Anim Cogn (2013) 16:677–684
DOI 10.1007/s10071-013-0603-5

ORIGINAL PAPER

Taking personality selection bias seriously in animal cognition research: a case study in capuchin monkeys (*Sapajus apella*)

F. Blake Morton · Phyllis C. Lee · Hannah M. Buchanan-Smith

Received: 12 October 2012 / Revised: 13 January 2013 / Accepted: 15 January 2013 / Published online: 31 January 2013

Abstract In most experimental work on animal cognition, researchers attempt to control for multiple interacting variables by training subjects prior to testing, allowing subjects to participate voluntarily, and providing subjects with food rewards. However, do such methods encourage selection bias from subjects' personalities? In this study, we trained eighteen zoo-housed capuchin monkeys (*Sapajus apella*) for two experiments, under conditions of positive reinforcement (i.e. food rewards) and free-choice participation. Using a combination of behavioral and rater-based methods, we identified and validated five personality dimensions in these capuchins (Assertiveness, Openness, Neuroticism, Sociability, and Attentiveness). Scores on Openness were positively related to individual differences in monkey task participation, reflecting previous work showing that such individuals are often more active, curious, and willing to engage in testing. We also found a negative relationship between scores on Assertiveness and performance on tasks, which may reflect the trade-offs between speed and accuracy in these animals' decision-making. Highly Assertive individuals (the most sociable within monkey groups) may also prioritize social interactions over engaging in research. Lastly, monkeys that consistently participated and performed well on both tasks showed significantly higher Openness and lower Assertiveness compared to others, mirroring relationships found between personality, participation, and performance among all participants. Participation and performance during training was clearly biased toward individuals with particular personalities (i.e. high Openness, low Assertiveness). Results are discussed in light of the need for careful interpretation of comparative data on animal cognition and the need for researchers to take personality selection bias more seriously.

Keywords Platyrrhines · Temperament · Cognitive experiment · Selection bias · Associative learning · Training

Electronic supplementary material The online version of this article (doi:10.1007/s10071-013-0603-5) contains supplementary material, which is available to authorized users.

F. B. Morton (✉) · P. C. Lee · H. M. Buchanan-Smith
Behaviour and Evolution Research Group and Scottish Primate Research Group, Psychology, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK
e-mail: morton.blake@gmail.com

Introduction

Selection bias is a statistical error that occurs when certain individuals or groups within a population are unequally represented in a study (Blaney and Millon 2009). Common sources of selection bias include: (1) self-selection of study subjects, (2) targeting a specific area or population, (3) ending a study after reaching a desired result, and/or (4) excluding data based on arbitrary grounds (Blaney and Millon 2009). Because certain individuals or groups are sampled more often than others, selection bias violates statistical assumptions about "randomly selected" data (e.g. Bornehag et al. 2004; Cahan and Gamliel 2006; Malani 2008).

In studies involving animal subjects, researchers must be wary of selection bias resulting from individual differences in personality (defined here as consistent individual differences in behavior and decision-making among animals; Gosling and John 1999). For example, in a study by Garamszegi et al. (2009), more exploratory birds were sampled more often than less exploratory birds. Similarly,

Wilson et al. (2011) reported that bolder fish were more often caught in open water, whereas more timid fish were typically found closer to places of refuge. If individuals with certain personality traits are over-sampled, and those traits interact with the study variable(s) of interest, then the resulting selection bias might lead to invalid conclusions about the population from which the data were obtained (Biro and Dingemanse 2009). This can also create issues when making comparisons between studies, particularly where subjects are sampled differently and/or are different in terms of their expression of specific personality traits.

In most experimental studies of animal cognition, researchers attempt to control for multiple interacting variables by training subjects prior to actual testing, allowing subjects to participate voluntarily ("free-choice" participation), and/or providing subjects with food rewards to encourage their participation, motivation, and attention (Benson-Amram and Holekamp 2012; Clark 2011; Fagot and Paleressompoulle 2009; Garber et al. 2012; Reid et al. 2008). Quite often, however, individuals are excluded from such studies if they do not participate regularly (e.g. Evans et al. 2008; Fagot and Paleressompoulle 2009; Rehbein and Moss 2002), or if they do not meet training criteria within a set timeframe (e.g. "100 trials for five consecutive days", Watson and Ward 1996; or "10 sessions of 100 trials each"; Bussey et al. 2008). Although such methods may reduce noise from variables like stress, motivation, or experience, they may also introduce personality selection bias into the study.

Individual differences in personality reflect differences in animals' participation and performance on a variety of learning and cognitive tasks. For example, slow-exploring guppies are often better spatial navigators (Burns and Rodd 2008), less neophobic birds have a tendency to approach novel objects faster and thus learn more quickly (Boogert et al. 2006), and extraverted humans are generally better at mental updating (Campbell et al. 2011). However, it is rarely acknowledged within the animal cognition literature whether or not individuals that proceed to the testing phase of an experiment have personalities that differ from those that fail to meet participation and training criteria (Carere and Locurto 2011; Sih and Del Giudice 2012). Considering that different personalities perform differently on tasks and that many studies of animal cognition rely on data obtained from small samples of subjects (e.g. Deaner et al. 2006), the potential risk of personality selection bias warrants further investigation.

In this study, we examined whether individual differences in personality would result in selection bias while training eighteen capuchin monkeys (*Sapajus apella*) for two experiments under conditions of positive reinforcement (i.e. food rewards) and free-choice participation. We first tested whether monkeys' scores on personality consistently and reliably predicted individual differences in participation and performance during training. Then, we tested for significant differences in the personalities of monkeys that did/did not regularly participate and perform well during training. We hypothesized that individual differences in the personality of each subject would be related to differences in participation and performance during training. Additionally, we hypothesized that monkeys that participated regularly and ultimately completed training would have different personalities compared to those individuals that failed to meet our participation and training criteria.

Methods

Study site and subjects

Brown capuchin monkeys were studied at the Living Links to Human Evolution Research Center (LL), located within the Royal Zoological Society of Scotland (RZSS), Edinburgh Zoo, UK. Capuchin monkeys are of great interest to comparative psychologists given their behavioral and cognitive similarities to great apes, such as relatively large brains, extractive foraging skills (e.g. tool-use), cooperative and food-sharing habits, and "cultural" traditions (Fragaszy et al. 2004).

Study subjects were from two breeding groups, and each of these groups cohabited with a group of common squirrel monkeys (*Saimiri sciureus*). At the time of study, the "East" group contained 2–3 adult males, 3 adult females, 3 juveniles, and 0–5 infants. The "West" group contained 2 adult males, 3 adult females, 4–5 juveniles, and 2–5 infants. Age of the study subjects ranged from 2 to 40 years for males (average 11.17 ± 13.72 years, $N = 12$ capcuhins), and 3–14 years for females (average 8.86 ± 3.63 years, $N = 7$ capuchins). All monkeys were captive born except the eldest male (Diablo), which was likely wild-born and came to LL as established members of the groups. One individual (Kato) was hand-reared. Both groups were housed in identically designed, but mutually exclusive, 189 m^3 indoor enclosures with natural light and near-permanent access to a 900 m^2 outdoor enclosure containing trees and other vegetation, providing ample opportunity to engage in natural behaviors. All monkeys received commercial TrioMunch pellets supplemented with fresh fruits and vegetables three times daily and were given cooked chicken and hardboiled eggs once every week. Water was available ad libitum at all times. Further details of housing and husbandry are provided in Leonardi et al. (2010). This study was noninvasive, approved by local ethics committees, and complied with regulations of the Association for the Study of Animal Behavior (ASAB 2012).

Personality assessment

Data on the personality of each monkey come from a previous study, and those methods and results can be found in Morton et al. (2013). Briefly, we studied capuchin monkey personality among 127 capuchins (>1 year old) across 7 international sites. Each subject was rated by one to seven raters (mean: 3.24 ± 1.61 raters), who were researchers ($N = 25$) and care staff ($N = 3$). Raters had at least 1 year of experience working with the individuals. Ratings were made on the Hominoid Personality Questionnaire (HPQ), which consists of 54 adjectives, each paired with one to three sentences defining it within the context of nonhuman primate behavior (Weiss et al. 2009, 2011). Inter-rater reliabilities were tested using ICC's, and all salient ratings were entered in a Principle Components Analysis, which reduced ratings into components, or "dimensions" of personality. This analysis revealed five distinct personality dimensions for brown capuchin monkeys: Assertiveness, Openness, Attentiveness, Neuroticism, and Sociability. Examples of the particular traits that clustered around each of these personality dimensions can be found in Table 1. Individual scores on each of the five personality dimensions were calculated, and validated against relevant behaviors recorded within each social group up to a year later (Morton et al. 2013). For example, scores on Assertiveness positively correlated with the amount of time each individual spent being aggressive toward others, while scores on Openness positively correlated with the amount of time individuals spent playing with others (Morton et al. 2013). The results of this analysis indicate that some element of behavioral consistency (i.e. personality) has been measured among these capuchins. More importantly, these findings demonstrate that the ratings do not merely reflect raters' implicit understanding of how personality traits *should* co-associate; rather, they reflect real-world behavioral patterns among subjects.

Table 1 Traits characteristic of the five personality dimensions identified in brown capuchin monkeys (from Morton et al. 2013)

Personality dimension	Examples of trait loadings[a]
Assertiveness	High: *bullying, aggressive, stingy*
	Low: *submissive, vulnerable, timid*
Openness	High: *innovative, active, curious*
	Low: *conventional, lazy, quitting*
Neuroticism	High: *excitable, distractible, erratic*
	Low: *decisive, predictable, stable*
Sociability	High: *sociable, affectionate, friendly*
	Low: *solitary, anxious, depressed*
Attentiveness	Low: *disorganized, unperceptive, clumsy, thoughtless, impulsive*

[a] Salient loadings |>0.4| from PCA of 54 traits using varimax-rotation (Morton et al. 2013)

Task apparatus and testing

All monkeys ($N = 18$, excluding infants) were given the opportunity to engage in training for two experiments. Training for the first experiment (hereafter referred to as "Task 1") was conducted between November 8, 2011 and January 13, 2012. Training for the second experiment (hereafter referred to as "Task 2") was conducted between February 15 and April 1, 2012.

Training took place in research cubicles, which were divided into two compartments (both 49.5 cm × 52.1 cm × 51.4 cm) and separated by a transparent plastic door that was halfway open. Participating monkeys could freely walk between the two compartments. Monkeys were able to access the cubicles via an open corridor leading from their main enclosure. Since the establishment of LL in 2008, subjects have been involved in a number of experimental studies, with a wide array of methodological designs (see MacDonald and Whiten 2011); however, the tasks and methods of administration in our study had not been used before.

After a monkey had voluntarily entered the research cubicles, the door was closed behind them, and training began. Prior to initiating a session, however, if the subject appeared agitated or distracted (e.g. rapidly moving between the two cubicles, looking away from the experimenter), the monkey was released back into the main enclosure. Also, as is typical of free-choice methods, subjects could control when they wished to initiate or end a training session, usually by either gesturing or pressing their hand on the cubicle door exit. If a subject signaled that it wanted to leave, the researcher stopped the test, opened the door, and released the monkey back into the main enclosure.

In Task 1, during each trial, a food reward was placed in front of one of two compartments. The location of the food reward (left or right compartment) was randomly selected for each new trial. The goal was for the monkey to learn that by moving into the compartment that had the food directly in front of it, the researcher would hand them the food. If the monkey failed to do this, no food was delivered, and the trial was ended. A subject's response on a given trial was considered to be "correct" when they sat inside the cubicle opposite to the food reward.

In Task 2, during each trial, two white-opaque cups were placed in front of one of the two compartments. The position of each cup (left or right compartment) was randomly selected for each new trial. The two cups differed in size, with one cup twice as tall (height: 19 cm, diameter: 6.4 cm) as the other cup (height: 9.5 cm, diameter:

6.4 cm). For this task, the goal was for the monkey to learn that by moving into the compartment facing the larger cup, they would receive a food reward that was hidden inside the cup. The larger cup was always the "winner", the smaller cup was always the "loser". If the monkey failed a trial, no food was delivered, and the trial was ended. A subject's response on a trial was considered to be "correct" when they sat inside the cubicle directly opposite to the larger cup.

For both tasks, none of the sessions contained situations where the correct response required subjects to remain within the same cubicle on four or more trials. Additionally, each monkey received up to 12 trials per session per day for 4 days a week until they met training criteria, or until they had been scoring at chance levels after 3 months. Each trial lasted for 5 s, and all trials were separated by 5–7 s. For each correct trial, subjects received a food reward (e.g. raisin or piece of papaya). All training sessions were video recorded, and later coded. Inter-observer reliability tests were conducted using a sub-sample of these data, whereby 36 trials from 5 monkeys (total 180 trials) were independently scored by two observers. Cronbach's alpha for each of the five monkeys was 1.0, indicating that both observers scored identically.

Participation was calculated for each monkey by dividing the number of sessions in which they participated by the total number of session offered to them, multiplied by 100. The performance of each individual was calculated for each task by dividing the total number of trials answered correctly by the total number of trials undergone, multiplied by 100. Based on a binomial test, we established that individuals would need to score $\geq$80 % of trials (i.e. >10 out of 12 trials) correctly in order to be statistically above chance. Individuals that scored $\geq$80 % on three consecutive sessions were considered to have learned the task, and their training subsequently ended.

Statistical analyses

Using a PASW 18.0 (SPSS, IBM corp., Chicago, USA) package, we performed forward step-wise multiple regressions to determine the relationships between monkeys' performance, participation, and scores on each personality dimension. Pearson correlations were used to examine the relationship between two variables (e.g. participation and a specific personality dimension). We also used Pearson correlations to determine whether significant associations between two variables were positive or negative. All statistical tests were two-tailed, and alpha was set to 0.05.

Where specified in our results, we used bootstrapping procedures to control for differences in sample size to test for differences in personality between monkeys that did/did not participate and perform well on our tasks. Bootstrapping procedures were conducted in Matlab by P. Hancock (Stirling University) as follows: A new dataset was created by selecting at random with replacement the personality scores of (1) monkeys that participated and performed well and (2) monkeys that did not. Then, for each sample of subjects, their mean scores on each personality dimension were calculated. This procedure was repeated 10,000 times; if for a given personality dimension, the mean of either sample came out significantly larger than the other on more than 9,750 occasions, this was defined as $p < 0.05$, two-tailed.

Results

Participation and performance on tasks

Task 1

Thirteen monkeys participated in Task 1 and participated on an average of 83.9 ± 27.9 % of sessions (range 10–100 %). One of these subjects participated in 10 % of sessions, 8 subjects participated in 100 % of sessions, and the remaining subjects participated between 50–100 %. Each participant received between 11–108 trials (mean: 73.8 ± 23.1 trials) over the course of the task, which was a function of how many sessions in which they participated. Average successful performance among participants was 73.8 ± 16.04 % (range 45.5–91.7 %). A total of eight monkeys learned the task within 60–84 trials (mean performance: 83.38 ± 4.5 %), which was faster than the 5 remaining monkeys who were still performing at chance levels by the end of the task (mean performance: 58.4 ± 16.0 %). Monkeys that met learning criteria for this task were the same eight individuals that participated on 100 % of sessions. Individual learning curves for each monkey are provided in Fig. 1 of the "Online Resource".

Task 2

Fifteen monkeys participated in Task 2 and participated in an average of 70.4 ± 36.4 % sessions (range 5.6–100 %). Four of these subjects participated in <50 % of sessions, 7 subjects participated in 100 %, and the remaining subjects participated somewhere in between 50–100 %. Participants received between 12–216 trials (mean: 100.8 ± 69.7 trials) which was a function of how many sessions in which they participated. Average successful performance among individuals was 65.0 ± 13.5 % (range 41.7–87.5 %). Five monkeys learned the task within 48–84 trials (mean performance: 80.0 ± 4.7 %), which was faster than the 10 remaining monkeys, who, as in Task 1, were still

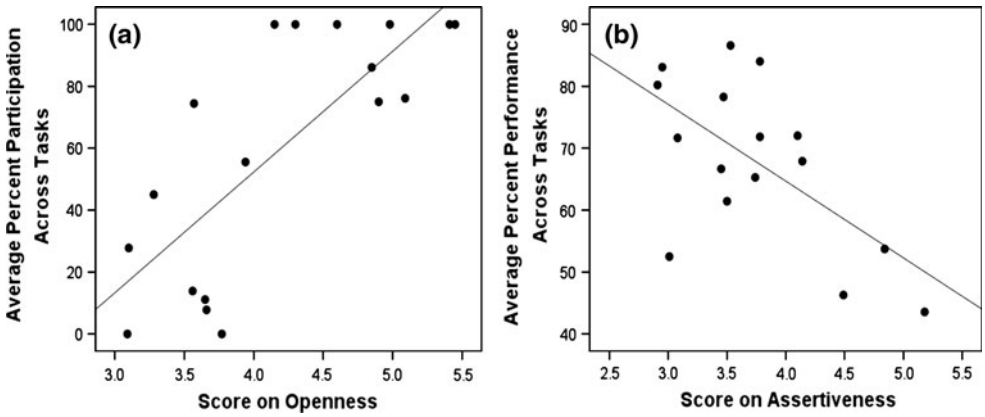

Fig. 1 Relationship between **a** individual scores on Openness and rate of participation (% sessions participated), and **b** individual scores on Assertiveness and performance (% trials correct). Because results were similar in both tasks, here, the data have been averaged across tasks for simplicity. Eighteen monkeys were given the option to participate on both tasks; twelve monkeys performed on both tasks

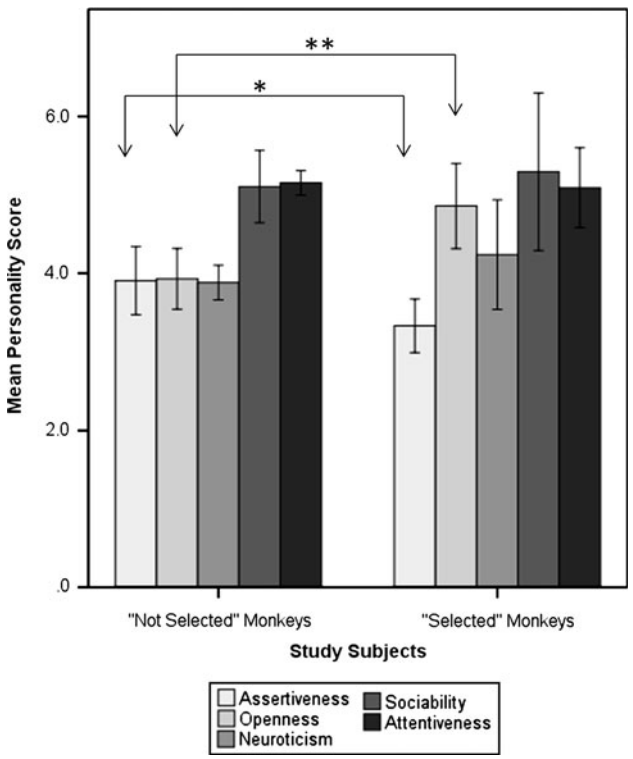

Fig. 2 Comparison of personality differences between monkeys that did/did not proceed to the testing phase of each task. "Selected" monkeys scored significantly lower on Assertiveness (bootstrapped $p = 0.025$) and higher on Openness (bootstrapped $p = 0.002$), compared to "Not Selected" monkeys

performing at chance levels by the end of the task (mean performance: 57.4 ± 9.2 %). Five of the seven monkeys that participated on 100 % of sessions also met learning criteria. Individual learning curves for each monkey are provided in Fig. 2 of the "Online Resource".

Personality, participation, and performance on tasks

For both tasks, individual differences in participation were positively related to scores on Openness (Table 2; Fig. 1a; Task 1: Adj. $R^2 = 0.414$, $F = 13.02$, $p = 0.002$, $df = 17$; Task 2: Adj. $R^2 = 0.495$, $F = 17.7$, $p = 0.001$, $df = 17$). In Task 1, among the participating monkeys, individual differences in performance were negatively related to scores on Assertiveness (Table 2; Fig. 1b) and positively related to scores on Openness (Table 2; Adj. $R^2 = 0.635$, $F = 21.9$, $p = 0.001$, $df = 12$). Individual scores on Assertiveness were not significantly related to scores on Openness ($r = -0.180$, $p = 0.474$, $N = 18$ monkeys). For Task 2, participating monkeys' average performance was negatively related to Assertiveness (Table 2; Fig. 1b; Adj. $R^2 = 0.223$, $F = 5.02$, $p = 0.043$, $df = 14$). None of the other personality dimensions were significantly related to participation or performance for either task (Table 2).

Personality differences in subjects that did/did not meet criteria

The same five monkeys met learning criteria on both tasks and were among those individuals that participated on 100 % of sessions. These "selected" individuals scored significantly higher on Openness (bootstrapped $p = 0.002$) and significantly lower on Assertiveness (bootstrapped $p = 0.025$), compared to other monkeys (Fig. 2); however, their scores on the other personality dimensions were not significantly different from other subjects (Fig. 2).

Discussion

We hypothesized that individual differences in personality were related to differences in participation and performance while training subjects for two experiments. We also hypothesized that monkeys that ultimately completed training (and would thus proceed to the testing phase of an experiment) would have personalities that differed from those individuals that failed to meet our criterion. Both hypotheses were supported by our results.

Table 2 Pearson correlations between individual scores on personality, participation, and performance (Task 1 and Task 2)

Personality type	Participation		Performance*	
	Task 1 ($N = 18$ monkeys)	Task 2 ($N = 18$ monkeys)	Task 1 ($N = 13$ monkeys)	Task 2 ($N = 15$ monkeys)
Assertiveness	$r = -0.007$	$r = -0.239$	$r = -0.816$	$r = -0.528$
	$p = 0.977$	$p = 0.339$	$p = 0.001$	$p = 0.043$
Openness	$r = 0.67$	$r = 0.724$	$r = 0.748$	$r = 0.434$
	$p = 0.002$	$p = 0.001$	$p = 0.003$	$p = 0.106$
Neuroticism	$r = 0.023$	$r = 0.210$	$r = 0.368$	$r = 0.426$
	$p = 0.929$	$p = 0.403$	$p = 0.217$	$p = 0.114$
Sociability	$r = 0.408$	$r = 0.381$	$r = 0.163$	$r = -0.135$
	$p = 0.093$	$p = 0.119$	$p = 0.594$	$p = 0.632$
Attentiveness	$r = -0.116$	$r = -0.172$	$r = -0.001$	$r = -0.037$
	$p = 0.647$	$p = 0.496$	$p = 0.996$	$p = 0.895$

* N of participants differs according to how many subjects participated in the task

Studies across multiple species have found that individuals with Open traits tend to be more active, exploratory, and curious in novel situations (Bacon 1980; Bergman and Kitchen 2009; Biondi et al. 2010; Kashdan et al. 2004). Thus, highly Open individuals are more willing to engage in cognitive testing and, as a consequence, generally complete a task faster than others (e.g. Herrelko et al. 2012; Bates and Shieles 2003; Watson and Ward 1996). Reflecting such studies, we found positive associations between differences in participation during training and monkeys' scores on Openness. Additionally, monkeys' scores on Openness were positively correlated with their performance on Task 1. Lastly, monkeys that consistently met training criteria on both tasks had significantly higher scores on Openness compared to other monkeys and were also among those individuals that participated 100 %.

In Task 2, however, despite there being a positive association between participation and Openness, two of the monkeys that participated in 100 % of sessions did not meet training criteria, and there was no significant association between monkeys' scores on Openness and performance on this task. Considering that we found a significant relationship between Openness and performance in Task 1, this could suggest that variables other than personality contributed to animals' decision-making during Task 2, such as differences in motivation, the social atmosphere within monkeys' groups at the time of testing, or task difficulty. Indeed, average participation and performance among all participants was lower in Task 2 compared to Task 1 (see "Results").

Monkeys' performance on each task was also associated with their score on Assertiveness. Several authors have proposed that individuals with aggressive personality traits exhibit a "sunk cost effect" (or "Concorde fallacy"; Dawkins and Carlisle 1976), whereby individuals that are more aggressive emphasize speed over accuracy in their decision-making and thus take longer to perform a task (reviewed in Sih and Del Giudice 2012). Although we did not specifically test speed over accuracy, our results support these predictions: First, Assertiveness was characterized by adjectives such as aggressive, bullying, and dominant (see Table 1; Morton et al. 2013). Second, scores on this personality dimension were positively correlated with the amount of time monkeys spent being aggressive toward others (Morton et al. 2013). Third, across all participants, individual differences in performance were negatively associated with scores on Assertiveness for both tasks; meaning that less Assertive individuals were able to solve both tasks faster. Lastly, individuals who consistently met learning criteria across both tasks scored significantly lower on Assertiveness compared to other monkeys. However, within our study groups, highly Assertive capuchins were among the most sociable individuals (Morton et al. 2013). Thus, relationships between Assertiveness and performance may also reflect other interactions between personality and decision-making, such as prioritizing social interactions over engaging in research.

Surprisingly, none of the other personality dimensions (Sociability, Neuroticism, and Attentiveness) were related to participation or performance, nor were mean scores on these dimensions significantly different between those subjects that did and did not pass the training phase of each task. Given that these dimensions are related to other aspects of subjects' behavior (e.g. scores on Sociability were positively associated with the amount of time monkeys spent in close proximity with others; Morton et al. 2013), the traits associated with Assertiveness and Openness may be particularly relevant to performance and participation within this capuchin population. However,

further work will be necessary to determine causal relationships among these variables. For instance, familiarity with the experimenter and/or research environment could underlie why individual differences in participation were not significantly associated with scores on Neuroticism or Sociability, while those of Assertiveness and Openness were.

Conclusions

Personality selection bias is potentially a global issue to the animal sciences (Biro and Dingemanse 2009). Within the context of animal cognition research, our findings indicate that subjects that meet participation/training criteria, and are thus involved in the testing phase of an experiment, can have personalities that differ from those of individuals that fail to meet the criteria. Thus, basic statistical assumptions about the experiment are violated. Cognitive researchers should therefore remain wary of making comparisons between studies, especially when studies differ in their methods of training or use of free-choice participation. Similarly, comparing experimental data between studies may not be valid if subjects differ in their expression of certain personality traits, particularly where small samples of subjects are described as being representative of a group or species.

Currently, throughout the animal cognition literature, individuals that never progress to the testing phase of an experiment are often excluded without further mention. Additionally, test subjects often come from a larger pool of individuals (e.g. socially housed animals), yet there is little or no discussion as to why some individuals were included in a study, while others were not. We should remain wary of any effects due to personality selection bias within such studies. In future, we therefore urge researchers to consistently assess and report personality differences among study subjects and discuss how such differences may have contributed to the outcome of an experiment. Doing so should greatly facilitate cognitive comparisons between studies.

Acknowledgments This research would not have been possible without the many people involved in the capuchin personality study, particularly A. Weiss, S. Brosnan, and B. Thierry. F.B.M. would also like to personally thank Prof. Andrew Whiten, Director of Living Links (RZSS), who gave permission to conduct research there, and all of the Living Links/RZSS staff and students who kindly provided support and assistance. We also thank K. Howie who gave helpful statistical advice, P. Hancock who kindly conducted the bootstrapping analysis used in this study, the BERG (University of Stirling) for their feedback, and J. Anderson, S. Brosnan, and the anonymous reviewers for providing useful comments on earlier drafts. F.B.M gratefully acknowledges the Charles A. Lockwood Memorial Fund and the University of Stirling for funding. The authors declare that they have no conflict of interest.

References

ASAB (2012) Guidelines for the treatment of animals in behavioural research and teaching. Anim Behav 83:301–309

Bacon ES (1980) Curiosity in the American black bear. Bears: their biology and management, vol. 4, a selection of papers from the fourth international conference on bear research and management, Kalispell, Montana, USA, February 1977, pp 153–157

Bates TC, Shieles A (2003) Crystallized intelligence as a product of speed and drive for experience: the relationship of inspection time and openness to *g* and Gc. Intelligence 31:275–287

Benson-Amram S, Holekamp KE (2012) Innovative problem solving by wild spotted hyenas. Proc R Soc B-Biol Sci. doi:10.1098/rspb.2012.1450

Bergman TJ, Kitchen DM (2009) Comparing responses to novel objects in wild baboons (*Papio ursinus*) and geladas (*Theropithecus gelada*). Anim Cogn 12:63–73

Biondi LM, Bo MS, Vassallo AI (2010) Inter-individual and age differences in exploration, neophobia and problem-solving ability in a neotropical raptor (*Milvago chimango*). Anim Cogn 13:701–710. doi:10.1007/s10071-010-0319-8

Biro PA, Dingemanse NJ (2009) Sampling bias resulting from animal personality. Trends Ecol Evol 24:66–68. doi:10.1016/j.tree.2008.11.001

Blaney PH, Millon T (2009) Oxford textbook of psychopathology. Oxford University Press, Oxford

Boogert NJ, Reader SM, Laland KN (2006) The relation between social rank, neophobia and individual learning in starlings. Anim Behav 72:1229–1239

Bornehag CG, Sundell J, Sigsgaard T, Janson S (2004) Potential self-selection bias in a nested case-control study on indoor environmental factors and their association with asthma and allergic symptoms among pre-school children. Scand J Public Health 34(5):534–543

Burns JG, Rodd FH (2008) Hastiness, brain size and predation regime affect the performance of wild guppies in a spatial memory task. Anim Behav 76:911–922. doi:10.1016/j.anbehav.2008.02.017

Bussey TJ, Padain TL, Skillings EA, Winters BD, Morton AJ, Saksida LM (2008) The touchscreen cognitive testing method for rodents: how to get the best out of your rat. Learn Memory 15:516–523. doi:10.1101/lm.987808

Cahan S, Gamliel E (2006) Definition and measurement of selection bias: from constant ratio to constant difference. J Edu Meas 43:131–144

Campbell AM, Davalos DB, McCabe DP, Troup LJ (2011) Executive functions and extraversion. Pers Individ Dif 51:720–725

Carere C, Locurto C (2011) Interaction between animal personality and animal cognition. Curr Zool 57:491–498

Clark FE (2011) Great ape cognition and captive care: can cognitive challenges enhance well-being? Appl Anim Behav Sci 135:1–12. doi:10.1016/j.applanim.2011.10.010

Dawkins R, Carlisle TR (1976) Parental investment, mate desertion, and a fallacy. Nature 262:131–133

Deaner R, van Schaik C, Johnson V (2006) Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. Evol Psychol 4:149–196

Evans TA, Beran MJ, Chan B, Klein ED, Menzel CR (2008) An efficient computerized testing method for the capuchin monkey (*Cebus apella*): adaptation of the LRC-CTS to a socially housed nonhuman primate species. Behav Res Methods 40:590–596. doi:10.3758/BRM.40.2.590

Fagot J, Paleressompoulle D (2009) Automatic testing of cognitive performance in baboons maintained in social groups. Behav Res Methods 41:396–404. doi:10.3758/BRM.41.2.396

Fragaszy DM, Visalberghi E, Fedigan LM (2004) The complete capuchin: the biology of the genus *Cebus*. Cambridge University Press, New York

Garamszegi LZ, Eens M, Török J (2009) Behavioural syndromes and trappability in free-living collared flycatchers, *Ficedula albicollis*. Anim Behav 77:803–812. doi:10.1016/j.anbehav.2008.12.012

Garber PA, Gomes DF, Bicca-Marques JC (2012) Experimental field study of problem-solving using tools in free-ranging capuchins (*Sapajus nigritus,* formerly *Cebus nigritus*). Am J Primatol 74:344–358. doi:10.1002/ajp.20957

Gosling SD, John OP (1999) Personality dimensions in nonhuman animals: a cross-species review. Curr Dir Psychol Sci 8:69–75

Herrelko ES, Vick S-J, Buchanan-Smith HM (2012) Cognitive research in zoo-housed chimpanzees: influence of personality and impact on welfare. Am J Primatol. doi:10.1002/ajp.22036

Kashdan TB, Rose P, Fincham FD (2004) Curiosity and exploration: facilitating positive subjective experiences and personal growth opportunities. J Pers Assess 82:291–305

Leonardi R, Buchanan-Smith HM, Dufour V, MacDonald C, Whiten A (2010) Living together: behavior and welfare in single and mixed species groups of capuchin (*Cebus apella*) and squirrel monkeys (*Saimiri sciureus*). Am J Primatol 72:33–47

MacDonald C, Whiten A (2011) The 'Living Links to Human Evolution' Research Centre in Edinburgh Zoo: a new endeavour in collaboration. Int Zoo Yearbook 45:1–17. doi:10.111/j.1748-1090.2010.00120.x

Malani A (2008) Patient enrollment in medical trials: selection bias in a randomized experiment. J Econom 144:341–351

Morton FB, Lee PC, Buchanan-Smith HM, Brosnan S, Thierry B, Paukner A, de Waal FBM, Widness J, Essler J, Weiss A (2013) Personality structure in brown capuchin monkeys: comparisons with chimpanzees, orangutans, and rhesus macaques. J Comp Psychol. doi:10.1037/a0031723

Rehbein L, Moss MB (2002) Explorations of three modes of spatial cognition in the monkey. Psicológica 23:139–163

Reid GMcG, Macdonald AA, Fidgett AL, Hiddinga B, Leus K (2008) Developing the research potential of zoos and aquaria. The EAZA research strategy EAZA executive office, Amsterdam. ISBN: 978-90-77879-10-8

Sih A, Del Giudice M (2012) Linking behavioural syndromes and cognition: a behavioural ecology perspective. Phil Trans R Sci B. doi:10.1098/rstb.2012.0216

Watson SL, Ward JP (1996) Temperament and problem solving in the small- eared bushbaby (*Otolemur garnettii*). J Comp Psychol 110:377–385

Weiss A, Inoue-Murayama M, Hong K-W et al (2009) Assessing chimpanzee personality and subjective well-being in Japan. Am J Primatol 71:283–292. doi:10.1002/ajp.v71:4

Weiss A, Adams MJ, Widdig A, Gerald MS (2011) Rhesus macaques (*Macaca mulatta*) as living fossils of hominoid personality and subjective well-being. J Comp Psychol 125:72–83. doi:10.1037/a0021187

Wilson ADM, Thomas RB, Keegan PM, Cooke SJ, Godin J-GJ (2011) Capture technique and fish personality: angling targets timid bluegill sunfish, *Lepomis macrochirus*. Can J Fish Aquat Sci 68:749–757. doi:10.1139/F2011-019