Int J Primatol
DOI 10.1007/s10764-009-9366-x

Cebus apella Tolerate Intermittent Unreliability in Human Experimenters

Sarah F. Brosnan · Frans B. M. de Waal

Received: 9 March 2009 / Accepted: 10 June 2009

Abstract Monkeys form expectations for outcomes based on interactions with human experimenters. Capuchins, a cooperative New World monkey species, not only anticipate receiving rewards that the experimenter indicates, but also apparently anticipate rewards based on what the experimenter has given to their partners. However, this could be due to subjects responding to either outcomes or experimenters. Here we examine whether capuchins will continue to interact with human experimenters who are occasionally unreliable. We tested 10 monkeys with a series of familiar human experimenters using an exchange task. The experimenters had never before participated in exchange studies with these monkeys, hence the monkeys learned about their behavior during the course of testing. Occasionally experimenters were unreliable, failing to give a reward after the monkey returned the token. The monkeys did recognize these interactions as different, responding much more quickly in trials following those that were nonrewarded than in other situations with the same experimenter. However, subjects did not change their preference for experimenters when given the opportunity to choose between the unreliable exchanger and another exchanger, nor did subjects learn to prefer reliable experimenters from watching other monkeys' interactions. Instead, subjects returned the tokens to the same location from which they received it. These results indicate that capuchins may not be sensitive to isolated instances in which experimenters are unreliable, possibly because of a strong bias to returning the token to the location from which it was donated.

Keywords capuchins · *Cebus apella* · exchange · expectation

S. F. Brosnan (✉)
Department of Psychology & Neuroscience Institute, Georgia State University, Atlanta, GA 30302, USA
e-mail: sbrosnan@gsu.edu

S. F. Brosnan · F. B. M. de Waal
Living Links, Yerkes National Primate Research Center, Atlanta, GA 30329, USA

F. B. M. de Waal
Psychology Department, Emory University, Atlanta, GA 30322, USA

Published online: 25 July 2009

Introduction

Monkeys and apes are sensitive to expectations, reacting negatively when expectations are violated (Brosnan and de Waal 2003; Tinkelpaugh 1928). For almost a century, it has been known that macaques (*Macaca spp.*) will respond negatively when a reward is surreptitiously switched by the experimenter, leaving them with a less desirable reward than anticipated (Tinklepaugh 1928). In this case, the monkeys formed expectations for what they should receive based on outcomes that were indicated for them. More recent work indicates that capuchins (*Cebus apella*) may also form expectations for what they should receive based on outcomes for their partners. They respond negatively when a partner receives a better reward than they do after having completed the same interaction with the experimenter (Brosnan and de Waal 2003; Fletcher 2008; van Wolkenten *et al.* 2007). In these cases, expectations were apparently based on reward outcomes, but primates may also form expectations about certain individuals with whom they interact.

In daily interactions, individuals face the question of whether they should continue interacting with existing social partners, or find a new partner. Although there is a cost to forming a new relationship, there is also a cost to continuing to interact with a social partner who is not fulfilling expectations, often known as cheating or defection. Thus, individuals may do best to evaluate interactions rapidly and cease interaction if necessary. Judgments of partner reliability are known in several species, e.g., fish (Bshary and Grutter 2006), including primates. Capuchins are much more likely to continue to cooperate with conspecific partners who share rewards with them (Brosnan *et al.* 2006; de Waal & Berger 2000). Chimpanzees, too, cooperate more frequently with individuals who share with them (Melis *et al,* 2006a, b). These apes are also able to evaluate interactions between human experimenters and choose to associate with the individual who is generous instead of one who is selfish (Russell *et al.* 2008; Subiaul *et al.* 2008), although other species of ape do not appear to make this distinction (Russell *et al.* 2008). Finally, New and Old World primates (including apes) appear to be sensitive to others' intents, distinguishing between experimenters' intentional and accidental actions (Call *et al.* 2004; Phillips *et al.* 2009; Wood *et al.* 2007). Once an individual has determined a partner to be unreliable, it may take effort to restart the relationship. In a cooperative situation, tamarins follow a tit-for-tat variation, called two-tits-for-a-tat, requiring 2 cooperative interactions from a previous defector before cooperation will restart (Chen and Hauser 2005).

Aside from individual learning, animals may learn with whom to cooperate from watching others. If a social partner interacts with 2 others, 1 of whom is reliable and 1 of whom is not, others observing this should choose to interact with the reliable partner. In the chimpanzee study mentioned earlier, chimpanzees were also able to learn which experimenters were generous by watching them interact with other chimpanzees (Subiaul *et al.* 2008). We know that capuchins acquire information from watching others (Perry *et al.* 2003), and individuals can learn which of 2 tokens yields a higher reward after watching a conspecific exchange only 20 times with an experimenter (Brosnan and de Waal 2004b). However, awareness of how a potential partner acts with a conspecific may not be informative regarding that individual's reactions with others, including the observer. To our knowledge, this has not been empirically tested in primates, though guppies do prefer to associate with conspecifics that have cooperated with them in the past (Dugatkin and Alfieri 1991), but they do not prefer to associate with the more cooperative of 2 conspecifics after watching them interact in a cooperative task (Brosnan *et al.* 2003).

We investigated how monkeys respond to a social partner who does not reliably complete an interaction. We did this using a familiar exchange paradigm, in which subjects exchange a token with a human experimenter to receive a reward (Brosnan and de Waal 2004b). We intentionally designed the experiments so that subjects were exposed to very few unreliable exchanges, to determine how they responded to an isolated incident of cheating rather than whether subjects could learn over time to avoid certain experimenters. Familiar experimenters with whom the monkeys had never done exchange tasks interacted with the monkeys, some of which failed to return the offered reward on some occasions (unreliable exchangers). For this, we predicted that capuchins would continue to interact with unreliable experimenters in situations in which the risk was low (experimenters rewarded the majority of exchanges), but would reduce interaction when the risk was high (many unrewarded exchanges). We then investigated whether the subjects continued to interact with unreliable exchangers, and whether they took an opportunity to interact with a new experimenter. We predicted that subjects would choose to interact with the second experimenter instead of the unreliable one, when given the option to do so. We finally investigated whether subjects would learn to avoid unreliable exchangers from observing interactions with other individuals. We predicted that after watching an unreliable exchanger, subjects would be more likely to interact with the other available experimenter.

Methods

Subjects

The subjects included 10 adult and subadult brown capuchins (3 adult males, 2 subadult males, and 5 adult females) housed in 2 social groups at the Yerkes National Primate Research Center, in Atlanta, GA. All but 1 of the adult females were pregnant or carrying a dependent offspring at some point during testing.

The groups in which the subjects lived were housed in 2 large, indoor/outdoor enclosures (de Waal 1997). Each enclosure contained ample 3-dimensional climbing space as well as trapezes, perches, and enrichment items. Purina small primate chow was provided twice a day, at *ca.* 0930 h and 1730 h. Monkeys received a tray consisting of fruit, vegetables, and bread with a protein solution every day at *ca.* 1730 h and running water *ad libitum*. We followed this feeding schedule regardless of the day's testing, and subjects were never food or water deprived.

We had previously trained subjects to enter transport cages, which allowed us to place individuals into a test chamber with their cooperation. Individuals were comfortable with the procedure and were well habituated to the test chamber. The test chamber was divided by a mesh partition into 2 equal sized (36×60×60 cm) compartments, and we conducted all testing in only 1 of them. The test chamber was backed by an opaque panel, which provided the subjects with vocal, but no visual or tactile, access to their groups. This allowed us to interact with subjects in a controlled manner with minimal distractions from the group. We always allowed dependent offspring into the test chamber with their mothers.

We had previously used all but 2 of the individuals in food sharing and cooperation studies in our laboratory and all had extensive experience interacting with humans. However, the individuals were all mother reared and none had been separated from the group or intensively trained in any way, e.g., language training, that could have caused them to associate more with humans than their own species.

Exchange Paradigm

We operationally defined exchange as the subject returning an inedible token to the experimenter, for which the subject received a food reward (Hyatt and Hopkins 1998). Unless otherwise noted, we defined exchange as the experimenter giving the token(s) to the subject, then standing in front of the test chamber with left hand outstretched, palm up, as a begging gesture, and holding the reward above the left hand with the right hand. Subjects received the reward on the placement of the token into the exchanger's left hand. If subjects threw tokens out of the test chamber or did not place them into the experimenter's hand, they were not rewarded.

All subjects had participated in 2 prior tasks involving exchange and therefore no training was required (Brosnan and de Waal 2004a, 2004b). In these experiments, subjects always received the rewards they were offered for completing the exchange interaction. In a previous study (Brosnan and de Waal 2004b), subjects watched a conspecific interact with the experimenter to learn the value of a set of tokens. We videotaped all sessions, together with time in hundredths of a second, on either a Super-VHS or digital video recorder and S. F. Brosnan later collected data from the videotapes. A second observer collected latency data during testing.

Subjects underwent a number of different experiments, each consisting of a procedure repeated multiple times. Throughout, test refers to an experimental type, trial to the procedure that was repeated multiple times per test, and session to the set of trials for a particular test, i.e., a test consisted of a multiple-trial session.

Human Experimenters

We required a series of novel exchangers, with whom the capuchins were familiar but with whom they had never exchanged. We could not use completely unfamiliar exchangers because the monkeys become extremely agitated by strangers and refuse to interact with them. In all but 1 case, we used the 5 novel individuals, each in only a single test. In the exceptional case (the Unreliable Exchanger test), S. F. Brosnan was the reliable exchanger, who never failed to provide a reward. All exchangers had worked directly with the monkeys as research assistants for a minimum of 6 mo and were familiar to all of the subjects from day-to-day interactions and other experimental procedures. However, none of the monkeys had ever participated in an exchange experiment with any of the novel individuals. Each subject saw a different experimenter for each of the tests, so there were no carryover effects between the experimenter's behavior in one test and another. Because there were a limited number of appropriate experimenters, we conducted each test only a single time on each of the 10 monkeys to avoid the possibility that the subjects learned responses over time.

The monkeys appear to discriminate between the individuals with whom they interact; however, to ensure that experimenters were visually distinct, each individual wore a differently colored, but otherwise identical, shirt over the standard laboratory clothes. The shirt colors were consistent with the role of the experimenter throughout testing. The procedure is sufficient for capuchins to distinguish between experimenters in other contexts (Chen *et al.* 2006).

Experimenter Preference Control

The Experimenter Preference Control established whether subjects preferred to return to one side of their enclosure over another or to one exchanger over another. If subjects preferred one exchanger over the other when both were reliable, then it would be difficult to verify whether preferences that emerged in later experiments were due to a random preference or to the actions of the experimenters.

Trials consisted of 20 exchanges with familiar exchangers. Exchangers took turns donating a token to the capuchin and exchangers switched position, e.g. left and right, after trial 10, so each exchanger donated 50% of the time from each side. Capuchins received the same reward regardless of their choice of the exchanger to whom to return the token.

The movements during exchange were stereotyped, to control for the subject cueing on unintentional cues. Each exchanger began 2 steps back from the test chamber. The exchanger giving the rock stepped forward; gave the rock to the subject directly, to maximize recognition by the subject of which exchanger was the donor; then stepped back parallel with the second exchanger. On cue from one exchanger, both stepped forward, placed their left hand on an X of tape on their side of a table in front of the test chamber, and placed their right hand, with the reward, up at the same height and distance from the cage, directly above the left hand. We used this table in all subsequent tests to verify that the experimenters' movements were coordinated and that neither was closer to the subject. We analyzed data to see if there was a tendency to 1) exchange only on one side, 2) exchange with only one individual, or 3) exchange with the donor of the item. The results were the baseline for comparison for later results.

Location Preference Control

To distinguish whether monkeys preferred to exchange to the same location versus the same exchanger, we conducted another test based on the Experimenter Preference Control. This test was identical, except the experimenters alternated position between the monkey's receipt of the rock and the offer to exchange. If subjects preferred to return to the donor, they needed to follow that individual; however, if they chose based on location, they should return to the non-donor. We analyzed the data to determine if there was a tendency to 1) exchange only on one side, 2) exchange with only one experimenter, or 3) exchange with the donor of the item. These results were the baseline for comparison for later results.

Unreliable Exchanger Test

Here we explored the reactions of the subjects during their first encounter with an unreliable exchanger. Virtually all of the subjects had been a part of this colony since their births, and the single exception had not participated in exchange experiments in his previous laboratory. Thus, we know that none had experienced an experimental test in which an exchange with an experimenter was not rewarded.

The methodology was similar to that of the control tests. The unreliable exchanger—one of the novel exchangers—placed a token in the test chamber, and then held up the reward being offered for exchange. The interaction commenced with 5 normal trials to verify baseline exchange latencies with the new exchanger. Next, the exchanger failed to give the offered reward on the 1st, 5th, 10th, 15th, and 20th of 20 trials, but instead placed the reward back in a pocket. Afterward, the known reliable exchanger (S. F. Brosnan) conducted 5 trials to see if any reactions were extrapolated across all exchangers, or were restricted to the cheater.

We expected subjects to alter their latency to exchange; remove the item, e.g., throw the item out of the front of the test chamber; or refuse to exchange, e.g., ignore the exchanger. We compared responses between the initial exchanges and 1) the 5 unrewarded exchanges, 2) the 15 rewarded exchanges, and 3) the final (rewarded) exchanges with the reliable exchanger.

New Experimenter Available Test

A subject exposed to an unreliable partner should be open to interactions with a novel partner whose reliability is untested. Here we gauged subjects' reactions when a new exchange partner was available to replace an unreliable exchanger. The test commenced with 10 trials in which a novel, unreliable exchanger failed to reward 50% of the trials, determined using a random number table before the test. For the second set of 10 trials, another novel exchanger, who was reliable, was present, and the subject could return the item to either exchanger. The unreliable exchanger always donated the item. To control for bias, the exchangers alternated sides after each exchange, but not between donation and return of the token. The subject in this test served as a model for the following test. We collected data on whether the subjects 1) ceased exchanging, 2) commenced exchanging with the reliable partner, or 3) continued to exchange with the unreliable partner.

Social Influence on Partner Choice

Identifying the tendencies of an exchange partner, e.g., cooperative or noncooperative, through observation of their interactions with a social partner instead of direct experience should benefit subjects. In this way, the subject could immediately invest more in a cooperative partner, hence increasing payoffs, and avoid interaction with a noncooperative partner, hence minimizing costs.

Subjects observed a social partner interact with an unreliable exchanger. (This was the New Experimenter Available test for the social partner.) We then gave the subjects 10 trials in which they could exchange with the unreliable exchanger or the reliable exchanger. Unlike in previous tests, both of the exchangers offered a token, allowing the subject to choose from whom to receive the item and, as in previous tests, both exchangers rewarded identically for exchange. The exchangers alternated position after every trial to control for side biases. We compared results to previous tests to ascertain the strength of any preference exhibited. Owing to the small number of available monkeys and exchangers, we conducted this test simultaneously with a New Experimenter Available test. (The social partner was actually performing this test.)

Statistics

Subjects showed a bias, preferring to return tokens to one side of the experimental chamber. Therefore, we performed the analysis using a side-correction procedure common to signal detection analysis. The Correct Choice Measure (CCM) A′ is a non-parametric measure of discrimination sensitivity that controls for biased guessing (Grier 1971). The CCM varies between 0 and 1, with 0.5 indicating chance performance (Stanislaw and Todorov, 1999). All statistical tests reported are on a total sample size of 10 individuals: 5 male and 5 female. We conducted comparisons between 2 dependent groups using the Wilcoxon signed rank sum test. Although the sample size was 10 for all tests except the Social Influence test, for some statistics we report smaller *N*, due to ties. Because the sample size was <15, we used only exact tests (Mundry and Fischer 1998). We used Friedman's test for comparisons of multiple dependent variables. We conducted some χ^2 tests with individual subjects' data, and owing to the small sample size, in every case we applied the Yates continuity correction. All statistics are 2-tailed.

Results

Exchanger Preference Control

In this test, capuchins could choose to return a token to the donor or to a second experimenter present during the test. Subjects preferred to return the token to the donor from whom the token was received (mean CCM=0.71, t=6.217, df =8, p<0.001; we did not include 1 individual whose data resulted in a 0 in the denominator when calculating the individual CCM).

In a second version of this test, explicitly controlling for the location bias, the exchangers switched positions after the donation of the token to the subject but before the monkey returned the token. Thus, we examined whether subjects were more likely to return the token to the donor or to the side on which the donation was made, e.g., were more likely to return it to the experimenter who had not donated the token. In this case, subjects preferred to return the token to the *other* experimenter—the one who had not donated the token, but was now occupying the position from which the token was received (Fig. 1; mean CCM=0.80, t=5.342, df=9, p<0.001)—indicating a preference for returning the token to a location, but not an individual.

Unreliable Exchanger Test

For this condition, the experimenter failed to give the reward on some exchanges. There were no refusals to exchange in any of the unreliable exchanger tests,

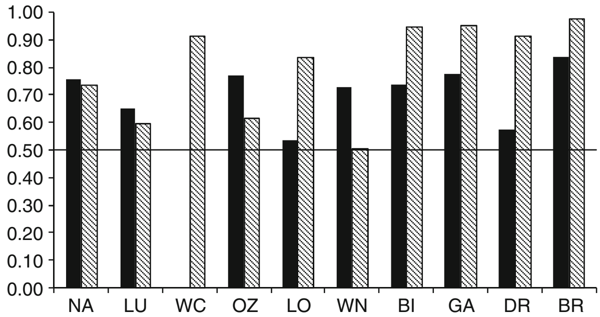

Fig. 1 Preference for the experimenter who donated the token (black bars) or the side on which the token was originally donated (hatched bars) for each of the 10 subjects. The *y*-axis indicates the choice for each individual, corrected for side bias using the Correct Choice Measure procedure, in which 0.5 is chance behavior. WC is missing a value for the token series because the CCM calculation resulted in a 0 in the denominator.

indicating that unreliability did not cause individuals to cease all interactions, although these monkeys will cease exchanging in other situations (Brosnan and de Waal, 2003). However, the latency to exchange did vary between conditions (Fig. 2: Friedman's test; $\chi^2=9.92$, $n=10$, df=4, $p=0.04$). Analyzing this in more detail, we found that the monkeys took the same amount of time to exchange with both experimenters in the first few, rewarded, trials with them, e.g., in the first 5 trials with the novel exchanger, before this experimenter failed to reward them and in the trials at the end with the familiar exchanger; Wilcoxon signed-rank sum test $T+=30$, $n=10$, $p=0.423$), indicating that during reliable series of exchanges the identity of the experimenter was irrelevant to them. These exchanges were also of the longest latency, on average. In addition, subjects did not differ in latency between their initial exchanges with the unreliable experimenter and those in which the experimenter failed to return a reward (initial vs. unrewarded: $T+=40$, $n=10$, $p=0.116$). However, subjects exchanged more rapidly when rewarded during these 20 trials than they had in the initial sessions (initial vs. rewarded: $T+=49$, $n=10$, $p=0.014$; initial vs. rewarded trials immediately after unrewarded trials: $T+=46$, $n=10$, $p=0.032$).

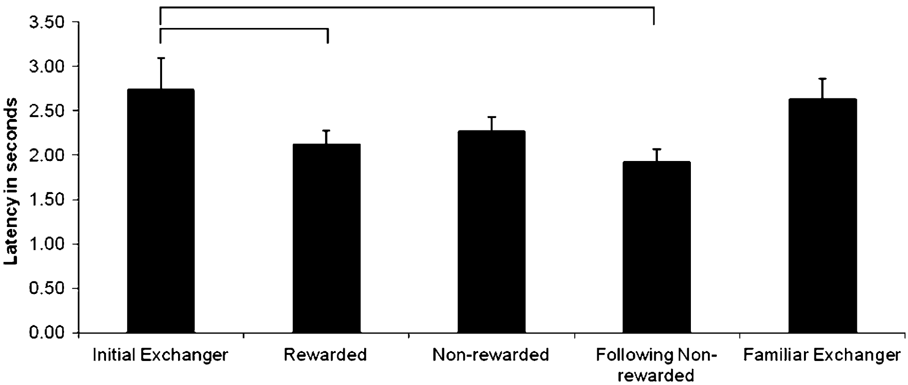

Fig. 2 The average latency for the subject to return the token in each of the conditions during the Unreliable Exchanger test. The lines over the bars indicate significant differences ($p<0.05$).

New Experimenter Available Test

In this test, in which cheating was much more common, i.e., every other trial, on average, rather than every 5th trial, there was no difference in latency between initial trials, trials in which the experimenter failed to return a reward, and trials following these episodes of unreliability (χ^2=2.6, df=1, p=0.273). When we considered the average latency to exchange for each individual when cheated versus trials immediately after cheating, we found a significant difference, with individuals exchanging more rapidly after they have been cheated (T+ =49, n=10, p=0.014).

Although we saw no response to the unreliable exchanger in this or in previous tests, it was possible that reactions were more subtle, such as choosing to interact with another individual. Nevertheless, in this test subjects were more likely to return the token to the experimenter who gave it to them (CCM=0.70, t=4.169, df=6, p= 0.006). Moreover, the magnitude of this preference did not differ from that of the original experimenter preference tests (original: t=–0.328, df=6, p=0.754; experimenters switch: t=0.325, df=6, p=0.756).

Social Influence on Partner Choice

For this condition, subjects observed their partners' interactions and then interacted with both experimenters. Owing to experimental difficulties, only 6 subjects completed this test (3 males and 3 females). Again, subjects chose to return their token to the side from which it was received, without any apparent interest in which experimenter was more reliable (CCM=0.75, t=12.558, df=3, p=0.001). Not all subjects completed all exchanges, suggesting the possibility that subjects preferred to exchange in situations in which the reliable exchanger donated the token. However, subjects showed no preference between returning the token to the cheating or noncheating exchanger (all χ^2 values are <0.62).

Discussion

We exposed capuchins to a series of experimenters with whom they could complete a familiar task, exchanging a token for a reward. Some of the experimenters were unreliable, routinely failing to give rewards to the subjects for completing the exchange, a situation they had never experienced in any previous experiment. Subjects appeared to recognize when exchangers were unreliable, i.e., when exchanges were unrewarded, as the latency to complete the exchange significantly changed during trials that were interspersed with nonrewarded exchanges, in particular after unrewarded trials, perhaps indicating uncertainty about the outcome of the trial. However, our original predictions were not met; when given the opportunity to choose between 2 exchangers, including the unreliable exchanger, subjects failed to show a change in preference away from the less reliable exchanger, regardless of whether the subject had experienced a large or small number of unrewarded exchanges. This was true both when the other exchanger was known to be reliable from previous interactions with the subject (S. F. Brosnan in the Unreliable Exchanger test) and when the other exchanger was novel, having never participated in an exchange task with the subject (New Partner Available test).

Instead of showing a preference for one experimenter or the other, subjects apparently used location to determine with whom to exchange, returning the token to whichever experimenter was standing at the location from which they received it. This led to the subjects preferring to interact with the donor, except in the second Experimenter Preference test, regardless of which experimenter was more reliable. This is in contrast to chimpanzees, which learned to prefer generous exchangers (Russell *et al.* 2008; Subiaul *et al.* 2008). Not surprisingly, based on these results, subjects also did not learn to prefer a reliable exchanger over one who was unreliable when watching another monkey participate (again, chimpanzees did learn to prefer the generous experimenter in this situation: Subiaul *et al.* 2008). The results appear to indicate that monkeys do not change their behavioral patterns when an experimenter occasionally fails to reward them for completing the exchange.

Although capuchins share many behaviors with chimpanzees, the species apparently behave differently when assessing the reliability of human experimenters. However, it is noteworthy that in the current experiment, even the unreliable experimenter gave the proffered reward on at least 50% of occasions, and more often in most conditions, whereas in the chimpanzee studies the selfish experimenter never gave a reward (Russell *et al.* 2008; Subiaul *et al.* 2008). Thus, the level of cheating in the current experiment may have been too low to elicit a response. It is also possible that the intermittent rewarding by the unreliable experimenter actually increased reinforcement for selecting the unreliable exchanger (Ferster and Skinner, 1957).

As with any study, there are several other possibilities that may explain negative results. First, we chose experimenters who were familiar to the monkeys and had good relationships with them. From a practical perspective, it is very difficult or impossible to coax monkeys to work with individuals whom they do not know. In addition, most normal interactions among conspecifics involve individuals from their social group, with whom they have some history, even if in a novel situation, as here. However, it may be that the history that each experimenter had with the monkeys predisposed the monkeys to believe that the experimenters were basically reliable, and that the events in this study were an anomaly. In this case, the instances of nonrewarding may have been taken in the larger context of the monkeys' relationships with the experimenters, rather than the monkeys basing their reactions solely on the current experiment.

In addition, we intentionally conducted only 1 session per pair because we were interested not in whether they could learn to prefer one experimenter over the other, but in their initial reaction to an unreliable experimenter. Note that capuchins did learn to prefer one token over another after only a single session of observing a partner interact with the tokens (Brosnan and de Waal, 2004b). However, Subiaul and colleagues (2008) found that while chimpanzees did not respond on initial interactions, they did so later on. Thus, the capuchins may learn which experimenter is more reliable after more extensive interactions.

A final possibility is that the monkeys were not sufficiently motivated because the focus of the interaction was with the experimenter, who was not a conspecific. However, this seems unlikely because capuchins are highly sensitive to other situations in which the experimenter varies treatment between individuals (Brosnan and de Waal 2003; Fletcher 2008; van Wolkenten *et al.* 2007), and chimpanzees responded to the reliability of experimenters in 2 previous experiments (Russell *et al.* 2008; Subiaul *et al.* 2008). Moreover, in a very similar situation involving paired monkeys exchanging different tokens for foods of different value, monkeys can learn which token is the higher-value of the 2 simply by watching a conspecific interact with a human experimenter (Brosnan and de Waal 2004b).

It is of note that the subjects did not choose randomly between experimenters in this situation, but routinely preferred to return the token to the experimenter on the side from which they had received it. Thus it is possible that preference for an experimenter was masked by the strong location bias. Similarly, chimpanzees' performance can be affected when choices must be made between 2 different objects, as opposed to focusing on a single object, masking subjects' abilities at tool-use tasks (Girndt *et al.* 2008). In future studies involving designs in which subjects must choose between 2 options, or 2 different experimenters, it is critical to take the location preference into account.

These results indicate that capuchins are not sensitive to isolated incidents of unreliability by the experimenter in a simple exchange task, and possibly that they are making social judgments based not solely on the previous interaction, but on the history of the relationship. However, this result could be due to a masking effect inherent in our experimental design, in which subjects showed strong location preference that may have overridden the effect of the manipulation. Thus, we also note the necessity of assessing experimental designs for potential masking effects before drawing strong conclusions.

Acknowledgments We thank Lisa Bradley, Jason Davis, Marietta Dindo, May Lee Gong, and Laura Mullen for assisting with data collection; 2 anonymous reviewers for comments on an earlier draft of the manuscript; and the animal care and veterinary staff of the Yerkes National Primate Research Center for care for our subjects. S. F. Brosnan was funded by a National Science Foundation Human and Social Dynamics Grant (SES 0729244) and the laboratory was funded by NSF grant (IOS-0718010) to the senior investigator. The YNPRC is fully accredited by the American Association for Accreditation for Laboratory Animal Care.

References

Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature, 425*, 297–299.

Brosnan, S. F., & de Waal, F. B. M. (2004a). A concept of value during experimental exchange in brown capuchin monkeys. *Folia Primatologica, 75*, 317–330.

Brosnan, S. F., & de Waal, F. B. M. (2004b). Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey, *Cebus apella*. *Journal of Comparative Psychology, 118*(2), 133–139.

Brosnan, S. F., Earley, R. L., & Dugatkin, L. A. (2003). Observational learning about predator inspection in guppies, *Poecillia retiuclata*. *Ethology, 109*, 823–833.

Brosnan, S. F., Freeman, C., & de Waal, F. B. M. (2006). Partner's behavior, not reward distribution, determines success in an unequal cooperative task in capuchin monkeys. *American Journal of Primatology, 68*, 713–724.

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in cleaner fish mutualism. *Nature, 441*, 975–978.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': chimpanzees understanding of human intentional action. *Developmental Science, 7*, 488–498.

Chen, M. K., & Hauser, M. D. (2005). Modeling reciprocation and cooperation in primates: Evidence for a punishing strategy. *Journal of Theoretical Biology, 235*, 5–12.

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *Journal of Political Economy, 114*(3), 517–537.

de Waal, F. B. M. (1997). Food transfers through mesh in brown capuchins. *Journal of Comparative Psychology, 111*(4), 370–378.

de Waal, F. B. M., & Berger, M. L. (2000). *Payment for labour in monkeys. Nature, 404*, 563.

Dugatkin, L. A., & Alfieri, M. (1991). Guppies and the TIT FOR TAT strategy: preference based on past interaction. *Behavioral Ecology and Sociobiology, 1991*(28).

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Fletcher, G. E. (2008). Attending to the outcome of others: Disadvantageous inequity aversion in male capuchin monkeys (*Cebus apella*). *American Journal of Primatology, 70*, 901–905.

Girndt, A., Meier, T., & Call, J. (2008). Task constraints mask great apes' ability to solve the trap-table task. *Journal of Experimental Psychology: Animal Behavior Processes, 34*(1), 54–62.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*, 424–429.

Hyatt, C. W., & Hopkins, W. D. (1998). Interspecies object exchange: Bartering in apes? *Behavioural Processes, 42*, 177–187.

Melis, A. P., Hare, B., & Tomasello, M. (2006a). Chimpanzees recruit the best collaborators. *Science, 311*, 1297–1300.

Melis, A. P., Hare, B., & Tomasello, M. (2006b). Engineering cooperation in chimpanzees: Tolerance constraints on cooperation. *Animal Behavior, 72*, 275–286.

Mundry, R., & Fischer, J. (1998). Use of statistical programs for nonparametric tests of small samples often leads to incorrect *P* values: examples from Animal Behaviour. *Animal Behaviour, 56*, 256–259.

Perry, S., Panger, M., Rose, L., Baker, M., Gros-Luis, J., Jack, K., *et al.* (2003). Traditions in wild white-faced capuchin monkeys. In D. M. Fragaszy & S. Perry (Eds.), *The Biology of Traditions: Models and Evidence*. Cambridge, UK: Cambridge University Press.

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M. & Santos, L. R. (2009). 'Unwilling' versus unable': Capuchins' (*Cebus apella*) Understanding of human intentional action? *Developmental Science*. doi:10.1111/j.1467-7687.2009.00840.x

Russell, Y. I., Call, J., & Dunbar, R. I. M. (2008). Image scoring in great apes. *Behavioural Processes, 78*, 108–111.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behaviour Research Methods, Instruments & Computers, 31*, 137–149.

Subiaul, F., Vonk, J., Okamoto-Barth, S., & Barth, J. (2008). Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Animal Cognition, 11*, 611–623.

Tinklepaugh, O. L. (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology, 8*, 197–236.

van Wolkenten, M., Brosnan, S. F., & de Waal, F. B. M. (2007). Inequity responses in monkeys modified by effort. *Proceedings of the National Academy of Sciences, 104*(47), 18854–18859.

Wood, J. N., Glynn, D. D., Phillips, B. C., & Hauser, M. D. (2007). The perception of rational, goal-directed action in nonhuman primates. *Science, 317*, 1402–1405.