



# Assessments made easier: examining the use of a rating-based questionnaire to capture behavioral data in rehabilitant orangutans (*Pongo pygmaeus morio*)

Florence Rocque<sup>1</sup> · Helene Chotard<sup>1</sup> · Kim Bard<sup>1</sup> · Jerome Micheletta<sup>1</sup> · Augustine Tuuga<sup>2</sup> · Sylvia Alsisto<sup>2</sup> · Marina Davila-Ross<sup>1</sup>

Received: 23 July 2021 / Accepted: 8 May 2022  
© The Author(s) 2022

## Abstract

Rehabilitation and release are commonly used for confiscated, surrendered, and rescued primates. To improve release efficacy it is important to generate accurate behavioral profiles of release candidates. Research on primates traditionally uses observer ratings to measure individual differences. This method is easily implemented, but its validity has been questioned. We evaluated whether observer ratings reflect behavioral data indicating forest adaptation in 18 free-ranging rehabilitant orangutans (*Pongo pygmaeus morio*). In 2017, we used a species-specific questionnaire to measure how often orangutans engaged in behaviors linked to living successfully in the wild (e.g., nest building) and the extent to which they express personality traits that may influence forest adaptation. We collected 11 months of observational data on 17 of the orangutans concurrently to validate the questionnaire items, and collected further questionnaire data for 16 of the individuals in 2019. We used regularized exploratory factor analysis (REFA) and parallel analysis to condense the ratings and determine that two factors could be reliably extracted. We conducted another REFA using the observational data, and calculated factor congruence coefficients following procrustean rotation. The first of the two factors represented forest skills and human aversion, and was congruent with observational data. The second factor reflected boldness, sociability, and exploration, and was not congruent with observational data. Ratings correlated significantly with observations for all five questionnaire items reflecting adaptation to forest life, and for three of seven items reflecting personality traits. We conclude that ratings can be a valid approach to obtain individual-based behavioral information reflecting forest adaptation in free-ranging rehabilitant orangutans, and may be particularly useful in summarizing behaviors relevant to forest adaptation that are otherwise challenging to gather in primates.

**Keywords** Rating methodology · Individual differences · Rehabilitation · Forest adaptation

---

Handling Editor: Joanna Setchell.

Extended author information available on the last page of the article

## Introduction

Animals in rehabilitation centers undergo behavioral and ecological rehabilitation to help them acquire the skills necessary to survive unsupported in the wild (Beck et al., 2007; Russon, 2009). When releases fail, it may indicate these skills have not been acquired adequately (Soorae, 2010). Differences in personality traits may also influence rehabilitation and release outcome for individuals (de Azevedo & Young, 2021; Powell & Gartner, 2011). For instance, measures of boldness corresponded with increased mortality in swift foxes (*Vulpes velox*: Bremner-Harrison et al., 2004) and wood mice (*Apodemus sylvaticus*: Stratton, 2015), but decreased mortality in Tasmanian devil release attempts (*Sarcophilus harrisi*: Sinn et al., 2014). Thus, data describing the rehabilitants' behavior are required to make informed decisions about the skills training individuals require, and about their release (Beck et al., 2007; Russon, 2009). Resources are often limited, and a systematic, validated method to obtain such information from experienced staff familiar with these animals would be beneficial when observational or experimental data cannot be collected quickly. However, although some rehabilitation centers have developed inventories to collect pre-release data (Russon, 2009), there is no standardized assessment method. By improving our understanding of what kind of questions generate valid ratings, we can develop questionnaires to assess rehabilitation and help to determine which individual characteristics influence rehabilitation success. Such methods could benefit rehabilitation programs by enabling them to share and compare data on animal behavior more easily.

Individual differences in animal behavior can be measured using three methods: behavioral observations, experiments, and ratings (Błaszczuk, 2020; Weiss, 2017). Of these, rating data can be obtained relatively quickly and easily (King et al., 2005), while behavioral observations and experiments may be more time consuming and challenging to implement (de Azevedo & Young, 2021; Freeman & Gosling, 2010). Experienced caregivers have provided reliable behavioral ratings for a wide range of animal species (e.g., primates: Iwanicki & Lehmann, 2015; rodents: Baker et al., 2016; and cetaceans: Úbeda et al., 2019). However, some researchers question the subjectivity of such ratings, as caretakers may anthropomorphize the animals studied (Uher, 2008), attributing characteristics to them which may not have been observed, or may be influenced by notable events and inflate scores for behaviors which are not typically expressed (Freeman & Gosling, 2010). For instance, a rater's perception of an animals' tendency to be aggressive may be inflated if they recently observed it engage in a fight, even if this is unusual behavior for the individual.

Among primates, rating instruments have been validated via significant correlations between ratings for questionnaire items and observed behaviors [chimpanzees (*Pan troglodytes*), gorillas (*Gorilla gorilla*), bonobos (*Pan paniscus*): Murray, 2011] and between behavioral constructs determined from such ratings and observed behaviors [macaques: (*Macaca mulatta*) Capitanio, 1999; chimpanzees: Freeman et al., 2013, Pederson et al., 2005; gorillas: Eckardt et al., 2015]. However, ratings may be more valid for questionnaire items describing specific behaviors than for items describing behaviors in more general terms. For example, to determine

curiosity, ratings for the item “individual often touches new objects (e.g., enrichment object)” were more consistent with experimentally assessed behaviors than adjective-based items (e.g., “individual is curious”) in orangutans, gorillas, bonobos, and chimpanzees (Uher & Asendorpf, 2008).

Orangutans are one of the most frequently released species, with growing numbers being cared for by rehabilitation centres across their remaining home ranges in Asia (Brent, 2007; Russon, 2009). It is important for successful rehabilitation that orangutans can adapt to independent life in the forest (Russon, 2009). As skill acquisition is integral to this process, individual adaptation may be impacted by traits such as sociability, via social information transmission (Dindo et al., 2010), and exploration, via independent learning, whereby more exploratory individuals may learn quicker about their surroundings (Germano et al., 2017). Such traits may be particularly relevant in a rehabilitation setting, as there is no further potential for maternal skill transfer.

Skills that orangutans are encouraged to develop throughout rehabilitation, such as tree climbing and nest building (Basalamah et al., 2018; Grundmann, 2006; Riedler et al., 2010; Russon, 2009), correspond with successful adaptation to the forest, and it is important to document them during rehabilitation. Terrestriality should also be considered as a factor in successful rehabilitation, because if rehabilitants spend too much time on the ground they may increase their chances of encountering parasites or predators and have reduced opportunities for species-typical foraging and nest building (Rijksen, 1978; Russon, 2009). For some individuals, choosing to stay more on the ground may be a consequence of learning from human caregivers (Grundmann, 2006; Riedler et al., 2010).

Displaying a species-typical ranging pattern may also be important in release success. When rehabilitation programs employ a ‘soft release’ approach, human caregivers provide supplementary food to rehabilitants, fostering gradual acclimatization to the forest (de Azevedo & Young, 2021), but fixed supplementary feeding platforms may restrict ranging for some rehabilitants if they maintain proximity to this food resource (Beck et al., 2007; Snaith, 1999). Orangutans who disperse further into the forest, and away from rehabilitation centers which can be congested with lots of individuals, may encounter more fruits, which are highly nutritious but scattered across wide areas and often only temporally available (Riedler et al., 2010; Russon, 2009). However, by feeding on supplementary food, rehabilitants avoid the risks of eating toxic foods, engaging in inexperienced locomotory techniques, or malnourishment (Kuze et al., 2008; Snaith, 1999).

It is also important to determine how interested individuals are in approaching humans when assessing rehabilitation (Riedler et al., 2010). Successfully released orangutans should not interact with humans (Russon, 2009), but human involvement in care at rehabilitation centres is often unavoidable (Beck et al., 2007). Orangutans, like most great apes, require care for emotional, as well as physical needs (Bard & Hopkins, 2018) and humans can provide the emotional care required for healthy infant development (Beck et al., 2007; Clay et al., 2015). However, the goals of providing appropriate support for infants earlier on must be balanced with helping them to gain independence and with minimizing human orientation later on in the process (Russon, 2009).

A valid approach to generate accurate behavioral profiles of release candidates is important to improve release efficacy (Beck et al., 2007; Russon, 2009). We collected questionnaire data from orangutan caregivers at a rehabilitation centre to test whether these ratings can generate valid behavioral profiles. We included items (questions) about species-specific behaviors indicating successful adaptation to the forest. We also included items reflecting personality traits based on predicted ecological significance (Bremner-Harrison et al., 2004; Uher & Asendorpf, 2008; Wolf & Weissing, 2012) to examine whether they corresponded with items directly related to forest adaptation. We assessed interrater reliability between caregivers and used this to help reduce the number of questions in our questionnaire, making it quicker to implement. We used dimension reduction analysis to determine which behaviors are most associated with one another and generate a behavioral construct that could be used to assess individuals. We assessed rating validity by determining whether ratings clustered similarly to observations of equivalent behaviors following data reduction, and whether ratings were correlated with these observations at a question level. If the questionnaire generates valid data on orangutan behavior in the forest, then we predict that ratings for individual items and factors relating to forest adaptation will correlate positively with observational data reflecting equivalent behaviors. As rating stability may also be important for making accurate post-release predictions in the future, we also explore whether ratings change over time, predicting that personality traits will be consistent over time (Wolf & Weissing, 2012) but that forest skills (nest building, tree climbing, etc.) will improve as rehabilitation progresses and the orangutans get older and more experienced (Riedler et al., 2010; Russon, 2009).

## Methods

### Orangutans and study site

We studied orangutans at Sepilok Orangutan Rehabilitation Centre (SORC), at the edge of the 43 km<sup>2</sup> Kabili Sepilok Forest Reserve, Sabah, Malaysia (Borneo). We included all free-ranging orangutans who were available at the time of data collection in 2017 and 2019. In 2017 we collected questionnaire data on 18 free-ranging rehabilitant northeast Bornean orangutans (*Pongo pygmaeus morio*): ten males and eight females, aged 5–17 years, half of the 18 subjects were 8+ years old. We also collected observational data for 17 of these subjects to validate the questionnaire. In 2019 we collected further questionnaire data on 16 of the original 18 orangutans: eight males and eight females, aged 7–19 years old. None of the subjects were related to one another or had mothers at the centre.

Depending on their age and experience, following veterinary assessment, orangutans at SORC are placed into one of three rehabilitation stages, consistent with other rehabilitation sites (Russon, 2009). Orangutans in Stage 1 of rehabilitation are typically infants and less experienced. They are allowed some supervised forest access each day and are cared for by local staff and local and international volunteers. SORC's staff care work included feeding orangutans daily at the center and forest

feeding platforms, supervising tourists and volunteers and weekly health checks for orangutans. Individuals in Stage 2 and 3 are juveniles (4–8 years), adolescents (8–11 years) and adults (11+ years) (Russon, 2006; Russon et al., 2016) and are allowed unrestricted ranging in the forest adjoining SORC. Caretakers encourage individuals in Stage 2 to return to the protection of night cages and to return to SORC between 11:00 and 14:30 for supplementary food. Individuals in Stage 3 are not given access to night cages but are offered supplementary food if present and are housed in cages at the veterinary clinic at SORC if they require medical care. All orangutans in this study were either in Stage 2 ( $n = 5$ ) or stage 3 ( $n = 13$ ) of rehabilitation in 2017, and as such, in a transitional phase prior to final release. In 2019, one of the Stage 2 individuals moved to Stage 3; this individual was not included in any analysis of rehabilitation stage effects.

Caretakers provided three buckets of fruit and vegetables across two feeding platforms during morning (09:30–10:30) and afternoon sessions (14:30–15:30). These feeding platforms at the center were situated approximately 170 m from one another and could be visited by other local wildlife, and by tourists who were permitted to observe the feedings from nearby viewing platforms. Individuals in the first stage of rehabilitation were in cages during feeding times, and our study animals also had access to any food that dropped below the cages. The infant cages were inaccessible to tourists and located further away from the forest than the night cages for orangutans in Stage 2.

## Data collection

**Rating questionnaire** We designed the questions based on previous rating schemes assessing personality in non-human primates that yielded reliable data across raters (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016), primarily Chotard (2020), and aiming to reflect orangutan behavior and adaptation to the forest. A key reason for developing the questionnaire on rehabilitant orangutans from SORC caregivers was for SORC to obtain systematic data on the individual orangutans and, consequently, to help wildlife authorities to make more informed decisions about releases. Afterwards, SORC and the research team decided to develop this study to systematically assess the application value of SORC's data.

In the questionnaire, we included items reflecting personality characteristics we deemed relevant for ecological rehabilitation; boldness, sociability, and exploration (Allard et al., 2019; Biro & Stamps, 2008; Dingemanse et al., 2004; Wolf & Weissing, 2012), and behaviors that caregivers could observe at the center that reflect forest-living skills and that have been used to assess forest adaptation in orangutan releases (Basalamah et al., 2018; Grunmann 2006; Riedler et al., 2010; Snaith, 1999). There were 29 items in the full questionnaire set, scored on a Likert scale from 1 (almost never) to 5 (very often); see Table 1. We administered the questionnaires in 2017 and 2019 to test for temporal consistency. Raters provided answers across all contexts in which they observed the orangutans unless specified in the questionnaire item, for example 'when eating next to stronger orangutans, takes food from nearby them' refers specifically to feeding.

**Table 1** Interrater reliability and temporal consistency for items in a questionnaire to describe the behavior of *Pongo pygmaeus morio* at Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia, in 2017 and 2019

Questionnaire items	Interrater reliability (ICC 1,1)		Temporal stability 2017–2019 (ICC 1,1)	Rating per item 2017 ( <i>n</i> = 18)		Rating per item 2019 ( <i>n</i> = 16)	
	2017 (4 raters)	2019 (5 raters)		Mean	SD	Mean	SD
	<i>P</i>			Mean	SD	Mean	SD
Climbs on trees <sup>b</sup>	0.584	0.475	<b>0.726</b>	3.5	0.9	3.8	0.8
Has good climbing skills	0.411	0.470	<b>0.792</b>	4.1	0.5	4.2	0.7
Stays on the ground <sup>b</sup>	0.186	0.662	<b>0.599</b>	3.2	0.6	3.4	0.8
Builds nests <sup>b</sup>	0.574	0.387	<b>0.575</b>	2.8	0.9	3.0	0.9
Builds steady nests	0.217	0.271	<b>0.513</b>	3.0	0.9	3.3	0.8
Stays in the forest all day <sup>b</sup>	0.526	0.419	<b>0.701</b>	3.1	0.8	3.0	0.8
Stays in the forest for a week or longer	0.635	0.407	<b>0.490</b>	1.8	1.1	2.6	0.9
Stays at the centre (even when out of the cage)	0.236	0.433	<b>0.439</b>	3.7	0.6	3.9	0.6
Stays in the centre during the day	0.545	0.331	0.000	3.9	0.6	4.0	0.6
Is feeding at the feeding platforms	0.089	0.087	0.333	3.6	0.4	3.6	0.4
Approaches and tries to touch staff <sup>b</sup>	0.532	0.557	<b>0.604</b>	2.7	0.9	2.6	0.9
Approaches and tries to touch volunteers or tourists	0.544	0.665	<b>0.778</b>	2.8	0.9	2.5	0.9
Keeps a distance from tourists	0.170	0.556	<b>0.545</b>	3.4	0.6	3.1	0.7
Watches staff	0.243	0.486	0.020	3.3	0.6	3.2	0.5
Watches tourists.	0.297	0.551	<b>0.552</b>	3.2	0.7	2.9	0.9
Is bold towards humans <sup>b</sup>	0.437	0.542	<b>0.476</b>	3.4	0.7	3.0	0.8
When eating next to stronger orangutans, takes food from nearby them <sup>a</sup>	0.542	0.239	0.260	2.9	0.7	2.6	0.6
When playing with stronger orangutans, plays rough with them	0.550	0.239	0.401	2.6	0.8	3.1	0.6
Is cheeky	0.641	0.481	<b>0.529</b>	3.0	1.0	3.1	0.5
Is bold <sup>b</sup>	0.274	0.287	<b>0.544</b>	3.2	0.6	3.3	0.5
Is shy towards other orangutans	0.084	0.263	0.057	3.1	0.5	2.6	0.6
Is friendly to other orangutans <sup>a</sup>	0.561	0.038	0.320	3.3	0.7	3.5	0.5

Table 1 (continued)

Questionnaire items	Interrater reliability (ICC 1,1)		Temporal stability 2017–2019 (ICC 1,1)		Rating per item 2017 (n = 18)		Rating per item 2019 (n = 16)	
	2017 (4 raters)	2019 (5 raters)	(16 subjects)	P	Mean	SD	Mean	SD
	Is social	0.537	0.000	<b>0.440</b>	<b>0.035</b>	3.5	0.7	3.5
When relaxing, is nearby other orangutans <sup>a</sup>	0.485	0.053	0.379	0.062	3.3	0.7	3.4	0.3
During feeding, sits nearby other orangutans	0.283	0.050	0.212	0.201	3.6	0.5	3.6	0.4
If there is new food (like a blue apple), walks to it to touch, sniff, or lick it	0.000	0.000	0.000	0.550	3.6	0.5	4.1	0.3
If there is a new object (for example: new ropes) touches it <sup>a</sup>	0.480	0.206	0.209	0.205	3.2	0.8	3.6	0.5
Is eager to learn	0.002	0.111	<b>0.543</b>	<b>0.010</b>	3.5	0.5	3.7	0.5
Is curious <sup>a</sup>	0.271	0.000	0.000	0.874	3.6	0.7	3.5	0.5

Questionnaire items included in further analysis are in bold. <sup>a</sup>Indicates items that included data from 2017 only in further analysis. <sup>b</sup>Indicates items that included data from 2017 and 2019 in further analysis. ICC values for items where moderate agreement (0.5) was reached between raters are in italics. ICC values for items that were stable over time are in bold.

The first author read the questionnaires to the raters if they were comfortable understanding English. Otherwise, a local research assistant read the questions in English/Malay with the help of written translations of the items and their meanings into Malay if the caretaker required clarification of any item. All raters responded in English. All raters had experience working with the orangutans (mean: 4.75 years; range: 1–14 years) and each orangutan was rated by three members of staff. All orangutans that were consistently present at the center were rated ( $n = 18$ ). There were seven raters in total; two rated in both years, two rated in 2017, and three in 2019. We asked raters not to discuss their scores with any other member of staff to avoid influencing answers. We administered items in a random order, so that those pertaining to a certain trait were not grouped together.

**Observational data for rating correlation** We collected observational data to validate the questionnaire items using two approaches: (1) live forest focal samples, to capture specific natural behaviors in the forest next to the center, such as nest building, and (2) video recordings at the center, to collect data on more nuanced behavioral actions. For instance, to collect data on boldness — consistent individual differences in risk-taking behavior (Réale et al., 2007) — we coded all risky actions directed towards other orangutans. This included obvious actions such as biting but also more subtle behaviors such as pulling and grabbing. We assessed personality traits around the center to ensure we could capture and record behavior in detail. We assumed that boldness, explorativeness, and sociability recorded around the center would be equivalent to behavior in the forest because personality traits are usually considered consistent across time and context (Koski, 2011; Wolf & Weissing, 2012).

**Live-scored focal samples** The first author collected focal samples on 17 orangutans from the forest and at the rehabilitation center between 06:30 and 18:00 from January to November 2017. A focal sample had to last at least 60 minutes to be included in analysis but was typically 2.5 hours long. We chose the minimum period of 60 minutes to include data from orangutans lost during sampling, and the maximum of 240 minutes to allow us to follow three orangutans for as long as possible each day while still allowing time to search for the next individual. We sampled all individuals in the morning, afternoon, and evening to account for behaviors which occur more at some times of day than at others, but we recorded individuals in Stage 2 of rehabilitation less often in evenings (5% of their recordings were after 16:00) than those in Stage 3 (18% of their recordings were after 16:00) as this is when they could return to sleep in night cages. We collected 646.75 hours of data (individual mean  $\pm$  SD: 38.04  $\pm$  12.18 h). At the start of each week, we selected the first orangutan to follow according to a randomized list; we then observed the next orangutan we encountered if we were lacking data for them at that time of day, if not we would continue searching.

We used all-occurrences sampling while following focal individuals in the forest. We recorded each event of nest building in trees, unsolicited approaches to staff (e.g., not when being called for feeding/treatment), bold actions towards

staff (chasing, biting, stealing from), and touching unusual objects (objects not consistently available to orangutans, such as tools and building materials). We also recorded the focal orangutans' canopy position every 15 minutes using three categories (tree, ground, and high human-made structure) to determine the proportion of time spent in the trees and on the ground. Finally, we recorded the GPS position at the same time, to determine whether the subject was at the center or in the forest.

**Video recordings at the center** We collected video recordings around feeding sessions at the rehabilitation center using focal sampling (Altmann, 1974). We chose orangutans according to a randomized schedule. Recordings lasted for 3 minutes, to provide enough data to code detailed behavioral events for as many orangutans as possible. We collected data at both feeding areas during morning (09:00–11:00) and afternoon (14:00–16:00) feedings, between January and June 2017. We collected 21.14 hours of data from 17 individuals (mean  $\pm$  SD: 1.24  $\pm$  0.26 h per individual). We modified our coding scheme from a scheme developed for chimpanzees (Chotard, 2020). To ensure interrater reliability in the coding, two additional coders coded 15% of the data (mean ICC (3,1) = 0.53, SD = 0.21,  $p < 0.05$ ).

## Statistical analysis

### Interrater reliability and questionnaire reduction

We simplified the 29-item questionnaire for ease of future application. We focused on key questions that we could validate with observational data and that might be more broadly applicable to other rehabilitation centers. We removed redundancies because we validated the selected items later. We retained items reflecting each of the three personality traits. We used ICC values, a widely used reliability index (Koo & Li, 2016), to assist in selecting questions to retain for further analysis. Because different raters rated different orangutans, we used a one-way random effects model, ICC (1,1) whereby both raters and orangutans are regarded as random (Shrout & Fleiss, 1979). Values less than 0.5 are considered indicators of poor reliability, 0.5–0.75 moderate, 0.75–0.9 good and  $> 0.9$  excellent (Koo & Li, 2016). We discarded items when there was no agreement (ICC = 0.000) between raters, and retained those with the highest interrater reliability estimates where there were redundancies.

### Temporal stability

We also used the ICC (1,1) to assess temporal stability in ratings between 2017 and 2019 for the 16 orangutans we studied in both years. When scores for items selected for further analysis had not changed significantly across the years ( $p < 0.05$ ), we used the mean of all the ratings in both years for each individual in subsequent

analysis (Jung & Lee, 2011). This approach to exploratory factor analysis is recommended to obtain more stable factor loadings (Wilson et al., 2018) when sample sizes are limited (Masilkova et al., 2020). We included data from 2017 and 2019 to represent the more stable aspects of the orangutans' behavior and to improve reliability as it combines six rather than three ratings. We used ratings from 2017 only in further analysis for items that were not temporally stable, as this was when we collected observational data.

### Reduction of the rating data

We conducted REFA with varimax rotation on the 12 questionnaire items we retained. We ran a parallel analysis (O'Connor, 2000) to generate eigenvalues from an equivalent random set of data. When eigenvalues generated by the REFA were larger than the values generated by the parallel analysis it indicated that the factor could be extracted reliably from the data (O'Connor, 2000). We used the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity ( $p < 0.01$ ) to make sure that the data were suitable for interpretation. We interpreted only item loadings exceeding |0.4|; we chose this as a conservative figure considering the sample size (Budaev, 2010).

We computed the composite score for each subject on each factor extracted from the REFA, and used Mann–Whitney tests to test the effects of sex and rehabilitation stage on these factor scores. We also used a Kruskal–Wallis test to test for the effect of age group (juvenile, adolescent, and adult) on factor scores. We excluded one individual that changed rehabilitation stage between years and from the rehabilitation stage analysis and three individuals that changed age categories between years from the age analysis.

### Testing the validity of the rating tool

We matched the items with observational data reflecting equivalent behavior to validate the ratings (Table 2).

To examine whether behavioral observations cluster similarly to ratings we conducted a REFA using the observational data and calculated congruence coefficients for the factors following procrustean rotation, a target rotation of factors to provide a statistical estimate of factor similarity (McCrae et al., 1996). We carried out Procrustes rotation using syntax developed by Fischer and Fontaine (2010). Coefficients exceeding 0.85 are considered to indicate fair replicability, and coefficients exceeding 0.95 are considered to indicate good replicability (Lorenzo-Seva & ten Berge, 2006).

To determine whether ratings were valid assessments of behavior at an item level, we used Spearman correlation coefficients (two-tailed) to compare the ratings with observational data (Table 2) because the data were non-normally distributed. We conducted all statistical tests using SPSS Statistics 25 (IBM,

**Table II** Behavioral data chosen to compare with ratings to validate questionnaire items describing the behavior of *Pongo pygmaeus morio* at Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia 2017–2019

Item	Measure of observed behavior
Climbs in trees <sup>a</sup>	Percentage of interval recordings the subject was found in a tree
Builds nests <sup>a</sup>	Rate of nest-building events (per hour)
Stays on the ground <sup>a</sup>	Percentage of interval recordings the subject was found on the ground
Stays in the forest all day <sup>a</sup>	Percentage of GPS recordings the subject was found in the forest surrounding the center
Is bold towards humans <sup>a</sup>	Rate of bold actions towards humans (stealing/chasing/biting (Chotard, 2020) (per hour)
Approaches humans <sup>a</sup>	Rates of unsolicited approaches to humans (per hour)
Is bold <sup>b</sup>	Rate of risky behaviors the subject directed towards a conspecific (8yrs +) who could cause some harm to the subject (i.e., not an infant), including jumping on, hitting, biting, pulling, grabbing, chasing, and pushing (Chotard, 2020; Clay et al., 2015) (per hour)
When eating next to stronger orangutans, takes food from nearby them <sup>b</sup>	Rate of taking food items from within 1 m of a conspecific aged (8yrs +) (per hour)
Is curious <sup>b</sup>	Rate of gazing at, touching, or manipulating objects (indicative of explorative tendency in apes (Damerius et al., 2017; Massen et al., 2013) (per hour)
If there is a new object (for example: new ropes), touches it <sup>a</sup>	Rate of interaction with unusually appearing objects (objects not consistently available to orangutans, such as tools and building materials) (per hour)
Is friendly <sup>b</sup>	Rate of affiliative social actions directed towards a conspecific (Capitanio, 1999; Koski, 2011) (per hour)
When relaxing is nearby other orangutan <sup>b</sup>	Percentage duration subject is resting within 3m of other orangutans

Data from: live-scored forest focal <sup>a</sup>, centre-based video recordings <sup>b</sup>

Chicago, IL, USA) with a significance level of  $p \leq .05$ . No statistical assumptions were violated.

## Ethical note

The Animal Welfare and Ethical Review Body (AWERB) of the University of Portsmouth granted approval for this research. The authors declare that they have no conflict of interest.

## Data availability

**Data Availability** The data sets supporting these findings are not publicly available, in accordance with Sabah government regulation due to the conservation status of the study population, but are available from the corresponding author on reasonable request.

## Results

### Interrater reliability, temporal consistency and selection of questionnaire items

Questionnaire items had ICC (1,1) estimates ranging from 0.000 to 0.641 in 2017 and 0.000 to 0.662 in 2019 (Table 1). Ratings were stable across the years for 17 of 29 items (Table 1).

The 12 items we chose for further analysis had ICC (1,1) estimates ranging from 0.186 to 0.584 in 2017. Ratings for nine of these 12 items had acceptable ( $> 0.5$ ) or nearing acceptable (0.4 to 0.5) reliability scores (Table 1). Of the items relating to tree climbing and nest building, we retained those with the highest interrater reliability estimates. We retained the item ‘stays on the ground’ as it was reliably rated in 2019 and was stable across the years. We retained ‘approaches and tries to touch staff’ from the items relating to human orientation as it is applicable to all rehabilitation centers (unlike items referring to tourists or volunteers), and due to the risks incurred in approaching humans post-release (Russon et al., 2016). From items relating to exploring the forest vs staying at the center, we retained the item ‘stays in the forest all day’ as it was scored reliably in 2017 and we could test it with our observational data, whereas we could not test staying in the forest a week or longer with observational data. From the personality-based traits we retained the adjective and behavior descriptive item with the highest interrater reliability estimates for exploration and sociability. For boldness, we retained both ‘is bold’ and ‘is bold towards humans’ as we considered the distinction relevant to successful rehabilitation. The behavior descriptor items for boldness, ‘when eating next to stronger orangutans, takes food from nearby them’ and ‘when playing with stronger orangutans, plays rough with them’ were both reliably rated in 2017, but neither were reliably rated in 2019 or stable across the years. We retained the former, considering the potential relevance of food competition to release success when competing for forest resources (Utami et al., 1997).

Ratings for seven of the 12 items chosen for further analysis were stable across the years so we used the mean across both years for the REFA (Table 1). We included data from 2017 in the REFA only for the five items that were not stable across the years: taking food from nearby conspecifics, being curious,

touching novel objects, being friendly, and resting nearby other orangutans (Table 1).

### Determining the behavioral structure based on rating data

The parallel analysis showed we could reliably extract two factors in the REFA. These factors explained 67.5% of the total variance. The Kaiser–Meyer–Olkin (KMO = 0.590) test and Bartlett’s test of sphericity ( $p < 0.01$ ) indicated that the data were suitable for interpretation. We labelled the first factor ‘Forest skills and Human aversion’ as it contained most of the ecologically based items with an expected positive impact on forest adaptation, and a lack of interest in approaching humans. The Forest skills and Human aversion factor explained 40.1% of the overall variance in ratings (Table 3). We labeled the second factor ‘Bold, Social and Explorative’. This factor explained 27.4% of the variance and contained one item reflecting boldness, two items reflecting sociability, and two items reflecting exploration.

Factor scores did not differ significantly between rehabilitation stage (Mann–Whitney test: Forest skills and Human aversion:  $U = 26.0$ ,  $p = 0.721$ , Bold, Social and Explorative:  $U = 23.0$ ,  $p = 0.506$ ), between males and females (Mann–Whitney test: Forest skills and Human aversion:  $U = 35.0$ ,  $p = 0.696$ , Bold, Social and Explorative:  $U = 23.0$ ,  $p = 0.146$ ), or between juveniles,

**Table III** Predicted impact of behavior on rehabilitation and item loadings from regularized exploratory factor analysis on two factors describing the behavior of *Pongo pygmaeus morio* in Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia 2017–2019

Questionnaire item	Predicted impact on rehabilitation	Factor loading	
		1	2
Climbs on trees	positive	0.893	
Stays on the ground	negative	−0.883	
Stays in the forest all day	positive	0.847	
Approaches and tries to touch staff	negative	−0.801	
Builds nests	positive	0.752	
Bold towards humans (bold)	negative	−0.534	0.463
When relaxing, is nearby other orangutans (social)	?		0.891
Is friendly to other orangutans (social)	?		0.777
If there is a new object (for example: new ropes) touches it (explorative)	?		0.704
Is curious (explorative)	?	−0.462	0.645
Is bold (bold)	?		0.542

Words in brackets indicate the general personality trait the item corresponds to. The item ‘When eating next to stronger orangutans, takes food from nearby them (bold)’ did not reach the threshold to load on any factor (factor 1: 0.386; factor 2: 0.243).

adolescents, and adults, (Kruskal–Wallis test: Forest skills and Human aversion:  $X^2(2) = 4.457$ ,  $p = 0.108$ , Bold, Social and Explorative:  $X^2(2) = 0.040$ ,  $p = 0.890$ ).

### Testing the validity of the rating tool

Congruence coefficients calculated following Procrustes rotation on the factor scores from the rating and observational dataset indicated factor-level congruence was highest for the Forest Skills and Human aversion factor (0.93), suggesting it is replicable. Congruence for the Bold, Social, Explorative factor (0.71) did not reach the threshold for fair replicability (0.85) (Lorenzo-Seva & ten Berge, 2006).

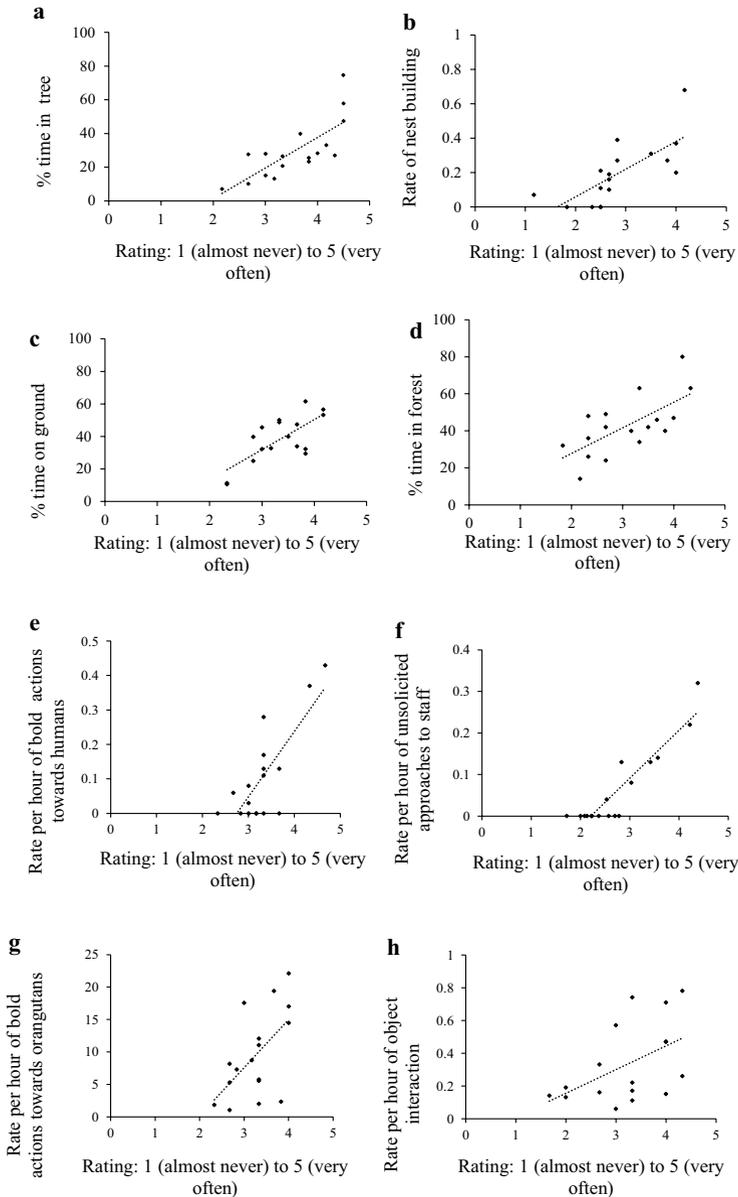
All items predicted to directly impact rehabilitation were significantly and positively correlated with observational data (Fig. 1, Table 4). Four items did not correlate significantly with the observed behaviors, two were more general personality-based items ('is friendly' and 'is curious'), and the other two reflected exploration and boldness where a specific behavior occurring in a particular context was described ('When eating next to stronger orangutans, takes food from nearby them' and 'When relaxing is nearby other orangutans' (Table 4).

### Discussion

These results suggest a rating-based questionnaire can gather valid behavioral data reflecting adaptation to the forest in free-ranging rehabilitant orangutans. Such data can prove challenging to collect using other methods. In general, we found a greater coherence between ratings and observations when the ratings described specific behaviors (for instance 'builds nests') rather than general adjectives (for instance 'is social'). Our research supports previous findings in great apes (Uher & Asendorpf, 2008), where ratings for items including descriptions of the behavior were more consistent with experimentally assessed behaviors than adjective-based items.

A pre-requisite of validity is that the rating instrument generates reliable data (Iwanicki & Lehmann, 2015). When we reduced the questionnaire set to make it easier to administer in future, most items we selected had at least moderate indicators of interrater reliability. We preferentially selected reliably rated items, but retained some with low ICC values where there was no reliable alternative because items with ICC values greater than 0 can still express true variance if they load onto factors following data reduction (Weiss et al., 2017; Wilson et al., 2018). The three items we selected that did not meet the threshold for moderate agreement all loaded onto factors following REFA.

Overall interrater reliability was lower than values reported for some other studies of non-human primates (Capitano, 1999; Freeman & Gosling, 2010). Items with no interrater reliability probably referred to unusual scenarios, for instance, referring to a new food ('like a blue apple') or were unclear for the raters, requiring adjustments to the explanations or revision of the items in future questionnaires. Our interrater reliability values may also be lower due to



**Fig. 1** Correlation between ratings for validated questionnaire items and behavioural observations of *Pongo pygmaeus morio* in Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia 2017-2019. Items: (a) ‘Climbs trees’, (b) ‘Builds nests’, (c) ‘Stays on ground’, (d) ‘Stays in the forest all day’, (e) ‘Bold towards humans’, (f) ‘Approaches staff’, (g) ‘Is bold’, (h) ‘If there is a new object, touches it’

the study design, in which different raters rated each orangutan. We used an ICC (1,1) model to account for this, which consistently results in lower estimates (Trevethan, 2017). This could be rectified in future studies.

**Table IV** Results of Spearman rank correlations ( $n = 17$ ) between ratings for questionnaire items and behavioral observational of *Pongo pygmaeus morio* in Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia 2017–2019

Questionnaire item	$r_s$	$P$
Climbs on trees	<i>0.755</i>	<i>&lt; 0.001</i>
Builds nests	<i>0.836</i>	<i>&lt; 0.001</i>
Stays on the ground	<i>0.630</i>	<i>0.007</i>
Stays in the forest all day	<i>0.626</i>	<i>0.007</i>
Is bold towards humans	<i>0.605</i>	<i>0.010</i>
Approaches humans	<i>0.769</i>	<i>&lt; 0.001</i>
Is bold	<i>0.533</i>	<i>0.027</i>
When eating next to stronger orangutans, takes food from nearby them	-0.198	0.456
Is curious	0.121	0.643
If there is a new object (for example: new ropes) touches it	<i>0.491</i>	<i>0.045</i>
Is friendly	-0.023	0.931
When relaxing is nearby other orangutan	0.261	0.312

Significant correlations are in italics.

We expected ratings for items related to forest adaptation (nest building, tree climbing, etc.) to change over time as rehabilitation progresses, but these changes were not significant between years. Change may occur too slowly to detect across our 2-year study. Overall progress may also appear limited due to sampling bias. The study subjects regularly visit the center by choice, meaning they spend less time in the forest than other individuals in the later stage of rehabilitation who have dispersed further. We did not study these individuals, whose scores may have been higher. We expect personality traits to be more stable than forest skills over time, despite some evidence to suggest that expression of such traits can differ with age (Russon, 2006; Weiss, 2017), where younger individuals appear bolder and more social than adults do (Baker et al., 2015; Carter et al., 2014; Massen et al., 2013). Our sample varied in age but many of the personality-based items were consistent across the years, suggesting developmental effects did not influence our findings. While the lack of significant change over time highlights the potential significance of individuality over age in assessing forest adaptation prior to release, it also illustrates the time-consuming nature of rehabilitation (Basalamah et al., 2018; Riedler et al., 2010; Russon, 2009). In addition, it may highlight methodological issues in research of this nature in apes, where sample sizes are often limited, reducing statistical power (Serdar et al., 2021; Tkaczynski et al., 2019). Increasing the sample size by expanding data collection to other rehabilitation centers would provide greater confidence when interpreting the data.

Successful forest adaptation is based on numerous social and ecological factors (Russon, 2009). The first factor determined by the REFA, 'Forest skills and Human aversion', grouped items most relevant for assessing adaptation to the forest (Riedler et al., 2010; Russon, 2009), whereas the second factor ('Bold, Social and Explorative') grouped personality-based items with no direct relevance to forest adaptation. Individuals with high scores in the Forest skills and Human aversion factor are

likely to be frequent tree climbers who stay in the forest all day, build nests, and are not human orientated. These behaviors have been linked to successful orangutan forest adaptation and survival in the forest (Basalamah et al., 2018; Grundmann, 2006; Riedler et al., 2010). Orangutans with such qualities may more closely mirror the arboreal nature of wild populations (Riedler et al., 2010; Snaith, 1999) and in doing so, learn the skills needed to survive independently in the forest more readily (Russon, 2002). Rehabilitants may enhance ecological skill and resource acquisition by spending time in the forest (Germano et al., 2017; Reader, 2015), which would be beneficial post-release when less support is available.

Orangutans with the likely advantageous qualities in the Forest Skills and Human aversion factor appear to avoid humans. Researchers found a similar negative relationship between human orientation and ecological behaviors reflecting positive forest adaptation in rehabilitant Sumatran orangutans (*Pongo abelii*, Riedler et al., 2010). The orangutans in that study who were less interested in humans resembled wild orangutans more in their behaviors. Specifically, their diets contained a higher proportion of fruits, and they showed more ground avoidance and superior nest-building behaviors than those who were more human orientated. The less human-orientated orangutans also associated more with experienced conspecifics, leading the authors to suggest increased social contact may have facilitated the learning of these behaviors (Riedler et al., 2010). However, the data we collected did not support an increase in forest skills being driven by a general sociability with other orangutans in the transitional phase prior to the final release, as items reflecting 'sociability' did not load positively onto the factor with forest skills. As captive orangutans have demonstrated a preference for learning from a dominant male (Dindo et al., 2011), it is important to consider whether individuals socialize with more experienced conspecifics primarily, and to what extent such socializing benefits information transmission in future studies.

Individuals with high scores on the second factor are confident, friendly, and curious, demonstrating an interest in novel objects. The factor did not include any of the behaviors we expected to directly impact rehabilitation, although we did not test a connection to release outcome. As the item 'is bold towards humans' loaded on the first factor rather than the second, this aspect of boldness may be most related to the forest adaptive behaviors (Basalamah et al., 2018; Grundmann, 2006). Boldness in general may therefore not necessarily be considered a negative quality in orangutan release. Indeed, an increase in the likelihood of food patch discovery under experimental conditions has been identified in bolder chacma baboons (*Papio ursinus*, Carter et al., 2013), suggesting that in this case boldness may in fact be beneficial to survival.

Sociability may enhance rehabilitation if suitable orangutan models are used to assist information transmission (Dindo et al., 2011; Riedler et al., 2010;), but it may also hamper it if maladaptive behaviors are learnt from others. The impact of sociability on orangutan release success may therefore vary depending on the individual and situation. This might be especially true for rehabilitant orangutans, who have more opportunities to be sociable than their wild counterparts due to group rearing and congregation for supplementary feeding (Kuze et al., 2008; Russon, 2009). Orangutans may also be able to access more human artefacts at rehabilitation

centers than in the wild. A willingness to interact with objects is often seen an indicator of exploratory tendency (Forss et al., 2015; Massen et al., 2013) but it may be problematic in this context. Future studies may therefore benefit from separating interaction with novel objects in general into interaction with human-derived or naturally occurring objects.

The Forest skills and Human aversion factor is promising for assessing rehabilitation as it groups all the positive expressions of forest adaptation, tree climbing, nest building etc. On a factor level, ratings clustered similarly to observations. While this replicability is promising for the overall use of factor scores to make valid assessments of rehabilitation progress, factor scores did not differ with age or rehabilitation stage as might be predicted for these behaviors. While the inclusion of more stable personality-based items to the constructs may have diluted the effect of age and experience, even at an item level, the differences between years for most of these behaviors were not significant. This might indicate a relative lack of sensitivity to skills and an over-reliance on age in the assignment of orangutans to rehabilitation stages. Supporting the validation provided by previous behavioral studies in different non-human primate species (Eckardt et al., 2015; Pederson et al., 2005), two-thirds of questionnaire items ratings correlated significantly with the observational data collected in 2017 (eight of 12). This supports our hypothesis that the questionnaire items reflect genuine individual characteristics which may be used to predict real-world outcomes (Iwanicki & Lehmann, 2015), in this case forest adaptation.

Items with more specific descriptions, and items related to the ecological aspects of behavior, appear to often allow quick, valid assessments, requiring little effort from their caregivers. Similarly, some behavioral traits in macaque (*Macaca mulatta*) personality dimensions can be reliably predicted when comparing rating and observational data (Capitanio, 1999) but other results are less consistent (Iwanicki & Lehmann, 2015; Uher & Visalberghi, 2016). Our results support past research showing items describing specific behaviors (e.g., climbs trees), better reflected observed behaviors than more general adjective behavioral terms (e.g., 'is social') (Tkaczynski et al., 2019; Uher & Asendorpf, 2008). However, two items describing a specific behavior occurring in a particular context ('reflecting exploration' and 'boldness') were not validated by observations in this study: 'When eating next to stronger orangutans, takes food from nearby them' and 'When relaxing is nearby other orangutans'. Behaviors such as these may not be conspicuous enough (Freeman & Gosling, 2010) to be noticed readily by the caregivers. It is also possible that validity would increase if the item specified a location for this behavior (e.g., feeding platforms). While video coding analysis remains preferable for identifying more nuanced behavioral actions (Uher & Asendorpf, 2008), by measuring personality in the forest we may gather more valid data with which to assess rehabilitation. This context would be particularly relevant if trait expression changes according to the environment (Tkaczynski et al., 2019).

Considering the size of orangutan populations housed in rehabilitation centers and with little sign of these numbers reducing (Brent, 2007; Russon, 2009), stakeholders should prioritize efforts to improve the success of release attempts.

Our questionnaire to assess rehabilitant orangutan behavior prior to release relies on caregivers' knowledge and is quick to complete, making it an inexpensive, practical, and efficient method to collect standardized data. Further application of this validated method could help identify individuals with skills requiring improvement, assist in assessing skill improvement over time, and help suggest an optimum time to release individuals. It might also be useful to compare pre-and post-release ratings or to examine the impact of skill improvements. This method is suitable for longitudinal data collection, which would enable progress to be tracked throughout rehabilitation, and help determine whether factor scores are related to post-release success. By comparing factor scores with post-release survival, we may be able to develop more accurate predictors of release outcome in the future.

Other rehabilitant populations may differ from ours, for instance in the degree of human habituation. This may have implications for raters' assessments (Tkaczynski et al., 2019), so the context of each population should be considered and the items adjusted if necessary when applying the questionnaire. While we provide some insight into traits that reflect adaptation to the forest, further examination of the impact of personality traits on forest behavior is required to determine whether they are relevant to the rehabilitation process. Our data support those from previous studies which suggest that while age and experience (rehabilitation stage) may influence adaptation to the forest in rehabilitants, individual progress can differ (Riedler et al., 2010; Russon et al., 2016). Studying individual differences in relation to release efforts is a relatively recent development (de Azevedo & Young, 2021; Powell & Gartner, 2011). Species-specific differences are evident (Bremner-Harrison et al., 2004; Sinn et al., 2014), but as an approach that could be administered relatively easily at multiple centers, this method may benefit release efforts for a wide range of primate species in the future.

**Author Contributions:** FR and MDR designed the study. FR, AT, SA, and MDR contributed to the data collection. FR and HC analyzed the data. FR wrote the manuscript, while MDR, KB, JM, and HC provided editorial advice.

**Acknowledgements** We thank Anne Russon for very helpful discussions on this topic, and the Sabah Wildlife Department, Sabah Biodiversity Centre, and Sepilok Orangutan Rehabilitation Centre for permitting this research and for support of the staff there throughout. We also thank Pravidn Segaran, Yuen Zhao Yong, Natalie Adams, Emily Corbi, Aurelie Saulton, Alice Judge, Fabio Crepaldi, Jesse Hay, Carla Osterburg, and Koray Yilmaz for their assistance in data collection, and the Royal Society Research Grant (140282) and Leakey Foundation Research Grant for funding. We also kindly thank the editor and anonymous reviewers who have provided valuable feedback on the manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Allard, S., Fuller, G., Torgerson-White, L., Starking, M., & Yoder-Nowak, T. (2019). Personality in zoo-hatched Blanding's turtles affects behaviour and survival after reintroduction into the wild. *Frontiers in Psychology*, *10*, 2324. <https://doi.org/10.3389/fpsyg.2019.02324>.
- Altmann, J. (1974). Observational study of behavior: sampling Methods. *Behaviour*, *49*(3–4), 227–266. <https://doi.org/10.1163/156853974X000534>.
- Baker, K., Lea, S., & Melfi, V. (2015). Comparative personality assessment of three captive primate species: *Macaca nigra*, *Macaca sylvanus*, and *Saimiri sciureus*. *Int J Primatol*, *36*(3), 625–646. <https://doi.org/10.1007/s10764-015-9843-3>.
- Baker, L., Lawrence, M., Toews, M., Kuling, S., & Fraser, D. (2016). Personality differences in a translocated population of endangered kangaroo rats (*Dipodomys stephensi*) and implications for conservation success. *Behaviour*, *153*(13–14), 1795–1816. <https://doi.org/10.1163/1568539X-00003380>.
- Bard, K., & Hopkins, D. (2018). Early socioemotional intervention mediates long-term effects of atypical rearing on structural covariation in gray matter in adult chimpanzees. *Psychol Sci*, *29*(4), 594–603. <https://doi.org/10.1177/0956797617740685>.
- Basalamah, F., Atmoko, S., Perwitasari-Farajallah, D., Qayim, I., Sihite, J., Noordwijk, M., Willems, E., & Schaik, C. (2018). Monitoring orangutan reintroduction: Results of activity budgets, diets, vertical use and associations during the first year post-release in Kehje Sewen Forest, East Kalimantan. *Indonesia. Biodiversitas*, *19*(2), 639–650. <https://doi.org/10.13057/biodiv/d190242>.
- Beck, B., Walkup, C., Rodrigues, M., Unwin, S., Travis, D., & Stoinski, T. (2007). *Best practice guidelines for the re-introduction of great apes*. SCC Primate Specialist Group of the World Conservation Union.
- Biro, P., & Stamps, J. (2008). Are animal personality traits linked to life-history productivity? *Trends in Ecology & Evolution*, *23*(7), 361–368. <https://doi.org/10.1016/j.tree.2008.04.003>.
- Błaszczuk, M. (2020). Primates got personality, too: Toward an integrative primatology of consistent individual differences in behavior. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(2), 56–67. <https://doi.org/10.1002/EVAN.21808>.
- Bremner-Harrison, S., Prodohl, P., & Elwood, R. (2004). Behavioural trait assessment as a release criterion: boldness predicts early death in a reintroduction programme of captive-bred swift fox (*Vulpes velox*). *Animal Conservation*, *7*(3), 312–320. <https://doi.org/10.1017/S1367943004001490>.
- Brent, L. (2007). Life-long well being: applying animal welfare science to nonhuman primates in sanctuaries. *Journal of Applied Animal Welfare Science*, *10*(1), 55–61. <https://doi.org/10.1080/10888700701277626>.
- Budaev, S. (2010). Using principal components and factor analysis in animal behaviour research: caveats and guidelines. *Ethology*, *116*(5), 472–480. <https://doi.org/10.1111/j.1439-0310.2010.01758.x>.
- Capitanio, J. (1999). Personality dimensions in adult male rhesus macaques: prediction of behaviors across time and situation. *American Journal of Primatology*, *47*(4), 299–320. [https://doi.org/10.1002/\(SICI\)1098-2345\(1999\)47:4<299::AID-AJP3>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2345(1999)47:4<299::AID-AJP3>3.0.CO;2-P).
- Carter, A., Marshall, H., Heinsohn, R., & Cowlshaw, G. (2013). Personality predicts decision making only when information is unreliable. *Animal Behaviour*, *86*(3), 633–639. <https://doi.org/10.1016/j.anbehav.2013.07.009>.
- Carter, A., Marshall, H., Heinsohn, R., & Cowlshaw, G. (2014). Personality predicts the propensity for social learning in a wild primate. *PeerJ*, *2*, e283. <https://doi.org/10.7717/peerj.283>.
- Chotard, H. (2020). *Measuring Personality in Great Apes: Special Focus on Naturally Occurring Contexts*. Unpublished PhD thesis. University of Portsmouth.
- Clay, A., Bloomsmith, M., Bard, K., Maple, T., & Marr, M. (2015). Long-term effects of infant attachment organization on adult behavior and health in nursery-reared, captive chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *129*(2), 145–159. <https://doi.org/10.1037/a0038901>.
- Damerius, L., Graber, S., Willems, E., & van Schaik, C. (2017). Curiosity boosts orang-utan problem-solving ability. *Animal Behaviour*, *134*, 57–70. <https://doi.org/10.1016/j.anbehav.2017.10.005>.
- de Azevedo, C., & Young, R. (2021). Animal personality and conservation: basics for inspiring new research. *Animals*, *11*(4), 1019. <https://doi.org/10.3390/ANI11041019>.
- Dindo, M., Stoinski, T., & Whiten, A. (2011). Observational learning in orangutan cultural transmission chains. *Biology Letters*, *7*(2), 181–183. <https://doi.org/10.1098/rsbl.2010.0637>.

- Dingemans, N., Both, C., Drent, P., & Tinbergen, J. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1541), 847–852. <https://doi.org/10.1098/rspb.2004.2680>.
- Eckardt, W., Steklis, H., Steklis, N., Fletcher, A., Stoinski, T., & Weiss, A. (2015). Personality dimensions and their behavioral correlates in wild Virunga mountain gorillas (*Gorilla beringei beringei*). *Journal of Comparative Psychology*, 129(1), 26–41. <https://doi.org/10.1037/a0038370>.
- Fischer, R., & Fontaine, J. (2010). Methods for investigating structural equivalence. In D. Matsumoto & F. Van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 179–215). Cambridge University Press. <https://doi.org/10.1017/CBO9780511779381.010>.
- Forss, S., Schuppli, C., Haiden, D., Zweifel, N., & van Schaik, C. (2015). Contrasting responses to novelty by wild and captive orangutans: novelty response in orangutans. *American Journal of Primatology*, 77(10), 1109–1121. <https://doi.org/10.1002/ajp.22445>.
- Freeman, H., & 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/10.1002/ajp.20833>.
- Freeman, H., Brosnan, S., Hopper, L., Lambeth, S., Schapiro, S., & Gosling, S. (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/10.1002/ajp.22168>.
- Germano, J., Nafus, M., Perry, J., Hall, D., & Swaisgood, R. (2017). Predicting translocation outcomes with personality for desert tortoises. *Behavioral Ecology*, 28(4), 1075–1084. <https://doi.org/10.1093/beheco/axx064>.
- Grundmann, E. (2006). Back to the wild: will reintroduction and rehabilitation help the long-term conservation of orangutans in Indonesia? *Social Science Information*, 45(2), 265–284. <https://doi.org/10.1177/0539018406063643>.
- Iwanicki, S., & Lehmann, J. (2015). Behavioral and trait rating assessments of personality in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, 129(3), 205–217. <https://doi.org/10.1037/a003931>.
- Jakuba, T., Polcová, Z., Fedáková, D., Kottferová, J., Mareková, J., Fejsáková, M., Ondrašovičová, O., & Ondrašovič, M. (2013). Differences in evaluation of a dog's temperament by individual members of the same household. *Society & Animals*, 21(6), 582–589. <https://doi.org/10.1163/15685306-12341314>.
- Jung, S., & Lee, S. (2011). Exploratory factor analysis for small samples. *Behavior Research Methods*, 43, 701–709. <https://doi.org/10.3758/s13428-011-0077-9>.
- King, J., Weiss, A., & Farmer, K. (2005). A chimpanzee (*Pan troglodytes*) analogue of cross-national generalization of personality structure: zoological parks and an African sanctuary. *Journal of Personality*, 73(2), 389–410. <https://doi.org/10.1111/j.1467-6494.2005.00313.x>.
- King, T., Chamberlan, C., & Courage, A. (2012). Assessing initial reintroduction success in long-lived primates by quantifying survival, reproduction, and dispersal parameters: western lowland gorillas (*Gorilla gorilla gorilla*) in Congo and Gabon. *Int J Primatol*, 33, 134–149. <https://doi.org/10.1007/s10764-011-9563-2>.
- Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropract Med*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Koski, S. (2011). Social personality traits in chimpanzees: temporal stability and structure of behaviourally assessed personality traits in three captive populations. *Behavioral Ecology and Sociobiology*, 65(11), 2161–2174. <https://doi.org/10.1007/s00265-011-1224-0>.
- Kuze, N., Sipangkui, S., & Malim, T. (2008). Reproductive parameters over a 37-year period of free-ranging female Borneo orangutans at Sepilok Orangutan Rehabilitation Centre. *Primates*, 49, 126–134. <https://doi.org/10.1007/s10329-008-0080-7>.
- Lorenzo-Seva, U., & ten Berge, J. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2(2), 57–64. <https://doi.org/10.1027/1614-2241.2.2.57>.
- Masilkova, M., Weiss, A., Slipogor, V., & Konečná, M. (2020). Comparative assessment of behaviorally derived personality structures in golden-handed tamarins (*Saguinus midas*), cotton-top tamarins (*Saguinus oedipus*), and common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, 134(4), 453. <https://doi.org/10.1037/com0000226>.
- Massen, J., Antonides, A., Arnold, A.-M., Bionda, T., & Koski, S. (2013). A behavioral view on chimpanzee personality: exploration tendency, persistence, boldness, and tool-orientation measured with

- group experiments: behavioral measures of chimp personality. *American Journal of Primatology*, 75(9), 947–958. <https://doi.org/10.1002/ajp.22159>.
- McCrae, R., Zonderman, A., Costa Jr., P., Bond, M., & Paunonen, S. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552–566. <https://doi.org/10.1037/0022-3514.70.3.552>.
- Murray, L. (2011). Predicting primate behavior from personality ratings. In A. Weiss, J. King, & L. Murray (Eds.), *Personality and temperament in nonhuman primates. developments in primatology: progress and prospects* (pp. 129–169). Springer. [https://doi.org/10.1007/978-1-4614-0176-6\\_6](https://doi.org/10.1007/978-1-4614-0176-6_6).
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32, 396–402. <https://doi.org/10.3758/BF03200807>.
- Pederson, A., King, J., & Landau, V. (2005). Chimpanzee (*Pan troglodytes*) personality predicts behavior. *Journal of Research in Personality*, 39(5), 534–549. <https://doi.org/10.1016/j.jrp.2004.07.002>.
- Powell, D., & Gartner, M. (2011). Applications of personality to the management and conservation of nonhuman animals. In M. Inoue-Murayama, S. Kawamura, & A. Weiss (Eds.), *From genes to animal behavior* (pp. 185–199). Springer. [https://doi.org/10.1007/978-4-431-53892-9\\_8](https://doi.org/10.1007/978-4-431-53892-9_8).
- Reader, S. (2015). Causes of individual differences in animal exploration and search. *Topics in Cognitive Science*, 7(3), 451–468. <https://doi.org/10.1111/tops.12148>.
- Réale, D., Gallant, B., Leblanc, M., & Festa-Bianchet, M. (2000). Consistency of temperament in big-horn ewes and correlates with behaviour and life history. *Animal Behaviour*, 60(5), 589–597. <https://doi.org/10.1006/anbe.2000.1530>.
- Réale, D., Reader, S., Sol, D., McDougall, P., & Dingemans, N. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews*, 82(2), 291–318. <https://doi.org/10.1111/j.1469-185X.2007.00010.x>.
- Riedler, B., Millesi, E., & Pratte, P. (2010). Adaptation to forest life during the reintroduction process of immature *Pongo abelii*. *International Journal of Primatology*, 31(4), 647–663. <https://doi.org/10.1007/s10764-010-9418-2>.
- Rijksen, H. D. (1978). A field study on Sumatran orangutans (*Pongo pygmaeus abelii*): ecology, behavior, and conservation (Doctoral dissertation)
- Russon, A. (2006). Acquisition of complex foraging skills in juvenile and adolescent orangutans (*Pongo pygmaeus*): developmental influences. *Aquatic Mammals*, 32, 500–510. <https://doi.org/10.1578/AM.32.4.2006.500>.
- Russon, A. (2009). Orangutan rehabilitation and reintroduction. In S. Wich, S. U. Atmoko, M. Setia, & C. P. van Schaik (Eds.), *Orangutans: geographic variation in behavioral ecology and conservation* (pp. 327–350). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199213276.003.0023>.
- Russon, A., Smith, J., & Adams, L. (2016). Managing human–orangutan relationships in rehabilitation. In M. Waller (Ed.), *Ethnoprimatology. Developments in primatology: progress and prospects* (pp. 233–258). Springer. [https://doi.org/10.1007/978-3-319-30469-4\\_13](https://doi.org/10.1007/978-3-319-30469-4_13).
- Russon, A. E. (2002). Return of the native: cognition and site-specific expertise in orangutan rehabilitation. *International Journal of Primatology*, 23(3), 461–478. <https://doi.org/10.1023/A:1014909431148>.
- Serdar, C., Cihan, M., Yücel, D., & Serdar, M. (2021). Sample size, power, and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochimica medica*, 31(1), 010502. <https://doi.org/10.11613/BM.2021.010502>.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>.
- Sinn, D., Cawthen, L., Jones, S., Pukk, C., & Jones, M. (2014). Boldness towards novelty and translocation success in captive-raised, orphaned Tasmanian devils. *Zoo Biology*, 33(1), 36–48. <https://doi.org/10.1002/zoo.21108>.
- Snaith, T. (1999). *The behaviour of free-ranging ex-captive orangutans in Tanjung Puting National Park*. Thesis, University of Calgary, Calgary, Canada.
- Soorae, P. S. (Ed.) (2010). *Global re-introduction perspectives: additional case-studies from around the globe*. IUCN/ SSC Re-introduction Specialist Group.

- Stratton, T. (2015). Use of personality to improve reintroduction success: the effects of behavioural variation within release groups. Doctoral dissertation, Nottingham Trent University. <https://doi.org/10.13140/RG.2.2.20727.98722>
- Tkaczynski, P., Ross, C., MacLarnon, A., Mouna, M., Majolo, B., & Lehmann, J. (2019). Measuring personality in the field: an in-situ comparison of personality quantification methods in wild barbary macaques (*Macaca sylvanus*). *Journal of Comparative Psychology*, *133*(3), 313–325. <https://doi.org/10.1037/COM0000163>.
- Trevethan, R. (2017). Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. *Health Serv Outcomes Res Method*, *17*, 127–143. <https://doi.org/10.1007/s10742-016-0156-6>.
- Uher, J. (2008). Comparative personality research: Methodological approaches. *European Journal of Personality*, *22*(5), 427–455. <https://doi.org/10.1002/per.680>.
- Uher, J., & Asendorpf, J. (2008). Personality assessment in the great apes: comparing ecologically valid behavior measures, behavior ratings, and adjective ratings. *Journal of Research in Personality*, *42*(4), 821–838. <https://doi.org/10.1016/j.jrp.2007.10.004>.
- Uher, J., & Visalberghi, E. (2016). Observations versus assessments of personality: a five-method multispecies study reveals numerous biases in ratings and methodological limitations of standardised assessments. *Journal of Research in Personality*, *61*, 61–79. <https://doi.org/10.1016/j.jrp.2016.02.003>.
- Utami, S., Wich, S., & Sterck, E. (1997). Food competition between wild orangutans in large fig trees. *International Journal of Primatology*, *18*, 909–927. <https://doi.org/10.1023/A:1026392012922>.
- Úbeda, Y., Ortín, S., & St. Leger, J., Llorente, M., & Almunia, J. (2019). Personality in captive killer whales (*Orcinus orca*): A rating approach based on the five-factor model. *Journal of Comparative Psychology*, *133*(2), 252–261. <https://doi.org/10.1037/com0000146>.
- Weiss, A. (2017). A human model for primate personality. *Proceedings of the Royal Society Biological Sciences*, *284*(1864), 20171129. <https://doi.org/10.1098/rspb.2017.1129>.
- Weiss, A., & King, J. (2015). Great ape origins of personality maturation and sex differences: A study of orangutans and chimpanzees. *Journal of Personality and Social Psychology*, *108*(4), 648–664. <https://doi.org/10.1037/pspp0000022>.
- Weiss, A., Wilson, M., Collins, D., Mjungu, D., Kamenya, S., Foerster, S., & Pusey, A. (2017). Personality in the chimpanzees of Gombe National Park. *Scientific Data*, *4*, 170146. <https://doi.org/10.1038/sdata.2017.146>.
- Wilson, V., Inoue-Murayama, M., & Weiss, A. (2018). A comparison of personality in the common and Bolivian squirrel monkey (*Saimiri sciureus* and *Saimiri boliviensis*). *Journal of Comparative Psychology*, *132*(1), 24–39. <https://doi.org/10.1037/com0000093>.
- Wolf, M., & Weissing, F. J. (2012). Animal personalities: consequences for ecology and evolution. *Trends in Ecology & Evolution*, *27*(8), 452–461. <https://doi.org/10.1016/j.tree.2012.05.001>.

## Authors and Affiliations

Florence Rocque<sup>1</sup>  · Helene Chotard<sup>1</sup>  · Kim Bard<sup>1</sup>  · Jerome Micheletta<sup>1</sup>  · Augustine Tuuga<sup>2</sup> · Sylvia Alsisto<sup>2</sup> · Marina Davila-Ross<sup>1</sup> 

✉ Florence Rocque  
Florence.Rocque@port.ac.uk

<sup>1</sup> Department of Psychology, University of Portsmouth, King Henry Building, King Henry I Street, Portsmouth PO1 2DY, UK

<sup>2</sup> Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia