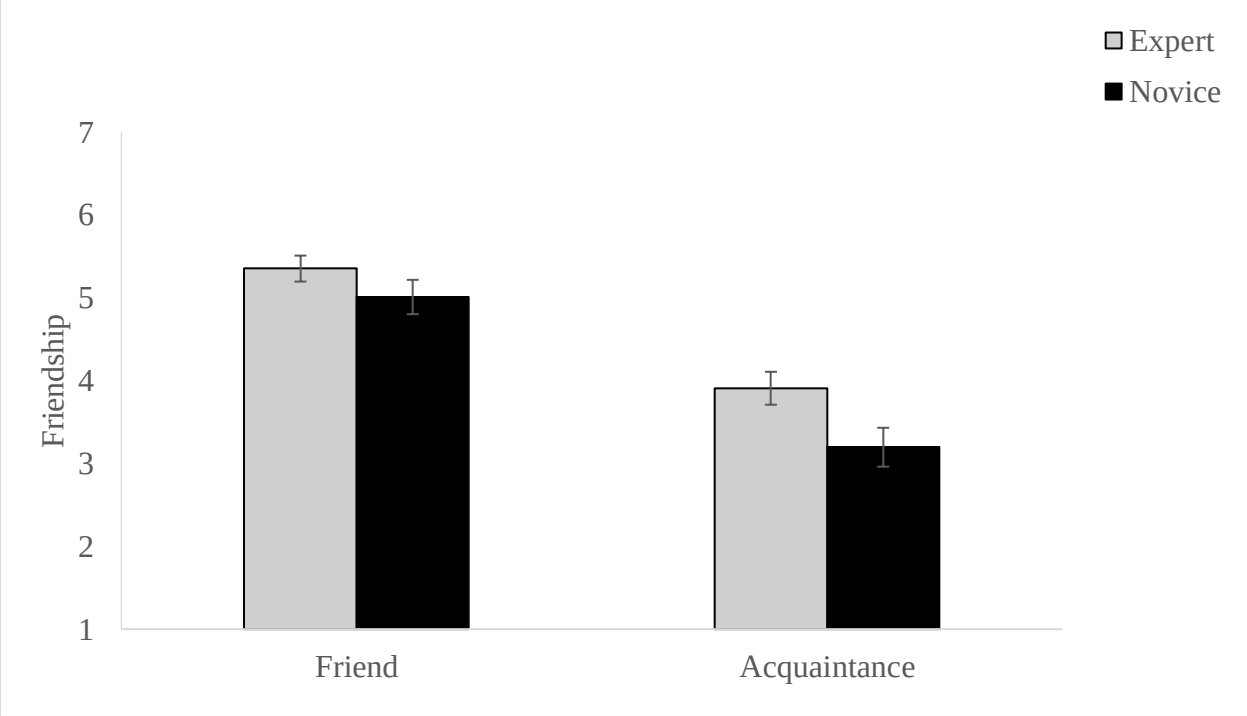

Figure 3.2a. Mean (SE) participants' friendship ratings. Participants reported a stronger friendship with a close friend than an acquaintance, and when the person providing the feedback was an expert versus a novice.

Perceptions of Benevolence. The ANOVA revealed a significant main effect of relationship, $F(1, 198) = 22.62$, $p < .001$, $\eta^2 = .10$, with participants perceiving greater benevolent intentions from a friend ($M = 4.79$, $SD = 1.31$) than an acquaintance ($M = 3.96$, $SD = 1.18$). We also found a significant main effect of expertise, $F(1, 198) = 9.97$, $p = .002$, $\eta^2 = .05$, such that participants perceived greater benevolence in critical feedback from an expert ($M = 4.65$, $SD = 1.18$) than a novice ($M = 4.10$, $SD = 1.32$). Finally, there was no relationship x expertise interaction, $F(1, 198) = 1.24$, $p = .266$, $\eta^2 = .00$. See Figure 3.2b.

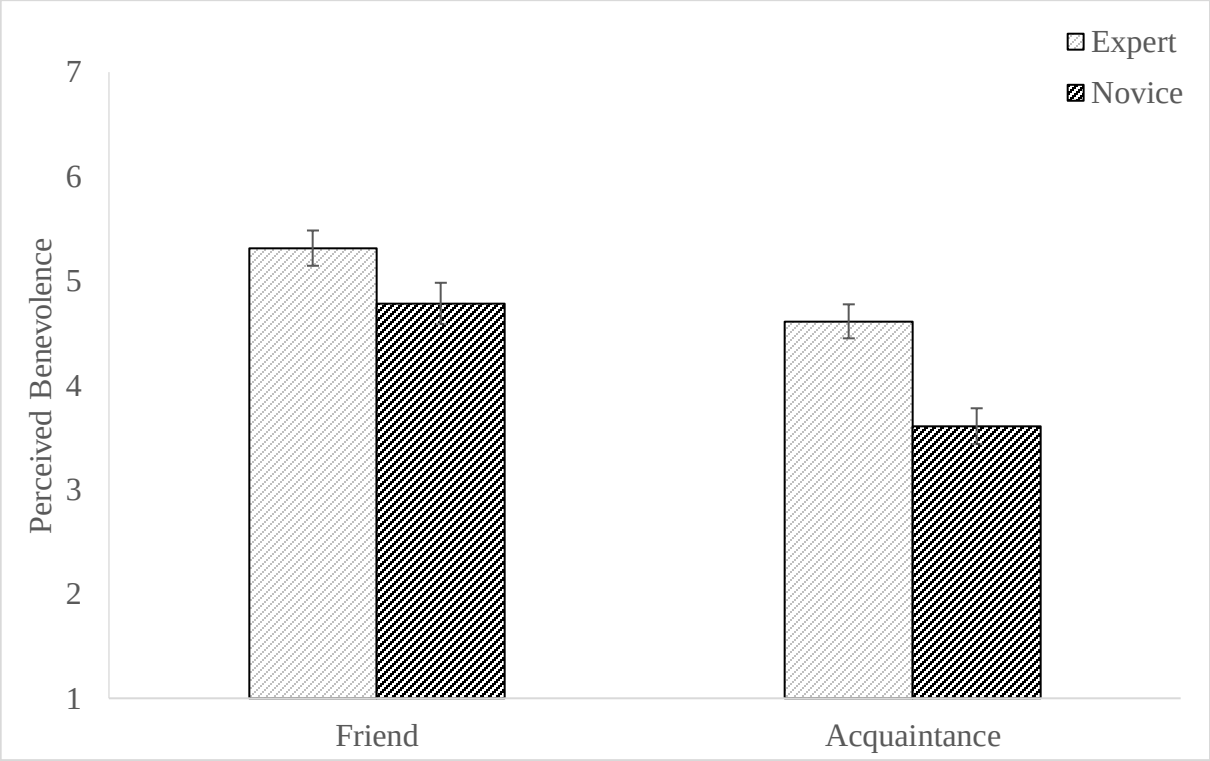

Figure 3.2b. Mean (SE) participants' rating of perceived benevolence. Participants perceived greater benevolence from a close friend than an acquaintance and from an expert than a novice.

Usefulness. There was a significant main effect of relationship, $F(1, 198) = 16.55$, $p < .001$, $\eta^2 = .08$, such that participants thought the response was more useful when it came from a close friend ($M = 4.88$, $SD = 1.52$) than an acquaintance ($M = 3.95$, $SD = 1.72$). There was also a main effect of expertise, $F(1, 198) = 15.20$, $p < .001$, $\eta^2 = .07$, with participants reporting the

critical feedback to be more useful when coming from an expert (M = 4.86, SD = 1.52) than a novice (M = 3.97, SD = 1.72) Finally, the ANOVA revealed no relationship x expertise interaction, $F(1, 198) = 0.97$, $p = .325$, $\eta^2 = .00$.

Appropriateness. The ANOVA revealed a significant main effect of relationship, $F(1, 198) = 14.07$, $p < .001$, $\eta^2 = .07$, such that participants rated the critical feedback as more appropriate when it came from a close friend (M = 4.58, SD = 1.68) than an acquaintance (M = 3.68, SD = 1.72). There was also a significant main effect of expertise, $F(1, 198) = 11.41$, $p = .001$, $\eta^2 = .05$, such that participants thought the critical feedback was more appropriate coming from an expert (M = 4.53, SD = 1.66) than a novice (M = 3.73, SD = 1.74). Finally, there was no relationship x expertise interaction, $F(1, 198) = 1.36$, $p = .245$, $\eta^2 = .00$.

Honesty. The ANOVA revealed a significant main effect of relationship, $F(1, 198) = 10.73$, $p = .001$, $\eta^2 = .05$, with critical feedback rated as more honest when coming from a close friend (M = 5.93, SD = 1.14) than an acquaintance (M = 5.37, SD = 1.28). We also found a significant main effect of expertise, $F(1, 198) = 8.82$, $p = .003$, $\eta^2 = .04$, with people rating the critical feedback as more honest when communicated by an expert (M = 5.91, SD = 1.01) than a novice (M = 5.39, SD = 1.41). There was no significant relationship x expertise interaction, $F(1, 198) = 0.13$, $p = .721$, $\eta^2 = .00$.

Discussion

Consistent with our hypotheses and previous results, Study 3.2 provides more evidence that people find critical feedback as less damaging to the relationship, and therefore more acceptable, when coming from a close friend than an acquaintance. Moreover, people again inferred close friends to have greater benevolent intentions than acquaintances.

In this study, we also find that critical feedback was less damaging to the relationship when coming from an expert than a novice. People also perceived greater benevolent intentions when receiving critical feedback from an expert as compared to a novice. This suggests that expertise in a specific area reduces people's perceptions of ill-intent when receiving critical feedback from others. Expertise in a specific area most likely leads people to believe that the feedback provided, whether critical or not, is rooted more in the expert knowledge than any ill-intent, and therefore, people are less likely to perceive non-benevolent intentions on the part of an expert.

Nevertheless, Study 3.2 also showed that people still perceive less damage to the relationship when a close friend delivers a critical truth versus an acquaintance. And that the relationship was less damaged when receiving a critical truth from an expert versus a novice does so. Of course, given that people already know their close friends better than their acquaintances, and that their close friends presumably also know them better than do their acquaintances, they most likely assume greater benevolent intent on the part of their close friends. What would happen if someone received critical feedback from a friend who knew them quite well, but this friend had non-benevolent intentions?

Study 3.3

Our previous study demonstrated that people emphasize friendship more than expertise when interpreting critical feedback. We argue that we found this difference because we assume benevolent intentions with friends. However, we may respond more favorably to critical criticism from friends for a number of reasons, including that they have had many more interactions with us as well as the fact that they may have a better idea about our personal preferences. In order to control for these factors, in Study 3 we examine if people respond less

favorably to critical truths from a friend who they have reason to believe no longer holds benevolent intentions. Here, the friend still holds expert knowledge about one's preferences and is more familiar than an acquaintance but in this case one cannot assume that the closeness of the relationship will allow one to predict the intentions of the person providing the critical feedback. Thus, if the intentions of the feedback provider are more important, then people should be more accepting of critical feedback with clear benevolent intentions from both close friends and acquaintances. However, if the relationship with the feedback provider is more important, (because the closeness of the relationship acts as a proxy for the expertise, about one's friend's preferences and tastes, accumulated during the history of the friendship) then people should still be more accepting of critical feedback from a close friend as compared to an acquaintance regardless of that person's intentions.

Method

Participants. We recruited two hundred participants (85 females, M = 38.67; SD = 11.40) in a manner identical to previous studies. Participants were assigned to one of four conditions: when the person was a close friend and had benevolent intentions (n = 50), or non-benevolent intentions (n = 49); when the person was an acquaintance and had benevolent (n = 50) or non-benevolent intentions (n = 51). The pre-registration for Study 3 can be found here: https://aspredicted.org/ue4pm.pdf.

Procedure. In a 2 (relationship: close friend, acquaintance) x 2 (intention: benevolent, non-benevolent) between-subjects design, participants were randomly assigned to read one of four vignettes presented via Qualtrics. As with Study 3.2, participants were asked to imagine a friend or acquaintance and read the following. The relationship conditions are in brackets with

close friend followed by acquaintance and the intention conditions are in parentheses with benevolent followed by non-benevolent:

> Imagine that you have a friend named Jordan. You have been friends with Jordan for [10 years and Jordan knows you really well because you have known each other for a long time/a few weeks, but Jordan does not know you particularly well because you have only known each other for a short time].
>
> Recently, you and Jordan (have been getting along well. As usual, Jordan has been pretty kind towards you/had a falling out and have since been having friction. You've noticed that Jordan has been pretty unkind towards you lately).
>
> One evening, you run into Jordan and the two of you start talking. During the conversation Jordan asks you if you recently got a new haircut. Happy it was noticed, you reply, "Yes! I really wanted a new look. What do you think?"
>
> Jordan replies, "Um, it's really short. I don't really think it flatters your face."

After reading the scenario, participants answered eight questions regarding their feelings toward Jordan; all were presented on a 7-point Likert scale. We again asked participants to rate how close they felt to Jordan, whether Jordan had their best interests in mind, whether Jordan was trying to protect them, hurt them, and how appropriate Jordan's response was. We also asked participants to rate their commitment to a friendship with Jordan (*1 = much less committed, 7 = much more committed*), whether the relationship had been damaged or strengthened (*1 = extremely damaged, 4 = neither damaged nor strengthened, 7 = extremely strengthened*). For the general friendship measure, we created a composite of the closeness, commitment, and relationship damage measures, as these three were also highly correlated. Perceptions of benevolence was again measured as a composite score of three measures: participants' evaluations of whether Jordan had their best interests in mind, whether Jordan was trying to protect them, and the hurt measure, reverse-scored, as all three were again highly correlated (see supplemental materials). We also asked participants how well they thought

Jordan knew their personal preferences (*1 = not at all well, 7 = extremely well*) as a manipulation check for the relationship. This manipulation check was successful as participants rated a friend as having greater knowledge of their personal preferences ((*M* = 5.17, *SD* = 1.52) than an acquaintance (*M* = 2.88, *SD* = 1.49), $t(198) = 10.78$, $p < .00$. Finally, participants provided brief demographic information.

Results

We again first looked at the overall pattern of means with 2 (relationship: acquaintance, close friend) x 2 (expertise: expert, novice) between-subjects ANOVAs on all dependent measures.

Friendship. The ANOVA revealed a significant effect of relationship, $F(1, 196) = 5.36$, $p = .022$, $\eta^2 = .03$, such that after receiving critical feedback, participants reported a stronger relationship with a friend (*M* = 3.79, *SD* = 1.26) than an acquaintance (*M* = 3.38, *SD* = 1.24). We also found a main effect of intention, $F(1, 196) = 71.20$, $p < .001$, $\eta^2 = .27$, with participants reporting a stronger friendship after receiving critical feedback from a person with benevolent (*M* = 4.32, *SD* = 1.19) versus non-benevolent (*M* = 2.85, *SD* = 1.28) intentions. Therefore, even if the person is a close friend, if they have non-benevolent intentions, participants perceived a negative effect on the friendship after receiving critical feedback. This shows that even if the person has prior knowledge of the participant's personal preferences (by being a close friend) their intentions are more important when evaluating the impact of the critical feedback on the friendship. Finally, there was no relationship x intention interaction, $F(1, 196) = 1.84$, $p = .177$, $\eta^2 = .00$. See Figure 3.3a.

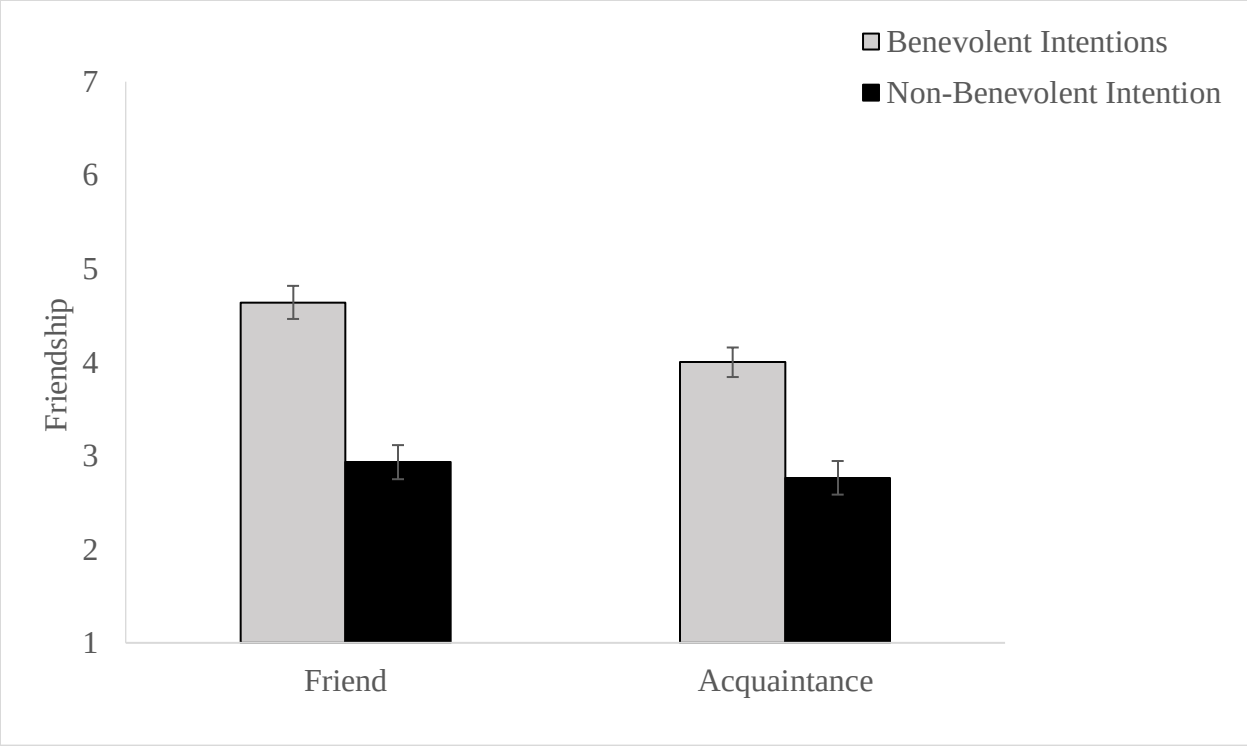

Figure 3.3a. Mean (SE) participants' friendship ratings. Participants reported a stronger friendship with a close friend than an acquaintance, and when the person providing the feedback had benevolent versus non-benevolent intentions.

Perceptions of Benevolence. The ANOVA revealed a significant main effect of relationship, $F(1, 196) = 8.92$, $p = .003$, $\eta^2 = .04$, with participants perceiving greater benevolent intentions from a friend ($M = 3.91$, $SD = 0.80$) than an acquaintance ($M = 3.56$, $SD = 0.83$). We also found a significant main effect of intention, $F(1, 196) = 26.54$, $p < .001$, $\eta^2 = .12$, such that participants perceived greater benevolence when the person had benevolent intentions ($M = 5.33$, $SD = 0.83$) as compared to when they did not ($M = 5.33$, $SD = 0.80$). Also, there was a significant relationship x intention interaction, $F(1, 196) = 4.12$, $p = .044$, $\eta^2 = .02$: when the person had benevolent intentions, participants rated greater benevolence from a close friend ($M = 5.33$, $SD = 1.27$) than an acquaintance ($M = 4.82$, $SD = 1.02$), $t(98) = 2.23$, $p = .028$; however, when the person had non-benevolent intentions, there was no statistical difference in

participants' perceptions of benevolence; that is, participants perceived low benevolence in both a close friend (M = 2.96, SD = 1.56) and an acquaintance (M = 3.00, SD = 1.46) in this case, $t(98) = 0.14$, $p = .893$. See Figure 3.3b.

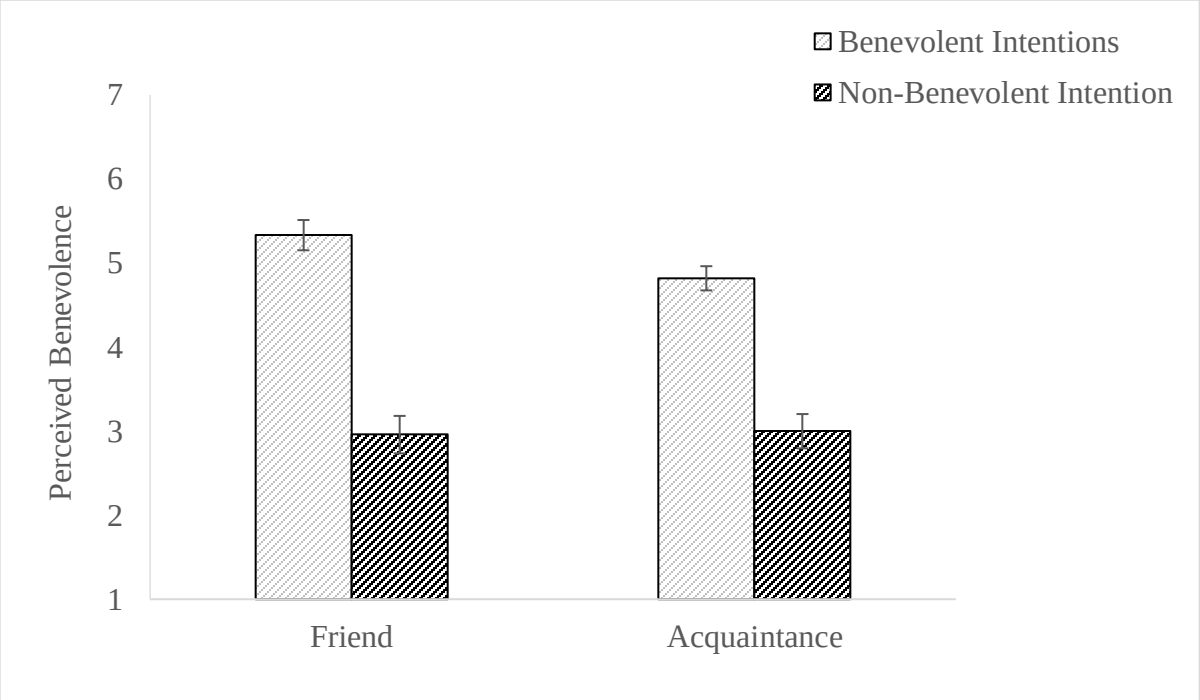

Figure 3.3b. Mean (SE) participants' rating of perceived benevolence. Participants perceived greater benevolence from a close friend than an acquaintance and when the feedback provider had benevolent versus non-benevolent intentions.

Appropriateness. The ANOVA revealed no effect of relationship, $F(1, 196) = 2.13$, $p = .146$, $\eta^2 = .01$. There was a significant main effect of intention, $F(1, 196) = 35.04$, $p < .001$, $\eta^2 = .15$, such that participants thought the critical feedback was more appropriate coming from a person with benevolent (M = 4.48, SD = 1.66) than non-benevolent intentions (M = 3.13, SD = 1.57). That is, critical feedback is more appropriate when the person has benevolent intentions and their previous knowledge of the participant (i.e. whether they were a close friend or acquaintance) did not affect perceived appropriateness. Finally, there was no relationship x intention interaction, $F(1, 196) = 1.38$, $p = .242$, $\eta^2 = .00$.

Discussion

The results of Study 3 show that when critical feedback is rooted in benevolent intentions, it is perceived more positively and is less likely to damage the friendship. People still infer greater benevolent intentions from close friends delivering critical feedback than acquaintances, but when non-benevolent intentions are clear, the relationship one has with the person (providing the critical feedback) is less important. Moreover, intentions were more important than the relationship when people considered the appropriateness of the critical feedback. That is, regardless of whether a friend or acquaintance gave critical feedback, if they possessed non-benevolent intentions, participants believed the critical feedback was similarly inappropriate. These data suggest that people care a great deal about a person's intentions when reacting to critical feedback and that, sometimes, intentions are more important than one's closeness (and the familiarity and prior personal knowledge that this brings) with the person providing the critical feedback.

General Discussion

Across all three studies, participants responded more positively to receiving feedback from a close friend than an acquaintance. Specifically, our results suggest that while prosocial lies are acceptable from both friends and acquaintances, participants perceived less damage to the relationship when receiving a critical truth from a friend than an acquaintance. Moreover, participants' perception of the benevolent intentions of the feedback provider were in line with their perceptions of the effect of the feedback on the relationship. That is, participants perceived greater benevolence when the feedback provider: told a prosocial lie, was a close friend, an expert, or a person with benevolent intentions. Taken together, these results suggest that although honesty may not always be the best policy, people are more accepting of critical truths from

close friends, those that they infer to hold benevolent intentions, or when they have information that suggests the critical truth is not a personal attack (i.e. the person is an expert).

These findings extend existing knowledge about how people react to prosocial lies versus critical truths when receiving feedback and the types of inferences people make in regards to the relationship and the person's intentions. Previous research has found that people are accepting of dishonesty when it is rooted in benevolence (Levine & Schweitzer, 2015) and this appears to hold true in interpersonal relationships as well. While honest but hurtful feedback can be potentially damaging to relationships (Vangelisti, 1994), these data suggest that honesty is more acceptable from close friends (as compared to acquaintances). This occurs for two reasons. First, infer benevolent intentions from close friends delivering critical truths. This is likely because people have greater personal knowledge about their close friends, and therefore, hold greater knowledge of their friends' intentions, which they believe to be more prosocial (Brendt, 1981). Indeed, friendships are characterized by help, intimacy, emotional security, and self-validation (Furman, 1996; Mendelson & Aboud, 1999; Sullivan, 1953; Weiss, 1974). Therefore, when a friend conveys a critical truth, it is less likely that people infer malevolent intent. It is likely that critical truths are more acceptable as the friendship progresses and people become closer. When one is in the early stages of a relationship, it is more common to be polite (Brown & Levenson, 1987) and prosocial lies may be a better strategy for a variety of reasons. Firstly, given that we have less personal knowledge about an acquaintance than a close friend, we are less likely to know if a critical truth will be interpreted as offensive. Second, given the limited personal knowledge we have about an acquaintance, we may not be aware of their personal preferences as we might be with a close friend. For instance, even if we dislike their new haircut, they may really like it and feel confident and happy with the choice that they made. As people become

closer friends, they tend to self-disclose more. Since self-disclosure is an example of honesty, this greater self-disclosure often increases feelings of intimacy (Chelune, Robinson, & Kommor, 1984; Greene, Derlega, & Mathews, 2006; Harvey & Omarzu, 1997; Laurenceau, Feldman-Barrett, & Pietromonaco, 1998; Reis & Shaver, 1988). Increased intimacy, honest self-disclosure, and the benevolence bias that we hold in relation to our friends all play a role in increasing the acceptability of honest but critical feedback. While it may not be acceptable for an acquaintance to tell us that we look bad on a night out, it may not only be more acceptable, but also expected, of a close friend to share such an opinion. Further, we put more weight on a friend's opinion than that of an acquaintance's (Berry & Shaw, 2017), and therefore, honesty may also be more expected from a friend than an acquaintance so it would be more acceptable when received. However, one may argue that people are more forgiving of a friend's violations versus an acquaintance's in general. Since the relationship with a friend is more established and known to continue, perceived to hold more benevolence, and more familiar in general, people may forgive their friends more easily. Although these studies did not test this possibility directly, previous research has found that people actually react more angrily at a friend for certain violations, such as failing to provide help (Sell, Sznycer, Al-Shawaf, Lim, Krauss, Feldman, Rascanu, Sugiyama, Cosmides, & Tooby, 2017).

Further, from these results, it also appears that people are more accepting of a critical truth when they have knowledge that the person conveying it has some expertise in the specific subject area. This greater acceptance of a critical truth from an expert may be indicative of greater trust placed on an expert's opinion, and therefore, greater inference of benevolent intent. Expertise is often linked to trustworthiness and credibility (Hovland, Janis, & Kelley, 1953), and research suggests that a person appears more credible if they exhibit a willingness to honestly

share their expert knowledge (Tseng & Fogg, 1999). Hence, critical feedback from an expert may be less hurtful (i.e. perceived as more benevolent) if the receiver believes there is some credibility behind what was shared. Moreover, the criticalness of the feedback may make the receiver feel it is more likely truth rooted in expertise (i.e. credible) than an insensitive statement. Further, such feedback from an expert may even make the critical truth feel less like a personal attack and more about the feedback provider's "objective knowledge". If expertise increases perceptions of benevolent intent, it would be interesting to see if this translates to credibility and veracity of the statement conveyed.

Interestingly, our results also suggest that explicit knowledge of the person's intentions is more impactful than the existing relationship. That is, when one is aware that the feedback provider's intentions may no longer be benevolent, the history of the relationship becomes less important than the person's intentions. While it seems intuitive that non-benevolent intentions would make critical truths less acceptable, it is noteworthy that one's history in terms of the friendship no longer holds the same weight as when the person has benevolent intentions. This suggests that a person's intentions have a greater effect on how critical feedback is interpreted as compared to a person's relationships. However, it is unclear whether intentions outweigh the relationship only for a specific interaction or over time also. It could be that the relationship is more important compared to an aggregate of intentions. Also, this may be true for more permanent relationships, such as kin or other close, long-term bond and therefore it would be beneficial to examine the juxtaposition of intentions and the relationship with feedback over time.

Limitations and future directions

While these results shed light on how people perceive critical feedback from friends versus acquaintances, they may be somewhat limited in their generalizability. First, all three experiments were vignette-style studies in which participants reacted to an imagined scenario with a real (Study 3.1) or ostensible (Study 3.2 and 3.3) friend or acquaintance sharing feedback. It could be argued that the way in which people respond to an imagined scenario might differ from their reactions to a real scenario. While this can certainly be the case, research suggests the use of vignettes can more valid and reliable measures of respondent opinion than other methods (Alexander & Becker, 1978). Moreover, it would be experimentally difficult to control the variance that would have been undoubtedly introduced by real friends or acquaintances providing feedback in comparable scenarios. Future research could attempt to create an environment in which the feedback provided is more controlled in order to test this in real-time interactions.

Another limitation is the scope of the subject matter participants received feedback on. Given that the feedback was in regards to a new haircut, participants may have been largely more accepting of prosocial lies from both friends and acquaintances because a new haircut is a low-stakes decision. It is possible that if the feedback was in relation to a higher-stakes decision or a more intimate topic, e.g., whether a participant's significant other did something that could be interpreted as detrimental to the relationship, then participants may have desired a critical truth from their close friends more than a prosocial lie. Future research should explore how people respond to critical truths and prosocial lies in response to subjects of varying levels of severity or intimacy. Further, it is possible that our results are specific to a friendship domain, future research should aim to explore if intentions hold more weight than the prior relationship for other types of social connections (e.g., family) as well.

Conclusion

We expect our friends to provide us with feedback and validation on most aspects of our lives. While we often desire honesty in our close friends, as a trait and important moral characteristic, the results of these studies suggest that, in low-stakes situations, people often prefer to be told prosocial lies by others instead. However, honesty appears to be more acceptable from close friends than acquaintances. We infer greater benevolent intentions from close friends than acquaintances, suggesting that the greater level of intimacy and familiarity that exists between close friends discounts perceptions of mal-intent when receiving critical truths. More broadly, these results highlight the important role intent plays in how people interpret the information they receive from others and how they use this information to inform the relationship.

GENERAL CONCLUSION

Humans have evolved with a need to be social. An important part of being social is initiating and maintaining close friendships. Friendships are vital to our learning (Allen, 1981; Gottman, Gonso, & Rasmussen, 1975; Schonert-Reichl, 1999), our happiness (Berscheid, Snyder, & Omoto, 1989) and our survival (Cohen, Gottlieb, & Underwood, 2004; Cole, Capitanio, Chun, Arevalo, Ma, & Cacioppo, 2015; DeScioli & Kurzban, 2009a; Holt-Lunstad, Smith, & Baker, Harris, & Stephenson, 2015; Holwerda, Deeg, Beekman, Tilburg, Stek, Jonker, & Schoevers, 2012; Yang, Boen, Gerken, Li, Schorpp, & Harris, 2016). Friends are often described as the family we choose. Therefore, it is important to study what people want in a friend, how people interact with their friends, and how people interpret their friend's actions.

To that end, in the first chapter of the current dissertation I examined how people reacted to a friend being generous and helpful. I found that, people experience friendship jealousy when their friend is generous or helpful to another friend (but not a family member or romantic partner). The results from Chapter 1 suggest that while people do want friends who are cooperative, kind, and generous (Baumard, André, & Sperber, 2013; Kenny, Mohr, & Levesque, 2001; Nelson, 2002; Peck, 1993; Rand & Nowak, 2013), they especially want their friends to be cooperative, kind, and generous to them above other friends (DeScioli & Kurzban, 2009a; Tooby & Cosmides, 1996). That is, while people care about the objective generosity of their friends, they pay special attention to how their friend benefits them in relation to other friends. I suggest that people respond negatively to their friend being prosocial toward other friends because they perceive that prosociality as a threat to their friendship. In Chapter 2, I examined the inferences people make when, after transgressing or being transgressed again in a conflict with an acquaintance, their friend reacts to the conflict. People appear to make inferences about both the

friendship and their friend's own morals. When responsible for the conflict, people infer less damage to the friendship and are less upset with their friend for failing to take their side. This demonstrates that people are less likely to expect their friends to be loyal if another, sufficient obligation (e.g., the friend's morality) appears to be the reason behind the lack of loyalty. People also infer that their friend would be more likely to commit the same transgression if the friend remains neutral during the conflict. Finally, Chapter 3 demonstrates that while people prefer prosocial lies to critical truths from both close friends and acquaintances, critical truths from a close friend are more acceptable, deemed more benevolent, and less damaging to the friendship. While honesty is a virtue, and we want our friends to be virtuous and moral, deception rooted in benevolence is often preferred. Therefore, while we want friends who are ideally generous, loyal, and honest, sometimes, we prefer our friends to be less generous, less loyal, and dishonest.

References

Alexander, C. S., & Becker, H. J. (1978). The use of vignettes in survey research. *Public opinion quarterly*, *42*, 93-104.

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*, 272.

Andreoni, J., & Bernheim, D. B. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77, 1607–1636.

Argyle, M., & Henderson, M. (1984). The rules of friendship. *Journal of social and personal relationships*, *1*, 211-237.

Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution & Human Behavior*, *27*, 344–360

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior*, *34*, 164–175.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33–38.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 749–753.

Barnett White, T. (2005). Consumer trust and advice acceptance: The moderating roles of benevolence, expertise, and negative emotions. *Journal of Consumer Psychology*, *15*, 141-148.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36,* 59–78.

Berndt, T. J., & McCandless, M. A. (2009). Methods for investigating children's relationships with friends. *Handbook of peer interactions, relationships, and groups*, 63-81.

Berry, Z., & Shaw, A. (2017, May). Epistemic vigilance in a social context: Recognizing when allegiances distort the truth. Poster presentation at the 29thAPS Annual Convention. Boston, MA.

Bevan, J. L., & Samter, W. (2004). Toward a broader conceptualization of jealousy in close relationships: Two exploratory studies. *Communication Studies*, *55*, 14–28.

Bloom, P. (2013). Just babies: The origins of good and evil. New York, NY, US: Crown Publishers/Random House.

Bok, S. (1978). *Lying: Moral choices in public and private life.* New York, NY: Pantheon.

Boles, T. L., Croson, R. T., & Murnighan, J. K. (2000). Deception and retribution in repeated ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, *83*, 235-259.

Brandt, A. C., Vonk, R., & van Knippenberg, A. (2011). Augmentation and discounting in impressions of targets described by third parties with ulterior motives. *Social Cognition*, *29*, 210-220.

Brandts, J., & Charness, G. (2003). Truth or consequences: An experiment. *Management Science*, *49*, 116-130.

Broomfield, K. A., Robinson, E. J., & Robinson, W. P. (2002). Children's understanding about white lies. *British Journal of Developmental Psychology, 20*, 47-65.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage.* New York, NY: Cambridge University Press.

Carr, A. (1968). Is business bluffing ethical? *Harvard Business Review, 46,* 143-155.

Choshen-Hillel, S., Shaw, A., & Caruso, E. M. (2015). Waste management: How reducing partiality can promote efficient resource allocation. *Journal of Personality and Social Psychology, 109,* 210-231.

Clark, M. S. (1984a). A distinction between two types of relationships and its implications for development. In J. C. Masters & K. Yarkin- Levin (Eds.), *Boundary areas in social and developmental psychology* (pp. 241-270). New York: Academic Press.

Clark, M. L., & Bittle, M. L. (1992). Friendship expectations and the evaluation of present friendships in middle childhood and early adolescence. *Child Study Journal*.

Clark, M. S. & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology, 37,* 12–24.

Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, *64,* 201-229

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13*, 99-125.

Cohen, S., & Willis, T. (1985). Stress and Social Support and the Buffering Hypothesis. *Psychological Bulletin*, 98, 310-357.

Croson, R., Boles, T., & Murnighan, J. K. (2003). Cheap talk in bargaining experiments: Lying and threats in ultimatum games. *Journal of Economic Behavior & Organization, 51,* 143-159.

Delton, A. W., & Krasnow, M. M. (2017). The psychology of deterrence explains why group membership matters for third-party punishment. *Evolution and Human Behavior, 49*, 1027-1033.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*, 13335-13340.

Delton, A. W., & Robertson, T. E. (2016). How the mind makes welfare tradeoffs: Evolution, computation, and emotion. *Current Opinion in Psychology*, *7*, 12-16.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: lies are told to those who care. *Journal of personality and social psychology*, *71*, 703.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, *74*, 63.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*, 979.

Derlega, V. J., Metts, S., Petronio, S., & Margulis, S. T. (1993). Sage series on close relationships. *Selfdisclosure. Thousand Oaks, CA, US*.

DeScioli, P. (2016). The side-taking hypothesis for moral judgment. *Current Opinion in Psychology, 7*, 23-27.

DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science*, *22*(4), 442-446.

DeScioli, P., & Kurzban, R. (2009a). The alliance hypothesis for human friendship. *PLoS ONE, 4*, 1-8.

DeScioli, P., & Kurzban, R. (2009b). Mysteries of morality. *Cognition*, *112*, 281–299.

DeScioli, P., & Kurzban, R. (2011). The company you keep: Friendship decisions from a functional perspective. In J.I. Krueger (Ed.), *Social judgment and decision making* (pp. 281–299). New York, NY: Psychology Press.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*, 477–496.

DeScioli, P., Kurzban, R., Koch, E. N., & Liben-Nowell, D. (2011). Best friends: Alliances, friend ranking, and the Myspace social network. *Perspectives on Psychological Science*, 6, 6–8.

Dungan, J., Waytz, A., & Young, L. (2015). The psychology of whistleblowing. *Current Opinion in Psychology*, *6*, 129-133.

Dunn, J., Ruedy, N. E., & Schweitzer, M. E. (2012). It hurts both ways: How social comparisons harm affective and cognitive trust. *Organizational Behavior and Human Decision Processes*, *117*(1), 2-14.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Fleming, R., & Baum, A. (1986). Social support and stress: The buffering effects of friendship. In *Friendship and social interaction* (pp. 207-226). Springer, New York, NY.

Fiske, A. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*, 689–89.

Ford, C. V, King, B.H., & Hollender, M.H. (1988). Lies and liars: Psychiatric aspects of prevarication. *American Journal of Psychiatry*, *145*, 554-562.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*, 148.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55-130.

Graham, G., & LaFollette, H. (1986). Honesty and intimacy. *Journal of Social and Personal Relationships, 3*, 3-18.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. P*sychological Inquiry, 23*, 101-124.

Greenberg, A. E., & Wagner, A. F. (2016). The Ripple Effects of Deceptive Reporting. *Swiss Finance Institute Research Paper*, 16-45.

Greene, K., Derlega, V. J., & Mathews, A. (2006). Self-disclosure in personal relationships. *The Cambridge Handbook of Personal Relationships*, 409-427.

Guerrero, L. K., & Afifi, W. A. (1995). Some things are better left unsaid: Topic avoidance in family relationships. *Communication Quarterly, 43*, 276-296.

Gurven, M. & Winking, J. (2008). Collective action in action: Prosocial behavior in and out of the laboratory. *American Anthropologist, 110*, 179–90.

Hackman, J., Danvers, A., & Hruschka, D. J. (2015). Closeness is enough for friends, but not mates or kin: Mate and kinship premiums in India and U.S. *Evolution and Human Behavior, 36*, 137-145.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98-116.

Hess, N. H., & Hagen, E. H. (2006). Psychological adaptations for assessing gossip veracity. *Human Nature, 17*, 337–354.

Hovland, C. I., I. L. Janis, and H. H. Kelley. (1953). Communication and Persuasion: Psychological Studies of Opinion Change. London: Oxford University Press.

Hruschka, D., Hackman, J., & Macfarlan, S. (2015). Why do humans help their friends? Proximal and ultimate hypotheses from evolutionary theory. In V. Zeigler-Hill, L. M. Welling, T. K. Shackelford (Eds.), *Evolutionary perspectives on social psychology* (pp. 255-266). Cham, Switzerland: Springer International Publishing.

Hughes, J. S. (2017). In a moral dilemma, choose the one you love: Impartial actors are seen as less moral than partial ones. *British Journal of Social Psychology, 56,* 561-577.

Johnston, A. M., Mills, C. M., & Landrum, A. R. (2015). How do children weigh competence and benevolence when deciding whom to trust? *Cognition, 144,* 76-90.

Jones, S. C. (1973). Self-and interpersonal evaluations: esteem theories versus consistency theories. *Psychological Bulletin, 79,* 185.

Jordan, J. J., Sommers, R., Bloom, P., & Rand, D. G. (2017). Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychological Science, 28,* 356-368.

Kant, I. (1785). Groundwork for the metaphysics of morals.

Kant, I., & Gregor, M. J. (1998). *Groundwork of the metaphysics of morals*. Cambridge, U.K: Cambridge University Press.

Kant, I. & Infield, L. Trans. (1963). Immanuel Kant: Lectures on Ethics. New York: Harper & Row.

Kapur, N. B. (1991). Why it is wrong to be always guided by the best: consequentialism and friendship. *Ethics, 101,* 483-504.

Keller, S. (2007). The Limits of Loyalty. Cambridge: Cambridge University Press.

Kennedy, J., & Schweitzer, M. E. (2015). Holding People Responsible for Ethical Violations: The Surprising Benefits of Accusing Others. *Academy of Management Proceedings*, *1*, 11258-11258.

Kenny, D. A., Mohr, C. D., & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin*, *127*, 128–141.

Kohlberg, L. (1981). *The philosophy of moral development: Moral stages and the idea of justice*. San Francisco: Harper & Row.

Kleiman-Weiner, M., Shaw, A., & Tenenbaum, J. B. (2017). Constructing social preferences from anticipated judgments: When impartial inequity is fair and why? *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 676-681). London, UK: Cognitive Science Society.

Knobloch, L. K., & Carpenter-Theune, K. E. (2004). Topic avoidance in developing romantic relationships: Associations with intimacy and relational uncertainty. *Communication Research*, *31*, 173-205.

Kraft, C., & Mayeux, L. (2016). Associations among friendship jealousy, peer status, and relational aggression in early adolescence. *The Journal of Early Adolescence*, *36*, 1–23.

Krems, J. A., Kenrick, D. T., & Neel, R. (2017). Individual perceptions of Self-actualization: What functional motives are linked to fulfilling one's full potential? *Personality and Social Psychology Bulletin*, *43*, 1337-1352.

Krems, J. A., Williams, K. E. G., Kenrick, D. T., & Aktipis, A. (2017, February). Friendship jealousy: A functional emotion for protecting friendships. In *Evolutionary Preconference.* Data blitz conducted *at 2017 Society for Personality and Social Psychology Annual Meeting*, San Antonio, TX.

Lang, F. R., & Carstensen, L. L. (1994). Close emotional relationships in late life: Further support for proactive aging in the social domain. Psychology and Aging, 9, 315-324

LaFollette, H., & Graham, G. (1986). Honesty and intimacy. *Journal of Social and Personal Relationships*, *3*, 3-18.

Lederach, J. P. (2005). *The moral imagination: The art and soul of building peace*. Oxford University Press.

Levine, E.E. (2015). Lie to others as you would have others lie unto you: community standards of deception. In *Academy of Management Proceedings*. (Vol. 2015, No. 1, p. 10533). Briarcliff Manor, NY 10510: Academy of Management.

Levine, E. E., & Cohen, T. R. (2018). You can handle the truth: Mispredicting the consequences of honest communication. *Journal of Experimental Psychology: General*, *147*, 1400.

Levine, E., Hart, J., Moore, K., Rubin, E., Yadav, K., & Halpern, S. (2018). The surprising costs of silence: Asymmetric preferences for prosocial lies of commission and omission. *Journal of Personality and Social Psychology*, *114*, 29.

Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology, 53*, 107-117.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes, 126*, 88-106.

Lewis, M. & Saarni, C. (1993). *Lying and deception in everyday life.* New York, NY: The Guilford Press.

Liberman, Z., & Shaw, A. (2017). Children use partial resource sharing as a cue to friendship. *Journal of Experimental Child Psychology*, *159*, 96–109.

Liska, G. (1962). *Nations in alliance: The limits of interdependence.* Baltimore, MD: Johns Hopkins Press.

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*, 433-442.

Lount Jr, R. & Pettit, N. (2012). The social context of trust: The role of status. *Organizational Behavior and Human Decision Processes*, *117*, 15-23.

Lukaszewski, A. W., Simmons, Z. L., Anderson, C., & Roney, J. R. (2016). The role of physical formidability in human social status allocation. *Journal of Personality and Social Psychology*, *110*, 385-406.

Lupoli, M. J., Jampol, L., & Oveis, C. (2017). Lying because we care: Compassion increases prosocial lying. *Journal of Experimental Psychology: General*, *146*, 1026.

Malhotra, D., & Murnighan, J. K. (2002). The effects of contracts on interpersonal trust. *Administrative Science Quarterly*, 534-559.

Mayer, J. E. (1957). The self-restraint of friends: A mechanism in family transition. *Social Forces*, 230-238.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal relationships*, *6*, 159-179.

Morgan, G. S. (2008). *Political Orientation, Values, and Attributions: The Motivated Correction Model of Ideological Reasoning* (Doctoral dissertation, University of Illinois at Chicago).

Myers, J. L. & Well, A. D. (1995). *Research Design and Statistical Analysis*, Lawrence Erlbaum, Hillsdale, New Jersey.

Nelson, J. (2002). *Building partnerships: Cooperation between the United Nations system and the private sector*. United Nations Publications, New York, NY.

Nezlek, J. B. (1993). The stability of social interaction. *Journal of Personality and Social Psychology, 65*, 930-941.

Niemi, L., & Young, L. (2017). Who sees what as fair? Mapping individual differences in valuation of reciprocity, charity, and impartiality. *Social Justice Research*, *30*, 438-449.

Noë, R., & Hammerstein, P. (1995). Biological markets. *Trends in Ecology & Evolution*, *10*, 336-339.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573-577.

O'Connor, K. M., & Carnevale, P. J. (1997). A nasty but effective negotiation strategy: Misrepresentation of a common-value issue. *Personality and Social Psychology Bulletin, 23*, 504-515.

Opotow, S., Gerson, J., & Woodside, S. (2005). From moral exclusion to moral inclusion: Theory for teaching peace. *Theory into practice*, *44*, 303-318.

Pillutla, M., Malhotra, D., & Murnighan, J. K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, 39, 448-455.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature, 432*, 499–502.

Pearce, W. B. (1974). Trust in interpersonal communication.

Peck, J. R. (1993). Friendship and the evolution of co-operation. *Journal of Theoretical Biology, 162*, 195-228.

Pietraszewski, D. (2016). How the mind sees coalitional and group conflict: The evolutionary invariances of n-person conflict dynamics. *Evolution and Human Behavior, 37,* 470–480.

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences, 17,* 413–425.

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology, 12,* 487–503.

Richter, S. J., & Payton, M. E. (2003). Performing Two-Way Analysis of Variance Under Variance Heterogeneity. *Journal of Modern Applied Statistical Methods,2*(1), 152-160.

Robinson, S. L. (1996). Trust and breach of the psychological contract. *Administrative Science Quarterly, 41*, 574-599.

Rubin, K., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In W. Damon & R.M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 571–645). New York, NY: Wiley.

Santoro, M. A., & Paine, L. S. (1993). Sears auto centers (Harvard Business School case 9-394-010). *Boston: Harvard Business School Publishing*.

Sedikides, C., Hoorens, V., & Dufner, M. (2015). Self-enhancing self-presentation: Interpersonal, relational, and organizational implications. In F. Guay, H. Marsh, D. M. McInerney, & R. G. Craven (Eds.), *International advances in self research. Self-concept, motivation and identity: Underpinning success with research and practice* (pp. 29-55). Charlotte, NC, US: IAP Information Age Publishing.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*, 32-70.

Sedikides, C., Campbell, W. K., Reader, G. D., & Elliot, A. J. (1999). The relationship closeness induction task. *Representative Research in Social Psychology*, *23*, 1-4.

Sell, A. N. (2011). The recalibrational theory and violent anger. *Aggression and Violent Behavior*, *16*, 381-389.

Sell, A., Sznycer, D., Al-Shawaf, L., Lim, J., Krauss, A., Feldman, A., Rascanu, R., Sugiyama, L., Cosmides, L., & Tooby, J. (2017). The grammar of anger: Mapping the computational architecture of a recalibrational emotion. *Cognition, 168*, 110–128.

Shaw, A. (2013). Beyond "to share or not to share." *Current Directions in Psychological Science*, *22*, 413–417.

Shaw, A. (2016). Fairness: What it isn't, what it is, and what it might be for. In D. Berch & D. Geary (Eds.), *Evolutionary perspectives on education and child development* (pp. 193–214). Cham, Switzerland: Springer International Publishing.

Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2016). The development of inequity aversion: understanding when (and why) people give others the bigger piece of the pie. *Psychological Science*, *27*, 1352-1359.

Shaw, A., DeScioli, P., Barakzai, A., & Kurzban, R. (2017). Whoever is not with me is against me: The costs of neutrality among friends. *Journal of Experimental Social Psychology*, *71*, 96–104.

Shaw, A., DeScioli, P., & Olson, K.R. (2012). Fairness versus favoritism in children. *Evolution and Human Behavior*, *33*, 736–745.

Shaw, A., & Knobe, J. (2013). Not all mutualism is fair and not all fairness is mutualistic. *Behavior and Brain Sciences, 36,* 100–101.

Shweder, R.A. (2000). Moral maps, 'first world' conceits and the new evangelists." In L. Harrison & S. Huntington (Eds.), *Culture matters: Cultural values and human progress* (pp. 158-177). New York, NY: Basic Books, Inc.

Stewart-Williams, S. (2007). Altruism among kin vs. nonkin: Effects of cost of help and reciprocal exchange. *Evolution and Human Behavior, 28,* 193-198.

Schweitzer, M. E., & Croson, R. (1999). Curtailing deception: The impact of direct questions on lies and omissions. *International Journal of Conflict Management, 10,* 225-248.

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes, 101*, 1-19.

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30,* 497-514.

Talwar, V., Murphy, S. M., & Lee, K. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development, 31,* 1-11.

Telfer, E. (1970, January). Friendship. In *Proceedings of the Aristotelian Society* (Vol. 71, pp. 223-241). Aristotelian Society, Wiley.

Tokuno, K. A. (1983). Friendship and transition in early adulthood. *The Journal of genetic psychology*, *143*, 207-216.

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.

Toris, C., & DePaulo, B. M. (1984). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology*, *47*, 1063.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology, 46*, 35-57.

Tseng, S., and B. J. Fogg. (1999). "Credibility and Computing Technology." *Communications of the ACM, 42*, 39–44.

Tyler, J. M., Feldman, R. S., & Reichert, A. (2006). The price of deceptive behavior: Disliking and lying to people who lie to us. *Journal of Experimental Social Psychology*, *42*, 69-77.

Virgil, J. M. (2007). Asymmetries in the friendship preferences and social styles of men and women. *Human Nature, 18*, 143-161.

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblowers dilemma and the fairness–loyalty tradeoff. *Journal of Experimental Social Psychology, 49*, 1027-1033.

Weber, J. M., Malhotra, D., & Murnighan, J. K. (2004). Normal acts of irrational trust: Motivated attributions and the trust development process. *Research in Organizational Behavior*, *26*, 75-101.

Wieselquist, J., Rusbult, C. E., Foster, C. A., & Agnew, C. R. (1999). Commitment, pro-relationship behavior, and trust in close relationships. *Journal of personality and social psychology*, *77*, 942.

Xue, M., & Silk, J. B. (2012). The role of tracking and tolerance in relationship among friends. *Evolution and Human Behavior*, *33*, 17-25.