Accepted for publication 16th October 2020 in the Journal of Comparative Psychology.

Studying animal innovation at the individual level: A ratings-based assessment in capuchin monkeys (*Sapajus* [*Cebus*] *sp.*)

F. Blake Morton[1], Hannah M. Buchanan-Smith[2], Sarah F. Brosnan[3], Bernard Thierry[4,5], Annika Paukner[6], Jennifer L. Essler[7], Christopher S. Marcum[8], Phyllis C. Lee[2]

[1] Psychology, University of Hull, Hull, United Kingdom

[2] Behaviour and Evolution Research Group, Psychology Division, Faculty of Natural Sciences, University of Stirling, UK

[3] Language Research Center, Department of Psychology, and Center for Behavioral Neuroscience, Georgia State University, USA

[4] Physiologie de la Reproduction et des Comportements, Centre National de la Recherche Scientifique, Institut National de la Recherche Agronomique, Université de Tours, Nouzilly, France

[5] Psychology Department, Nottingham Trent University, UK

[6] Penn Vet Working Dog Center, School of Veterinary Medicine, University of Pennsylvania, USA

[7] National Institutes of Health, Bethesda, Maryland, USA

[8] Behaviour and Evolution Research Group, Psychology Division, Faculty of Natural Sciences, University of Stirling, UK

Author Note

We have no known conflicts of interest to declare.

We are very grateful to the many raters who contributed item ratings for this study, and to Drs Alexander Weiss, Jane Widness, and Frans de Waal for their assistance in acquiring those data. We thank Prof. Andrew Whiten for permission to conduct research at Living Links, and the Living Links staff for their support and assistance during data collection. F. B. M. thanks the University of Stirling and the Primate Society of Great Britain for funding. Research support for C. S. M. was provided by the Intramural Research Program of the National Institutes of Health. Finally, we thank the reviewers and editor for their helpful comments on the manuscript.

Correspondence concerning this article should be addressed to Dr Blake Morton, Psychology, University of Hull, HU6 7RX, United Kingdom. E-mail: b.morton@hull.ac.uk.

Abstract

Large-scale studies of individual differences in innovative behaviour among nonhuman animals are rare because of logistical difficulties associated with obtaining observational data on a large number of innovative individuals across multiple locations. Here we take a different approach, using observer ratings to study individual differences in innovative behaviour in 127 brown capuchin monkeys (*Sapajus* [*Cebus*] *sp.*) from 15 social groups and 7 facilities. Capuchins were reliably rated by 1 to 7 raters (mean 3.2 ± 1.6 raters/monkey) on a 7-point Likert scale for levels of innovative behaviour, task motivation, sociality, and dominance. In a subsample, we demonstrate these ratings are valid: rated innovation predicted performance on a learning task, rated motivation predicted participation in the task, rated dominance predicted social rank based on win/loss aggressive outcomes, and rated sociality predicted the time that monkeys spent in close proximity to others. Across all 127 capuchins, individuals that were rated as being more innovative were significantly younger, more social, and more motivated to engage in tasks. Age, sociality, and task motivation all had independent effects on innovativeness, whereas sex, dominance, and group size were non-significant. Our findings are consistent with long-term behavioural observations of innovation in wild white-faced capuchins. Observer ratings may therefore be a valid tool for studies of animal innovation.

Introduction

Some species have a proclivity for behavioural innovation, in which individuals of those species use new or modified behaviours to solve new or existing problems (Lee, 1991; Reader & Laland, 2003). Innovation has significant links with intelligence (Lee & Therriault, 2013; Ramsey et al., 2007), species differences in brain size (Lefebvre, 2013; Lefebvre et al., 2004; Reader, 2003; Reader & Laland, 2002), the evolution of tool use and culture (Biro et al., 2003; Boesch, 1995; Lefebvre, 2013; Reader et al., 2011; Tian et al., 2018), and the breadth of a species' ecological niche (Ducatez et al., 2015; Overington, Griffin, et al., 2011). At the proximate level, a range of dispositional and situational factors likely play a role in generating innovative behaviour (Amici et al., 2019; Brosnan & Hopper, 2014; Griffin & Guez, 2014; Lee, 1991; Lee & Moura, 2015; Moura & Lee, 2004; Ramsey et al., 2007; Reader & Laland, 2003). At its core, being "innovative" requires, at the very least, being able to *discover* (implicitly or explicitly) novel or modified behaviours (Ramsey et al., 2007; Reader & Laland, 2003). Unless an animal learns from its innovative action, and can repeat that action, the discovery will be lost from the repertoire of the individual.

Large-scale studies on individual differences in animal innovation are relatively few in number firstly because observations on innovative behaviour itself are rare, and secondly because of logistical difficulties (e.g. time, money, and standardising methods) associated with documenting innovations across a large, multi-site sample of individuals (Biro et al., 2003; Haslam et al., 2009). Observer ratings may help overcome such limitations. Indeed, a growing number of studies have shown that observer ratings are a reliable and valid tool for assessing a wide variety of behaviours and cognitive traits in animals (Freeman et al., 2013; Freeman & Gosling, 2010; Morton, Lee, & Buchanan-Smith, 2013; Morton, Lee, Buchanan-Smith, et al., 2013; Morton et al., 2015; Weiss et al., 2011; Weiss et al., 2012). Ratings also enable researchers to obtain data on multiple variables across a large sample of subjects within a reasonable timeframe, with the same definitions and methods (e.g. 7-point Likert scales) used consistently across observers, locations, and subjects to facilitate comparability.

In the current study, we obtained observer ratings on innovative behaviour within a large, multi-site sample of captive brown capuchin monkeys (*Sapajus* [*Cebus*] *sp*.). To help explain individual variation in innovative behaviour, we considered six variables (age, sex, dominance, task motivation, group size, and sociality) often linked to innovation that may reflect a myriad of reasons why individuals might be innovative, such as individual differences in personality (Benson-Amram et al., 2013; Henke-von der Malsburg & Fichtel, 2018; Huebner & Fichtel, 2015), physiology (Hopper et al., 2014), brain development and decline (Roskos-Ewoldsen et al., 2008), behavioural ecological niche (Aplin & Morand-Ferron, 2017; Giraldeau & Lefebvre, 1987; Liker & Bokony, 2009), and experience (Daveri & Parisi, 2015; Huebner & Fichtel, 2015). While many other factors may contribute to innovative behaviour, we opted to limit the number of variables to avoid oversaturating our model.

As with any study of animal innovation where subjects cannot be monitored continuously across their lifespan, it was not possible in the current study to observe and verify “new” innovations in our capuchins. Thus, to begin to assess the validity of observer ratings on capuchins’ innovative behaviour, we tested, in a subsample of our capuchins, whether the ratings could predict a relevant psychological construct *related* to innovative behaviour, specifically monkeys’ associative learning abilities. Being willing and able to discriminate and learn associatively from one’s actions can play an important role in the innovative process (Reader & Laland, 2003). If, for example, an animal cannot discriminate between old versus new actions, and learn new associations from its actions, then the chances of making a new discovery (i.e., making an association and repeating the innovative behaviour in the future) will be very limited. Under experimental conditions, animals that are more innovative are better at solving associative learning tasks (Griffin et al., 2013; Overington, Cauchard, et al., 2011). Thus, in the current study, we predicted that highly innovative monkeys would perform better on an associative learning task than less innovative individuals.

To further assess the validity of our observer ratings, we determined whether the same factors that predicted innovative behaviour across our entire sample of capuchins were consistent with findings from a 10-year observational study of innovations in wild white-faced capuchins (*Cebus capucinus*) (Perry et al., 2017). Specifically, we predicted that, like white-faced capuchins, individual differences in our capuchins' age and sociality (defined in terms of the amount of time individuals spent within proximity to others) would be important negative and positive predictors of their innovative behaviour, respectively, whereas sex and dominance (defined in terms of avoids, cowers, flees, and supplants) would show minimal, non-significant effects.

Method

Ethics

This study was approved by the ethics committee of the Psychology Division at the University of Stirling, the Living Links committee at the Royal Zoological Society of Scotland (RZSS), and complied with APA and ASAB ethical guidelines ("Guidelines for the treatment of animals in behavioural research and teaching," 2012).

Subjects

Subjects were 127 captive brown capuchins that were at least 1 year old, belonging to 15 social groups from 5 sites in the United States, 1 site in the UK, and 1 site in France (Table S1). Across all sites there were 60 males and 67 females. Age ranged from 1 to 40 years and the mean age was 11.0 years (*SD* = 8.9). To test the validity of item ratings, eighteen of these monkeys were observed at the Living Links to Human Evolution Research

Centre, affiliated with the Royal Zoological Society of Scotland (RZSS), U.K. Further details of housing and husbandry are provided in the 'Supplementary Information' (SI).

Observer ratings

Ratings were collected between 2010 and 2011 for a previous study (Morton, Lee, Buchanan-Smith, et al., 2013). Raters consisted of 25 researchers and 3 care staff who had known their subjects for at least one year. Definitions and scales for observer ratings on capuchins' innovative behaviour, sociality, dominance, and task motivation came from items from the Hominoid Personality Questionnaire (Morton, Lee, Buchanan-Smith, et al., 2013; Weiss et al., 2009). Each subject was rated by one to seven raters (3.2±1.6 raters per monkey) on each item based on the frequency of monkeys' behaviour on a 1 (absent) to 7 (very common) scale. Ratings were averaged across raters for each monkey. Measures of innovative behaviour came from the "innovation" item in the HPQ, which defined such behaviour as "the subject engages in new or different behaviours that may involve the use of objects or materials or ways of interacting with others". We later asked some of these raters to provide a few examples of innovative behaviour in their monkeys. For instance, one rater reported that a monkey was observed using a stick on several occasions to reach chow from under the fence, which other monkeys in the group did not do (Leverett and Rossetti, personal communication). In another instance, a rater reported that one of their monkeys would take a piece of wood, break pieces off of it, and then use it to scratch or comb its back, which had not been seen in any other monkey in that group by any rater (Leverett and Rossetti, personal communication).

Measures of dominance came from the "dominance" item in the HPQ, which was defined as "the subject is able to displace, threaten, or take food from other individuals; or the subject may express high status by decisively intervening in social interactions". Measures of sociality came from the "sociability" item in the HQP, which was defined as "the subject seeks and enjoys the company of other individuals and engages in amicable, affable, interactions with them". Measures of task motivation came from the "curiosity" item in the HPQ, which was defined as "the subject has a desire to see or know about objects, devices, or other individuals; this includes a desire to know about the affairs of other individuals that do not directly concern the subject".

Two intraclass correlations (Shrout & Fleiss, 1979) were used to determine interrater reliabilities for subjects rated by at least two raters. The first, ICC(3,1), indicates the reliability of individual ratings. The second, ICC(3,k), indicates the reliability of the mean of k ratings. Of the sample, 121 capuchins (out of the total 127 subjects) were rated by at least two raters (M = 3.35; SD = 1.57). Collectively, there was high inter-observer agreement across each item per monkey: dominance [ICC(3,1)=0.57, ICC(3,k)=0.82], innovation [ICC(3,1)=0.57, ICC(3,k)=0.82], sociability [ICC(3,1)=0.57, ICC(3,k)=0.82], and curiosity [ICC(3,1)=0.57, ICC(3,k)=0.82] (Morton, Lee, Buchanan-Smith, et al., 2013). Since there was no evidence that raters were unreliable, mean ratings for each item for all 127 monkeys were included in our analyses.

Raters' reliabilities were as good or even better than similar ratings reported in studies of humans and other animals (Freeman & Gosling, 2010; Gartner et al., 2014; McCrae & Costa, 1987). Because our raters passed the ICC reliability criteria, this also meant that no single rater was significantly biased towards over- or under-rating a given monkey (e.g. if they witness more behaviours compared to the other raters). Indeed, raters were instructed not to discuss their ratings and to make their ratings based on their *own* observations (not those mentioned by other people). Regarding the innovation ratings specifically, the Likert scale helped to ensure that raters made their ratings on the basis of behavioural frequency – not just one-off observations. Ratings data were normally distributed, not skewed, indicating that ratings were not biased towards raters recalling particular occasions of striking innovation in some monkeys but not others.

Testing the validity of observer ratings

Behavioural data (Table S3) were collected by an independent observer on the 18 capuchin monkeys at Living Links up to a year after those monkeys were rated on items. These data were used to validate interpretations of behaviour derived from ratings:

Innovative behaviour

Data on the Living Links capuchins' performances on a discrimination learning task were used to validate innovative behaviour ratings. While all 18 subjects were given the opportunity to voluntarily participate in the task, 15 of these monkeys participated. Testing occurred between 15 February 2012 and 1 April 2012, at 12 trials per session, four sessions per week. Monkeys were tested individually in cubicles to ensure all animals had the opportunity to engage in testing. The goal of the task was for individuals to learn the location of a hidden food reward by discriminating between two cups that were different sizes (details in SI). Learning performance was calculated for each monkey by dividing the total number of trials they completed correctly by the total number of trials they underwent, multiplied by 100.

Task motivation

Motivated animals are, of course, likely to voluntarily participate in tasks that require them to use their cognitive abilities (Skinner, 1938). Data on rates of voluntary participation in the learning task (see 'Innovative behaviour' above) were available for all 18 of the Living Links monkeys and therefore used to validate ratings on task motivation. Participation was calculated by dividing the number of sessions the monkey engaged in by the total number of session offered to them, multiplied by 100 (Morton, Lee, & Buchanan-Smith, 2013).

Sociality

Data on monkeys' time spent in close proximity to other group members were available on 18 of the Living Links capuchins, and therefore used to validate ratings on sociality. Capuchins who spend more time in close proximity with other group members are more sociable; they are more likely to engage in affiliative acts like grooming, food sharing, and coalitionary support (Morton et al., 2015), which is very typical of wild and captive capuchins (Ferreira et al., 2006; Fragaszy et al., 2004). Focal observations on all 18 monkeys' spatial proximity to others were made between May and August, 2011, totalling 3 hours per individual. Monkeys were sampled evenly between 9:00 and 17:30. Using point sampling methods (Martin & Bateson, 2007), group members within two body lengths from the focal were recorded at 1-min intervals for ten minutes per animal per day. On a given point sample, if no monkey was within two body lengths, the focal was described as "solitary". Scores were recorded at 1-min intervals and calculated within 10-minute sessions. Monkeys were observed on rotation across all 19 individuals; meaning, most of the time a given monkey was observed once a day, but on 20 occasions a monkey was observed more than once. On these occasions, sampling was separated by at least 21 minutes (M= 220.7 minutes, SD= 160.2 minutes).

Dominance

To test whether dominance ratings reflect social rank of individuals, social dominance was determined using data that were available on 18 of the Living Links capuchins (Morton, Lee, Buchanan-Smith, et al., 2013; Morton et al., 2015) by calculating David's Scores (DS) using data on win/loss outcomes during monkey's aggressive interactions (Gammell et al., 2003). All occurrences of fighting within the group were recorded while performing focal sampling of individuals outlined above (see '*Sociality*').

Analyses

In the subsample of 18 monkeys used to validate ratings, we used Pearson correlations to examine relationships between individual differences in item ratings, behaviours, and task performance. Across the entire sample (N=127 monkeys), age was skewed but normalised with a log (base=10) transformation. A linear mixed effects model was used to test for independent effects of age, sex, dominance, task motivation, sociality, and group size on innovative behaviour. This approach facilitates unbiased linear estimation of coefficients and robust standard errors that are adjusted for the clustering of animals by including random effects variance components for social group (intercept) and group size (slope). For this model, we calculated the percent adjusted R^2 that a particular covariate contributes to the full model, which we estimated using the leave-one-out method. As our 'group' variable captured information about location, and group size is a group-level variable, models were fit using linear mixed models with random intercept for group and random slope for group size. While bounded between 1 and 7, our dependent variable (innovative behaviour) and our key independent variables (sociality, task motivation, and dominance) are not discrete. Rather, because we measured them using a robust multi-rater design where values were averaged across raters as discussed above, they are continuous variables within the bounds. To bolster our argument that a linear model is appropriate for these analyses, we performed Shapiro-Wilk tests for the normality of each of these variables (Royston, 1982), though only our dependent variable need meet this assumption.

All Pearson correlations and log transformations were performed in SPSS 24 (IBM Corp., Chicago, IL, USA). Multivariate analyses were performed in the latest development release of R (R Core Team, 2019) using the "lmerTest" library for tests of linear mixed models (Kuznetsova et al., 2017).

Results

Validity of observer ratings

Ratings on innovative behaviour were significantly and positively related to performance on the discrimination learning task when all participants were included in the analysis (r=0.598, P=0.019, N=15 monkeys) and when only those participants that participated in >80% of sessions were included (r=0.787, P=0.02, N=8 monkeys). Ratings on task motivation were significantly and positively related to participation in the novel task (r=0.618, P=0.006, N=18 monkeys). Dominance ratings were significantly and positively related to social rank (r=0.833, P<0.001, N=18 monkeys). Sociality ratings were significantly and positively related to the amount of time individuals spent with other group members (r=0.495, P=0.037, N=18 monkeys).

Independent effects between innovative behaviour and sociality scores

One monkey was rated by a single rater. Given that ratings for monkeys with more than one rater were reliable, and that ratings were valid (see above), we included this individual with the remaining 126 monkeys for the following analysis.

A linear mixed effects regression model revealed that across all 127 capuchins, sociality, motivation to engage in tasks, and age all had independent and significant effects on innovativeness, whereas sex, dominance, and a random effect of group size did not (Table 1). Individual differences in innovative behaviour were significantly and positively related to sociality and task motivation, but negatively related to age (Figure 1).

Table 1

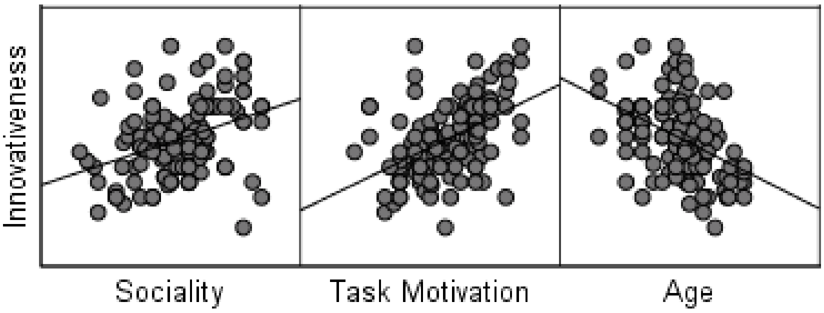
Independent effects of sociality, age, sex, dominance, and task motivation on individual differences in capuchins' ratings on innovative behaviour

	Estimate	Robust S.E.	Robust t	%R^2	Pr(>|t|)
(Intercept)	**2.17**	**0.66**	**3.29**	**---**	**<0.01**
Sociality	**0.22**	**0.09**	**2.44**	**8.37**	**0.02**
log(Age, base = 10)	**-0.79**	**0.31**	**-2.49**	**9.66**	**0.01**
Sex	0.05	0.18	0.27	0.18	0.79
Dominance	-0.05	0.06	-0.90	1.42	0.37
Task Motivation	**0.36**	**0.09**	**4.09**	**21.17**	**<0.001**

Note. Significant results ($P < 0.05$) in boldface. N in all cases = 127 monkeys. % R^2 is the percent contribution to the full model adjusted R^2 of a particular covariate by the leave-one-out method. Model fit statistics: Approximate Adjusted $R^2 = 0.351$, F-test: 13.07 on 5 and 120 d.f., $P < 0.0001$. Random effects variance components were of trivial size (Social Group Intercept < 0.002 and Group Size Slope < 0.005).

Figure 1

Independent associations between capuchins' innovative behaviour and individual differences in sociality, task motivation, and age (in years) (N=127 capuchins)



The small amount of variation explained by group size warranted retaining the covariate in the model as a random effect. We also ran a linear mixed model with an equivalent specification as our generalised estimating equation. The variance component associated with "location" was 0.004 which is negligible. The resulting random effects ("Supplementary information") differed only slightly in magnitude and thus any concern over a location or group bias is unfounded. With the exception of Dominance, each test resulted in our failure to reject the null that each variable was drawn from an underlying normal distribution. For Dominance, the deviation from normality is explained by the fact that dominance in these groups was highly distributed across individuals. Moreover, the shape of the histogram of this variable (Figure S1 and S2) suggested that it was drawn from an underlying uniform distribution which is supported by a Kolmogorov-Smirnov test (Conover, 1971) of uniformity (statistic=0.149, p-value=0.117) (Table S4). Such deviations might be problematic for the linear model as an outcome (dependent variable) but it is fine for an independent variable. Finally, the scatterplots of the dependent variable against the independent variables showed no observable heteroscedasticity that would indicate a violation of the underlying linearity of the relationship per the assumptions of the Pearson-product moment correlation or the linear model estimation.

Discussion

We used reliable observer ratings to study innovative behaviour in a large, multi-site sample of 127 brown capuchins. In a subsample of these capuchins, we found that the ratings predicted real-world behavioural patterns that were independently recorded up to a year later: ratings on innovative behaviour were correlated with performance on an associative learning task, task motivation scores were correlated with participation in the task, dominance scores were correlated with social rank based on win/loss aggressive outcomes, and sociality scores were correlated with the amount of time spent with other group members. Across all 127 monkeys, the independent effects of age, sociality, sex, and dominance reflected those reported in wild white-faced capuchins (Perry et al., 2017), ruling out captivity and methodological limitations of ratings as likely explanations for our results. Collectively, our findings support the notion that observer ratings may be a valid tool for studies of innovation.

As previously discussed, researchers very rarely have the luxury of being able to follow the same population continuously across generations to observe and verify new innovations. Thus, novel psychometric tasks (e.g. giving animals a novel puzzle feeder) are often used as an objective approach to experimentally induce animals to innovate (Benson-Amram et al., 2013; Henke-von der Malsburg & Fichtel, 2018; Huebner & Fichtel, 2015). Such approaches, however, come with their own limitations. For instance, it can be difficult to establish whether more frequent innovators are simply more motivated, less distracted, or have better experience or opportunities to engage in testing than other individuals. For this reason, psychometric tasks are not necessarily any more objective than observer ratings. Thus, much like on-going discussions from the animal personality literature (Freeman et al., 2013), future studies will likely benefit from using a combination of psychometric and ratings data to further test convergent validity between methods to study innovation.

In a similar vein, the psychological mechanisms that drive innovative behaviour in capuchins and other animals remain largely unknown (Ramsey et al., 2007). Studies of the common myna (*Sturnus tristis*) have shown that more frequent innovators are better at solving discrimination learning tasks, but do not perform as well on reversal learning tasks, suggesting that the associative learning underpinnings of the discrimination task were more relevant to innovation within this species than flexible learning (Griffin et al., 2013). As demonstrated in a subsample of our monkeys, ratings may reflect at least the associative learning processes related to capuchins' innovative behaviour (Griffin et al., 2013; Overington, Cauchard, et al., 2011; Ramsey et al., 2007; Reader, 2003). To better understand the psychological underpinnings of innovation in capuchins, we encourage studies to use a broader range of tasks varying in complexity and design, particularly those measuring other types of learning, inhibitory control, and intelligence (Huebner & Fichtel, 2015; Lee & Therriault, 2013).

Regarding our measure of sociality (i.e. time in close proximity to others), Morton et al. (2015) found that proximity loads onto the same factorial component as coalitions, food sharing, and grooming; meaning, at least in capuchins, all of these more "subtle forms" of sociality simply map onto the same thing: affiliative behaviour. Nevertheless, future work might consider whether these and other specific forms of sociality are better predictors of innovativeness, particularly time spent grooming, sharing food, and watching others while feeding. Using social network analysis can also provide a multi-dimensional approach to sociality for comparison.

Finally, captive animals are unlikely to face the same level of ecological pressure as in the wild (e.g. no predation risk), and can have a tendency to be more innovative than wild individuals of the same species (Benson-Amram et al., 2013). Nevertheless, as previously discussed, our findings are consistent with those found in wild capuchin monkeys (Perry et al., 2017). Future comparisons between captive and wild brown capuchins using the same or similar methods can therefore provide *complimentary* insight into the innovativeness of this species, for instance, in terms of controlling for factors like inter-group competition and predator vigilance, which might impact the amount of time wild (but not captive) capuchins can devote to being innovative.

Proximate underpinnings of capuchin innovation

We suggest at least two testable scenarios for why sociality might be positively correlated with innovative behaviour in brown capuchins. First, like most group-living primates, capuchins use strategies such as grooming, coalitions, and food sharing to achieve greater social embeddedness within their group (Ferreira et al., 2006; Fragaszy et al., 2004; Morton et al., 2015; Tiddi et al., 2012), and being more social may reduce stress, improve infant survival, provide better access to food and mating opportunities, and, in turn, lead to better fitness (Kalbitzer et al., 2017; Ostner & Schulke, 2018; Silk, 2007; Silk et al., 2003; Silk et al., 2009). Thus, a positive association between innovative behaviour and sociality may arise if, for example, being innovative enables individuals to concurrently improve their social status within groups. Second, individuals that are more social may simply have better opportunities in terms of the time and energy they can devote to experiment and engage in learning compared to less social individuals (Kummer & Goodall, 1985). Such opportunities may not necessarily be used to improve one's social status *per se* (e.g. foraging and self-directed innovativeness). This latter scenario might arise if sociality is a means through which capuchins solve an otherwise ecological problem (e.g. resource acquisition and protection from predators), and in turn, allow more time and/or opportunities for innovative behaviour.

Examining longitudinal associations between capuchins' innovative behaviour and sociality will help tease apart these and other possibilities.

While motivation may be the sole factor underlying individual differences in innovation in some species (van Horik & Madden, 2016), it only had a partial effect in our capuchins. Nevertheless, behavioural and cognitive traits are useless without animals being motivated enough to perform them, and so delineating possible interactions between task motivation (a situational effect) and personality (a dispositional effect) is required to better understand how innovative behaviour might be generated within these animals. Our findings may reflect food-related motivation (i.e. a situational effect) since capuchins' scores on task motivation were positively correlated with their willingness to participate in a task that involved food rewards. On the other hand, capuchins are naturally curious and readily investigate novel situations (Fragaszy & Adams-Curtis, 1991; Visalberghi & Guidi, 1998). Thus, their motivation to engage in innovative behaviour could be underpinned by personality traits like curiosity, exploration, persistence, or neophobia (Benson-Amram et al., 2013; Daniels et al., 2019; Kidd & Hayden, 2015; Overington, Cauchard, et al., 2011).

At least three possible scenarios could explain the negative association between capuchins' age and innovative behaviour. First, younger, smaller-bodied capuchins may not possess the necessary physical strength and dentition that older, larger-bodied capuchins have, which in turn could make innovations more necessary for them (Kummer & Goodall, 1985; Reader & Laland, 2001). Second, older capuchins may be less innovative due to age-related decreases in general playfulness and objective manipulation compared to younger individuals, which may reduce their probability of making innovative "discoveries" (Visalberghi & Guidi, 1998). Third, ageing may place constraints on innovative behaviour due to age-related neurological decline (Massimiliano, 2015; Roskos-Ewoldsen et al., 2008; Zwoinska et al., 2017).

While sex differences in psychological traits, including those related to innovativeness, have been reported in various birds and mammals (Amici et al., 2019; Boogert et al., 2011; Reader & Laland, 2001), we found no evidence of a significant and independent effect of sex on innovation within brown capuchins. Again, these findings are similar to those reported in white-faced capuchins whereby males and females show minimal differences in innovation (Perry et al., 2017). It is unclear why some species show sex differences in innovation while others do not, and so further studies are needed.

Implications for other species

Cross-species comparisons using the same or similar methods will help with modelling (in relative terms) how different factors shape innovation throughout the animal kingdom. Beyond capuchins, observer ratings have been used to study the behaviour of many other animals, such as other primates (Freeman & Gosling, 2010), horses (*Equus ferus*) (Lloyd et al., 2008), hyenas (*Crocuta crocuta*) (Gosling, 1998), cats (*Felis spp.*) (Gartner et al., 2014), deer (*Dama dama*) (Bergvall et al., 2011), and elephants (*Loxodonta africana* and *Elephas maximus*) (Lee & Moss, 2012; Seltmann et al., 2018). Researchers may therefore benefit from testing the validity of ratings to study innovative behaviour in these and other species. Such studies should consider using different items for innovation across specific domains (e.g. foraging, social, play, and others), and – for group-living species – specify within the definitions of those items that “new behaviours” should be new to the entire group, not just the individual.

Conclusions

Due to the logistical difficulties of conducting large-scale observational studies of animal innovation, we took a different approach using a large dataset of reliable ratings to study the innovative behaviour of brown capuchins. Ratings were valid predictors of real-world behavioural outcomes within a subsample of these capuchins, and factors associated with innovative behaviour across our *whole* sample were consistent with observations on wild capuchins. Observer ratings may therefore provide researchers with a valid approach to studying innovation in capuchins and, perhaps, other species as well.

References

Amici, F., Widdig, A., & Majolo, B. (2019). A meta-analysis of interindividual differences in innovation *Animal Behaviour, 155*, 257-268. https://doi.org/https://doi.org/10.1016/j.anbehav.2019.07.008

Aplin, L. M., & Morand-Ferron, J. (2017). Stable producer–scrounger dynamics in wild birds: sociability and learning speed covary with scrounging behaviour. *Proceedings of the Royal Society B: Biological Sciences, 284*(1852), 20162872. https://doi.org/https://doi.org/10.1098/rspb.2016.2872

Benson-Amram, S., Weldele, M. L., & Holekamp, K. E. (2013). A comparison of innovative problem-solving abilities between wild and captive spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour, 85*(2), 349-356. https://doi.org/https://doi.org/10.1016/j.anbehav.2012.11.003

Bergvall, U. A., Schapers, A., Kjellander, P., & Weiss, A. (2011). Personality and foraging decisions in fallow deer, *Dama dama*. *Animal Behaviour, 81*(1), 101-112. https://doi.org/https://doi.org/10.1016/j.anbehav.2010.09.018

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., & Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition, 6*(4), 213-223. https://doi.org/https://doi.org/10.1007/s10071-003-0183-x

Boesch, C. (1995, Feb). Innovation in wild chimpanzees (*Pan troglodytes*). *International Journal of Primatology, 16*(1), 1-16. https://doi.org/https://doi.org/10.1007/BF02700150

Boogert, N. J., Fawcett, T. W., & Lefebvre, L. (2011). Mate choice for cognitive traits: a review of the evidence in nonhuman vertebrates *Behavioral Ecology, 22*(3). https://doi.org/https://doi.org/10.1093/beheco/arq173

Brosnan, S. F., & Hopper, L. M. (2014). Psychological limits on animal innovation. *Animal Behaviour, 92*, 325-332. https://doi.org/https://doi.org/10.1016/j.anbehav.2014.02.026

Conover, J. W. (1971). *Practical nonparametric statistics*. John Wiley & Sons.

Daniels, S. E., Fanelli, R. E., Gilbert, A., & Benson-Amram, S. (2019). Behavioral flexibility of a generalist carnivore. *Animal Cognition, 22*, 387-396. https://doi.org/https://doi.org/10.1007/s10071-019-01252-7

Daveri, F., & Parisi, M. L. (2015). Experience, innovation, and productivity: empirical evidence from Italy's slowdown. *ILR Review, 68*(4), 889-915. https://doi.org/https://doi.org/10.1177/0019793915586971

Ducatez, S., Clavel, J., & Lefebvre, L. (2015). Ecological generalism and behavioural innovation in birds: technical intelligence or the simple incorporation of new foods? *Journal of Animal Ecology, 84*(1), 79-89. https://doi.org/https://doi.org/10.1111/1365-2656.12255

Ferreira, R. G., Izar, P., & Lee, P. C. (2006). Exchange, affiliation, and protective interventions in semifree-ranging brown capuchin monkeys (*Cebus apella*). *American Journal of Primatology, 68*(8), 765-776. https://doi.org/https://doi.org/10.1002/ajp.20277

Fragaszy, D. M., & Adams-Curtis, L. E. (1991). Generative aspects of manipulation in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 105*(4), 387-397. https://doi.org/https://doi.org/10.1037/0735-7036.105.4.387

Fragaszy, D. M., visalberghi, E., & Fedigan, L. M. (2004). *The complete capuchin: The biology of the genus Cebus*. Cambridge University Press.

Freeman, H. D., Brosnan, S. F., Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Gosling, S. D. (2013). Developing a comprehensive and comparative questionnaire for measuring personality in chimpanzees using a simultaneous top-down/bottom-up design. *American Journal of Primatology, 75*(10), 1042-1053. https://doi.org/https://doi.org/10.1002/ajp.22168

Freeman, H. D., & Gosling, S. D. (2010). Personality in nonhuman primates: A review and evaluation of past research. *American Journal of Primatology, 72*(8), 653-671. https://doi.org/https://doi.org/10.1002/ajp.20833

Gammell, M. P., De Vries, H., Jennings, D. J., Carlin, C. M., & Hayden, T. J. (2003). David's score: a more appropriate dominance ranking method than Clutton-Brock et al.'s index. *Animal Behaviour, 66*(3), 601-605. https://doi.org/https://doi.org/10.1006/anbe.2003.2226

Gartner, M. C., Powell, D. M., & Weiss, A. (2014). Personality Structure in the Domestic Cat (*Felis silvestris catus*), Scottish Wildcat (*Felis silvestris grampia*), Clouded Leopard (*Neofelis nebulosa*), Snow Leopard (*Panthera uncia*), and African Lion (*Panthera leo*): A Comparative Study. *Journal of Comparative Psychology, 128*(4), 414-426. https://doi.org/https://doi.org/10.1037/a0037104

Giraldeau, L. A., & Lefebvre, L. (1987). Scrounging prevents cultural transmission of food-finding behavior in pigeons. *Animal Behaviour*, *35*(2), 387-394. https://doi.org/https://doi.org/10.1016/S0003-3472(87)80262-2

Gosling, S. D. (1998). Personality dimensions in spotted hyenas (*Crocuta crocuta*). *Journal of Comparative Psychology*, *112*(2), 107-118. https://doi.org/https://doi.org/10.1037/0735-7036.112.2.107

Griffin, A. S., & Guez, D. (2014). Innovation and problem solving: A review of common mechanisms. *Behavioural Processes*, *109*(Part B), 121-134. https://doi.org/https://doi.org/10.1016/j.beproc.2014.08.027

Griffin, A. S., Guez, D., Lermite, F., & Patience, M. (2013). Tracking changing environments: innovators are fast, but not flexible learners. *PLoS ONE*, *8*(12), e84907.

Guidelines for the treatment of animals in behavioural research and teaching. (2012). *Animal Behaviour*, *83*(1), 301-309. https://doi.org/https://doi.org/10.1016/j.anbehav.2011.10.031

Haslam, M., Hernandez-Aguilar, A., Ling, V., Carvalho, S., de la Torre, I., DeStefano, A., Du, A., Hardy, B., Harris, J., Marchant, L., Matsuzawa, T., McGrew, W., Mercader, J., Mora, R., Petraglia, M., Roche, H., Visalberghi, E., & Warren, R. (2009). Primate archaeology. *Nature*, *460*, 339-344. https://doi.org/https://doi.org/10.1038/nature08188

Henke-von der Malsburg, J., & Fichtel, C. (2018). Are generalists more innovative than specialists? A comparison of innovative abilities in two wild sympatric mouse lemur species. *Royal Society Open Science, 5*(8), 180480. https://doi.org/https://doi.org/10.1098/rsos.180480

Hopper, L. M., Price, S. A., Freeman, H. D., Lambeth, S. P., Schapiro, S. J., & Kendal, R. L. (2014). Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success. *Animal Cognition, 17*(4), 835-847. https://doi.org/https://doi.org/10.1007/s10071-013-0715-y

Huebner, F., & Fichtel, C. (2015). Innovation and behavioral flexibility in wild redfronted lemurs (*Eulemur rufifrons*). *Animal Cognition, 18*(3), 777-787. https://doi.org/https://doi.org/10.1007/s10071-015-0844-6

Kalbitzer, U., Bergstrom, M. L., Carnegie, S. D., Wikberg, E. C., Kawamura, S., Campos, F. A., Jack, K. M., & Fedigan, L. M. (2017). Female sociality and sexual conflict shape offspring survival in a Neotropical primate. *Proceedings of the National Academy of Sciences of the United States of America, 114*(8), 1892-1897. https://doi.org/https://doi.org/10.1073/pnas.1608625114

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron, 88,* 449-460. https://doi.org/https://doi.org/10.1016/j.neuron.2015.09.010

Kummer, H., & Goodall, J. (1985). Conditions of innovative behavior in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 308,* 203-214.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/https://doi.org/10.18637/jss.v082.i13

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence, 41*(5), 306-320. https://doi.org/https://doi.org/10.1016/j.intell.2013.04.008

Lee, P. C. (1991). Adaptations to environmental change: an evolutionary perspective. In H. O. Box (Ed.), *Primate responses to environmental change* (pp. 39-56). Springer. https://doi.org/https://doi.org/10.1007/978-94-011-3110-0_2

Lee, P. C., & Moss, C. J. (2012). Wild female African elephants (*Loxodonta africana*) exhibit personality traits of leadership and social integration. *Journal of Comparative Psychology, 126*(3), 224-232. https://doi.org/https://doi.org/10.1037/a0026566

Lee, P. C., & Moura, A. C. D. (2015). Necessity, unpredictability and opportunity: an exploration of ecological and social drivers of behavioural innovation. In A. Kaufman & J. Kaufman (Eds.), *Animal Creativity and Innovation* (pp. 317-329). Elsevier. http://store.elsevier.com/Animal-Creativity-and-Innovation/isbn-9780128006481/

Lefebvre, L. (2013). Brains, innovations, tools and cultural transmission in birds, non-human primates, and fossil hominins. *Frontiers in Human Neuroscience, 7*, 245. https://doi.org/https://doi.org/10.3389/fnhum.2013.00245

Lefebvre, L., Reader, S. M., & Sol, D. (2004). Brains, innovations and evolution in birds and primates. *Brain Behavior and Evolution, 63*, 233-246. https://doi.org/https://doi.org/10.1159/000076784

Liker, A., & Bokony, V. (2009). Larger groups are more successful in innovative problem solving in house sparrows. *Proceedings of the National Academy of Sciences of the*

United States of America, 106(19), 7893-7898. https://doi.org/https://doi.org/10.1073/pnas.0900042106

Lloyd, A. S., Martin, J. E., Bornett-Gauci, H. L. I., & Wilkinson, R. G. (2008). Horse personality: Variation between breeds. *Applied Animal Behaviour Science, 112*(3-4), 369-383. https://doi.org/https://doi.org/10.1016/j.applanim.2007.08.010

Martin, P., & Bateson, P. P. G. (2007). *Measuring behaviour : an introductory guide* (3rd ed.). Cambridge University Press.

Massimiliano, P. (2015). The effects of age on divergent thinking and creative objects production: a cross-sectional study. *High Ability Studies, 26*(1135), 93-104. https://doi.org/https://doi.org/10.1098/rstb.1985.0020

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90. https://doi.org/https://doi.org/10.1037/0022-3514.52.1.81

Morton, F. B., Lee, P. C., & Buchanan-Smith, H. M. (2013). Taking personality selection bias seriously in animal cognition research: a case study in capuchin monkeys (*Sapajus apella*). *Animal Cognition, 16*, 677-684. https://doi.org/https://doi.org/10.1007/s10071-013-0603-5

Morton, F. B., Lee, P. C., Buchanan-Smith, H. M., Brosnan, S. F., Thierry, B., Paukner, A., de Waal, F. B. M., Widness, J., Essler, J. L., & Weiss, A. (2013). Personality Structure in Brown Capuchin Monkeys (*Sapajus apella*): Comparisons With Chimpanzees (*Pan troglodytes*), Orangutans (*Pongo spp.*), and Rhesus Macaques (*Macaca mulatta*). *Journal of Comparative Psychology, 127*(3), 282-298. https://doi.org/https://doi.org/10.1037/a0031723

Morton, F. B., Weiss, A., Buchanan-Smith, H. M., & Lee, P. C. (2015). Capuchin monkeys with similar personalities have higher-quality relationships independent of age, sex, kinship and rank. *Animal Behaviour*, *105*, 163-171. https://doi.org/https://doi.org/10.1016/j.anbehav.2015.04.013

Moura, A. C. D., & Lee, P. C. (2004). Capuchin stone tool use in Caatinga dry forest. *Science*, *306*(5703), 1909-1909. https://doi.org/https://doi.org/10.1126/science.1102558

Ostner, J., & Schulke, O. (2018). Linking sociality to fitness in primates: A call for mechanisms. *Advances in the Study of Behavior*, *50*, 127-175. https://doi.org/https://doi.org/10.1016/bs.asb.2017.12.001

Overington, S. E., Cauchard, L., Cote, K. A., & Lefebvre, L. (2011). Innovative foraging behaviour in birds: What characterizes an innovator? *Behavioural Processes*, *87*(3), 274-285. https://doi.org/https://doi.org/10.1016/j.beproc.2011.06.002

Overington, S. E., Griffin, A. S., Sol, D., & Lefebvre, L. (2011). Are innovative species ecological generalists? A test in North American birds. *Behavioral Ecology*, *22*(6), 1286-1293. https://doi.org/https://doi.org/10.1093/beheco/arr130

Perry, S. E., Barrett, B. J., & Godoy, I. (2017). Older, sociable capuchins (*Cebus capucinus*) invent more social behaviors, but younger monkeys innovate more in other contexts. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7806-7813. https://doi.org/https://doi.org/10.1073/pnas.1620739114

Ramsey, G., Bastian, M. L., & van Schaik, C. (2007). Animal innovation defined and operationalized. *Behavioral and Brain Sciences, 30*(4), 393-407. https://doi.org/https://doi.org/10.1017/S0140525X07002373

Reader, S. M. (2003). Innovation and social learning: individual variation and brain evolution. *Animal Biology, 53*(2), 147-158. https://doi.org/https://doi.org/10.1163/157075603769700340

Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1567), 1017-1027. https://doi.org/https://doi.org/10.1098/rstb.2010.0342

Reader, S. M., & Laland, K. N. (2001). Primate innovation: Sex, age and social rank differences. *International Journal of Primatology, 22*(5), 787-805. https://doi.org/https://doi.org/10.1023/A:1012069500899

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the United States of America, 99*(7), 4436-4441. https://doi.org/https://doi.org/10.1073/pnas.062041299

Reader, S. M., & Laland, K. N. (2003). *Animal innovation*. Oxford University Press. Table of contents only http://www.loc.gov/catdir/toc/fy044/2004296817.html

Publisher description http://www.loc.gov/catdir/enhancements/fy0620/2004296817-d.html

Roskos-Ewoldsen, B., Black, S. R., & McCown, S. M. (2008). Age-related changes in creative thinking. *Journal of Creative Behavior, 42*(1), 33-59. https://doi.org/https://doi.org/10.1002/j.2162-6057.2008.tb01079.x

Royston, P. (1982). Algorithm AS 181: The W test for normality. *Applied Statistics, 31*(2), 176-180. https://doi.org/https://doi.org/10.2307/2347986

Seltmann, M., Helle, S., Adams, M. J., Mar, K. U., & Lahdenpera, M. (2018). Evaluating the personality structure of semi-captive Asian elephants living in their natural habitat. *Royal Society Open Science, 5*(2). https://doi.org/https://doi.org/10.1098/rsos.172026

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bulletin, 86*(2), 420-428. https://doi.org/https://doi.org/10.1037//0033-2909.86.2.420

Silk, J. B. (2007). The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 539-559. https://doi.org/https://doi.org/10.1098/rstb.2006.1994

Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science, 302*(5648), 1231-1234. https://doi.org/https://doi.org/10.1126/science.1088580

Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., Wittig, R. M., Seyfarth, R. M., & Cheney, D. L. (2009). The benefits of social capital: close social bonds among female baboons enhance offspring survival. *Proceedings of the Royal Society B: Biological Sciences, 276*(1670), 3099-3104. https://doi.org/https://doi.org/10.1098/rspb.2009.0681

Skinner, B. F. (1938). *The behavior of organisms*. Appleton-Century-Crofts.

Team, R. C. (2019). *R: A language and environment for statistical computing.* In (Version Development version: "Unsuffered Consequences" compiled on 2019-08-05) R Foundation for Statistical Computing. https://www.r-project.org/

Tian, M., Deng, P., Zhang, Y. Y., & Salmador, M. P. (2018). How does culture influence innovation? A systematic literature review. *Management Decision, 56*(5), 1088-1107. https://doi.org/https://doi.org/10.5465/AMBPP.2018.15506abstract

Tiddi, B., Aureli, F., & Schino, G. (2012). Grooming up the hierarchy: The exchange of grooming and rank-related benefits in a New World primate. *PLoS ONE, 7*(5), e36641. https://doi.org/https://doi.org/10.1371/journal.pone.0036641

van Horik, J. O., & Madden, J. R. (2016). A problem with problem solving: motivational traits, but not cognition, predict success on novel operant foraging tasks. *Animal Behaviour, 114,* 189-198. https://doi.org/https://doi.org/10.1016/j.anbehav.2016.02.006

Visalberghi, E., & Guidi, C. (1998). Play behaviour in young tufted capuchin monkeys. *Folia Primatologica, 69*(6), 419-422. https://doi.org/https://doi.org/10.1159/000021665

Weiss, A., Adams, M. J., Widdig, A., & Gerald, M. S. (2011). Rhesus macaques (*Macaca mulatta*) as living fossils of hominoid personality and subjective well-being. *Journal of Comparative Psychology, 125*(1), 72-83. https://doi.org/https://doi.org/10.1037/a0021187

Weiss, A., Inoue-Murayama, M., Hong, K. W., Inoue, E., Udono, T., Ochiai, T., Matsuzawa, T., Hirata, S., & King, J. E. (2009). Assessing chimpanzee personality and subjective well-being in Japan. *American Journal of Primatology*, *71*(4), 283-292. https://doi.org/https://doi.org/10.1002/ajp.20649

Weiss, A., Inoue-Murayama, M., King, J. E., Adams, M. J., & Matsuzawa, T. (2012). All too human? Chimpanzee and orang-utan personalities are not anthropomorphic projections. *Animal Behaviour, 83*(6), 1355-1365. https://doi.org/https://doi.org/10.1016/j.anbehav.2012.02.024

Zwoinska, M. K., Maklakov, A. A., Kawecki, T. J., & Hollis, B. (2017). Experimental evolution of slowed cognitive aging in *Drosophila melanogaster*. *Evolution*, *71*(3), 662-670. https://doi.org/https://doi.org/10.1111/evo.13156