Lynx, n. s. (Praha), **48**: 183–198 (2017). ISSN 0024-7774 (print), 1804-6460 (online)

Methods for measuring mammalian personalities: In which animals and how accurately can we quantify it?

Metody měření personality savců: U kterých zvířat a jak přesně ji dokážeme měřit?

Barbora ŽAMPACHOVÁ[1,2], Eva LANDOVÁ[1,2] & Daniel FRYNTA[1,2]

[1] National Institute of Mental Health, Topolová 748, 250 67 Klecany, Czech Republic
[2] Department of Zoology, Faculty of Science, Charles University, Viničná 7, 128 44 Praha, Czech Republic

received on 21 November 2017

Abstract. The study of personality, and individual differences in behaviour has experienced a steady rise in popularity in the past years. In this review and meta-analysis, we aim to introduce the concept of personality and related phenomena. A behavioural trait should meet two basic conditions to be considered a personality trait – it should be consistent (1) in time and (2) across contexts. In mammals, the two most common orders in personality studies are primates and rodents. We therefore introduce different approaches to personality testing in these two orders. Primate personality studies are based on psychology studies and often rely on the observer's ratings. Rodent personality studies originate in the studies of physiology and use an experimental approach. We present a more detailed overview of methodological issues of repeatability as a statistical tool for measuring consistency across time. The classic methods of computing repeatability do not consider habituation and other trends which may become confounding factors and lead to underestimation of repeatability. We also discuss consistency across contexts and different understandings of the context definition. We illustrate the variability of personality studies in mammals with a meta-analysis of repeatability estimates. We found that repeatability of behaviour depends on the methodology of behavioural testing and statistical analyses used, but also the number of test repetitions and differences between the focal behaviours. Repeatability decreased with more repetitions and the tests of aggressiveness and exploratory behaviour yielded lower repeatability estimates than the tests of activity.

Key words. Personality, repeatability, behavioural syndrome, aggressiveness, exploratory behaviour, Rodentia, Primates.

Definition of personality

The animal personality topic has experienced a boom in popularity over the last two decades, with 520 articles in the last four years, 135 of these studying mammals. Out of these, only four orders were featured in more than five studies: 48 focused on rodents, 37 on primates, 26 on carnivores, and 15 on ungulates. The studies do not just state the existence of consistent inter-individual differences in the focal species, but focus on the ecological and evolutionary importance of personality (Dall et al. 2004, Sih et al. 2004a, b, Réale et al. 2010), its ontogeny (Stamps & Groothuis 2010a, b, Šimková et al. 2017), and correlations of personality traits with physiological parameters (Koolhaas et al. 2010, Careau & Garland 2012). There is also some discussion about how the personality is maintained. The prevailing opinion is that individual

differences are created and maintained by some sort of trade-off, either a life-history trade-off (different life-history strategies favour different personalities, Wolf et al. 2007) or a fluctuating selection pressure, created by rapidly changing environmental conditions that favour different behavioural responses (Dingemanse & Réale 2013).

If we consider the number of published papers on personality from different perspectives, it might be surprising to see how little agreement there is in the methodology of measuring individual differences in animal behaviour. Many different fields (e.g., ethology, comparative psychology, neurobiology, behavioural ecology) approach the inter-individual differences between animals from different angles. This plethora of different terms, methods of quantifying behaviour, and interpretations of behavioural traits may seem rather unorganised and confusing to taxonomists or zoologists who are interested in what the animal personality really is. Therefore, the aim of this article is to guide the reader through the most common methods and terms used in the studies of mammalian personality, as well as their respective issues.

More specifically, the goals were to (1) introduce the concept of personality and related terms, (2) provide a basic overview of methodology of measuring personality in the two most studied mammalian orders (primates and rodents), (3) point out some methodological issues of studying personality, (4) analyse repeatability estimates reported in published articles to provide an indicative insight into factors influencing the temporal stability of behaviour, and finally (5) provide some methodological guidelines for future studies of personality.

Psychologists usually use questionnaires to assess personality, however this method has only a limited applicability in animal behaviour research (see below). Ethologists therefore needed to devise new methods allowing them to search for behavioural traits that could be used to characterize personality of the focal individual. These traits should be as objectively observable as possible and easily quantifiable (e.g., the latency to approach novel object or attack a conspecific in a neutral environment). They are referred to as personality traits and their values as personality scores. Based on the definition of personality (Carere & Eens 2005, Réale et al. 2007, Careau & Garland 2012, Réale & Dingemanse 2012), the behaviour should meet three conditions to be considered a personality trait: (1) inter-individual variability that is (2) consistent across time and (3) contexts (Réale et al. 2007). Here we illustrate the methodology of personality research on two major mammalian orders: primates (Primates) and rodents (Rodentia).

The research of individual variability usually distinguishes between the personality (or temperament as its synonym, Carere & Eens 2005, Réale et al. 2007) and behavioural syndrome (Sih et al. 2004a, b). While the concept of personality or temperament usually focuses on only one aspect of behaviour (e.g., exploration) and emphasizes a temporal stability of individual differences (Réale et al. 2007, Careau et al. 2015), the behavioural syndrome describes correlations between various behavioural traits (e.g., more aggressive individuals being also more active) and does not stress the demand for temporal consistency (Sih et al. 2004a, b, Sih & Bell 2008).

Approaches to personality research in primates

Primate personality research is heavily influenced by the results of human personality research, which has a long history, beginning with Hippocrates and his four temperaments and continuing throughout the 19th and 20th century (Hjelle & Ziegler 1992). Psychologists have successfully used the concept of personality axes to analyse human personalities. An axis of personality is usually a set of functionally similar behaviours in which we can find a gradient of individual

variability. This axis is usually derived from a multivariate analysis of different intercorrelated behavioural variables. In human psychology, one of the most popular paradigms is the "Big Five" model – five major axes of variation in personality. They are usually labelled as conscientiousness (which is expressed in deliberation and self-discipline), openness (expressed in openness to experience, imagination, creativity, curiosity), extraversion (sociability, assertiveness, activity), agreeableness (trust, cooperation, lack of aggression), and neuroticism (anxiety, depression, vulnerability to stress) (Digman 1990, McCrae & John 1992). In 1999, Gosling and John speculated that the "Big Five" might be also found in animals (Gosling & John 1999) which started the search for five personality axes in subhuman species. Today, the five personality axes in animals are considered to be aggressiveness (agonistic reactions to conspecifics), sociability (reactions to absence/presence of conspecifics, excluding agonistic behaviour), exploration (reactions to new stimuli), boldness (reactions to a risky but not novel situation), and activity (a general level of the individual's activity), which can however slightly influence each other, e.g., a general activity influencing exploratory behaviour (Réale et al. 2007).

There are three basic approaches to the quantification of primate personality. The first one is a very common approach relying on various questionnaires derived from human personality questionnaires. Each primate is usually rated on a scale of 1–7 (alternatively 1–5) in a number of aspects (e.g., fearfulness, sociability, playfulness) by people who are familiar with the focal individual – typically caregivers or researchers working with the individual for a long time. This method is most common in apes (Weiss et al. 2007, 2011, 2012, Morton et al. 2013, Robinson et al. 2017), hence the name of the most popular questionnaire – Hominiod Personality Questionnaire. In this measure, caretakers are asked to score whether the animal is above, below, or average (on a 1–7 scale) for 51 selected traits, e.g., dominant or curious (Weiss et al. 2009). It is, however, also used for a variety of primate species (Konečná et al. 2012, Morton et al. 2013, Adams et al. 2015, Robinson et al. 2016).

The second possibility, although not very common, is to rate the personality based on the observation of animals in normal conditions without any experimental manipulation (Koski 2011, Massen & Koski 2014, Staes et al. 2015).

The third approach is experimental and mirrors personality research in other mammals. The most common tests focus on boldness and exploratory behaviour, usually in the form of a novel object or novel food test (Dammhahn & Almeling 2012, Koski & Burkart 2015).

Even though the results of these methods (questionnaires, observation, experiments) can be interpreted as yielding the same personality axis, their comparability is somehow questionable, as they do not always correlate across the methods (Staes et al. 2016). One of the biggest critical arguments against the questionnaire method is its subjectivity. Another problem is that this method cannot accurately measure the temporal stability of behaviour. Even if the questionnaires were administered repeatedly to the respondents, the results would reflect the stability of their subjective view on behaviour rather than the stability of the behaviour itself.

Personality in rodents

Even though the personality issue is considered a new study area and the most cited reviews and concepts are less than twenty years old, individual differences were discussed in the studies of rodent behaviour much earlier (Archer 1973). Researchers recognized the inter-individual variability and used it to select lines with different aggressiveness (Benus et al. 1987) or different exploratory tendencies (DeFries et al. 1978).

Rodent personality research started with the term "coping styles" (Benus et al. 1989, 1991, Koolhaas et al. 1999). Coping styles divide the animals according to their reaction to a highly stressful stimulus. The individuals labelled as proactive tried to actively escape the stress, while the reactive individuals responded with immobility.

The questionnaire approach that we can see in primates (see above) and some other larger mammals (e.g., companion dogs (Svartberg 2005), cats (Gartner & Weiss 2013), or even spotted hyenas (*Crocuta crocuta*) kept in zoos (Gosling 1998)) is almost non-existent in rodents, with one exception concerning pet rats (Korpela 2011). The most common personality aspects studied in rodents are exploratory behaviour and boldness, quantified using a standard testing procedure.

In the studies of explorative behaviour, the most common test is the open field test (Hall 1934, Archer 1973, Walsh & Cummins 1976), and its variation, the hole board test. In the open field, the animal is introduced to an unfamiliar environment (usually some sort of arena to prevent the animal from escaping) and an observer marks the time spent active/inactive and some characteristic behavioural traits (e.g., self-grooming, rearing on hind legs, jumping, see Lantová et al. 2011). In the hole board test (Boissier & Simon 1962, File 2001), there is a number of holes in the arena floor to provide additional stimuli for exploratory behaviour (see Fig. 1).

Another way to test exploratory behaviour is the novel object test, in which an unfamiliar object is introduced into the animal's home cage or another familiar environment (Barnett 1958, Cowan 1976, 1977). However, the tests of exploratory behaviour always face the problem of what can be considered as novel for the animal. If the animal is presented with the same object or introduced to the same environment twice, for the second time the stimulus may no longer be considered as novel by the animal, therefore we are not measuring exploratory behaviour. If we present a different stimulus each time, we do not know whether the results are influenced by the type of stimulus (e.g., one object is of a preferred colour, some shapes of the arena are more stressful than others) and any comparison might be problematic (Heyser & Chemero 2012).

Other tests usually focus on measuring boldness or anxiety (the terminology is not quite unified). The open field, hole board, and the novel object tests can yield some behavioural traits we can interpret as boldness, e.g., self-grooming as a marker of low boldness (Žampachová et al. 2017). However, some of the tests specifically focus on boldness.

One variation of the open field is the so-called light-dark test, when the animal is allowed to choose between a well-lit part of the arena (stressful environment) and a dark part (Aulich 1976). The bolder animals are expected to spend more time in the light part of the arena. Another boldness test is the startle test, which focuses on the animal's reaction to sudden noise or movement (Glowa & Hansen 1994), or the handling test, when the reaction to being handled by humans is observed (Martin & Réale 2008a, b). A very specific rodent test is the elevated plus maze and its derivatives (Pellow et al. 1985, Deacon 2013). The animal is placed on an elevated platform in the shape of a cross or a plus (hence the name). Two arms of this platforms have non-transparent walls (so-called "closed" arms), the other two have no walls ("open" arms) (see Fig. 2). The more time the animal spends on the open arms, the bolder it is considered.

The other axes of personality are studied less often. Aggressiveness and sociability are sometimes studied with the resident-intruder test (Van Oortmerssen & Bakker 1981, Koolhaas et al. 2013), when one individual is introduced to a home cage of the other individual and their interaction is observed. A similar method is the neutral cage test, where both individuals interact in a neutral environment (Čiháková & Frynta 1996, Suchomelová et al. 1998, Munclinger & Frynta 2000, Suchomelova & Frynta 2000). This test can be also used for inter-species

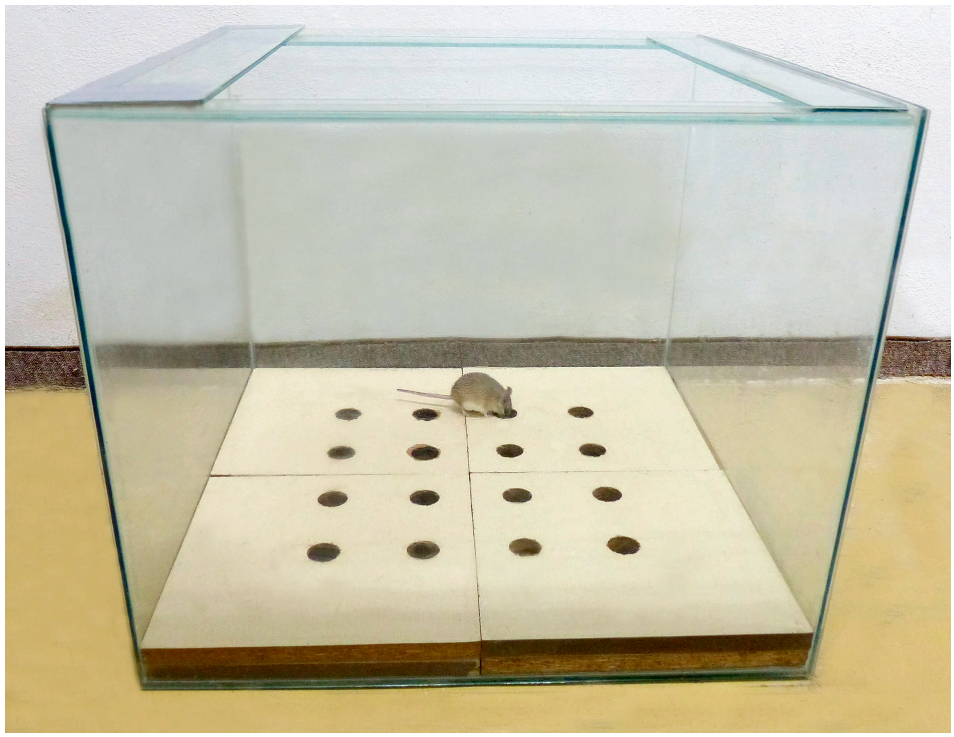

Fig. 1. Example of the hole board test using *Acomys cilicicus* as a subject. The animal is exhibiting head--dipping (looking inside the hole on the floor), a common measure of exploratory behaviour in this test. Photo by Barbora Kaftanová.
Obr. 1. Příklad využití hole board testu u *Acomys cilicicus*. Zvíře provádí head-dipping (nahlížení do děr v podlaze), což je považováno za proměnnou odrážející míru exploračního chování v tomto testu. Foto Barbora Kaftanová.

comparisons (Frynta et al. 2005). Activity can be measured either by ethogram of each individual or by a voluntary wheel running test (Dewsbury 1980, Sherwin 1998, Knab et al. 2009, Meek et al. 2009).

Temporal stability versus habituation

The temporal stability of behavioural traits associated with personality has been considered in studies more often since the meta-analysis by Bell et al. (2009). This study established the generally accepted average repeatability of behaviour r=0.37. Repeatability is usually a number between 0 and 1, reflecting the proportion of variability that can be explained by differences between individuals (Lessells & Boag 1987). In an unadjusted form it is called the agreement repeatability, because it tests the reproducibility of scores of different individuals (McGraw & Wong 1996, Hayes & Jenkins 1997, Nakagawa & Schielzeth 2010, Biro & Stamps 2015),

but it does not account for habituation to the testing procedure. The problem of calculating the repeatability of traits subject to a temporal change has been addressed in detail by Biro & Stamps (2015). When the temporal change is the same in all the animals (e.g., they all habituate at the same rate), the effect of habituation can be corrected for by statistical tools. The result is called the consistency repeatability. This method is useful especially for the tests of exploratory behaviour and boldness. These behaviours are subject to habituation, but consistency repeatability

Fig. 2. Example of the elevated plus maze (EPM) with *Acomys cilicicus* as a subject. The animal is sitting on an "open arm" of the maze, therefore exhibiting a "bold" behaviour. Photo by Barbora Kaftanová.
Obr. 2. Příklad vyvýšeného křížového bludiště (EPM) u *Acomys cilicicus*. Zvíře sedí na "otevřené" části bludiště, tedy projevuje se jako "odvážné". Foto Barbora Kaftanová.

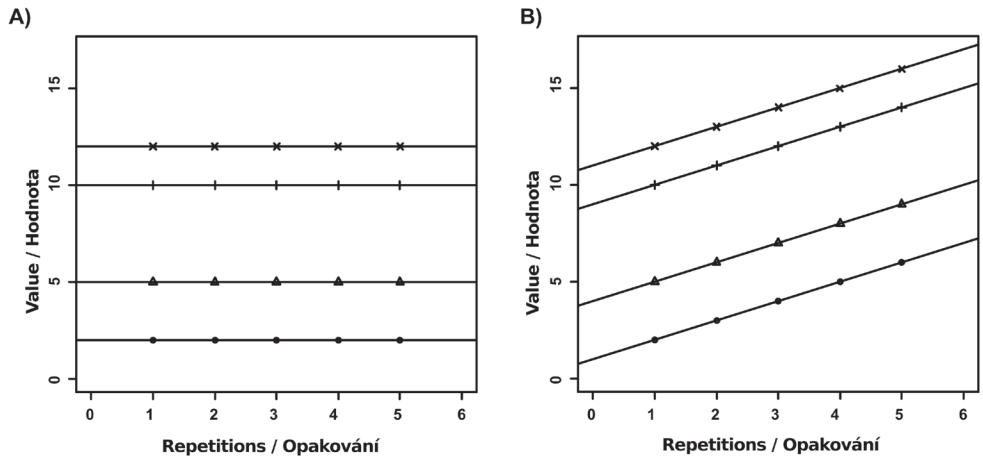

Fig. 3. Hypothetical examples of two personality traits, in which (A) describes a situation when the trait is stable and is not subjected to a temporal trend, therefore both agreement repeatability and consistency repeatability would be high. (B) shows a trait which is consistent, but with a distinct trend, therefore we would find low agreement repeatability. This trait, even though it reflects personality, would be repeatable only after correcting for the temporal trend (consistency repeatability). The different symbols represent values measured in different individuals.
Obr. 3. Hypotetický příklad dvou personalitních rysů, kde část (A) popisuje situaci, kdy je rys stabilní a nepodléhá žádnému trendu, tedy jak opakovatelnost bez korekce (agreement repeatability), tak opakovatelnost, beroucí v úvahu čas (consistency repeatability), by byly vysoké. Část (B) ukazuje personalitní rys, který je konzistentní, ale s výrazným časovým trendem, tedy opakovatelnost bez korekce (agreement repeatability) by byla nízká. Toto chování, přestože se v něm projevuje personalita, by bylo dobře opakovatelné pouze po statistické korekci, zohledňující časový trend (consistency repeatability). Různé symboly označují hodnoty chování, naměřené u různých jedinců.

allows us to establish whether the individual maintains its relative scores to other individuals (e.g., if the boldest individual stays the boldest throughout the habituation process). The consistency repeatability can be high and provide an evidence of personality, even when agreement repeatability is low (see Fig. 3).

It has however been shown that individuals may also differ in their reaction to an environmental variation (Koolhaas et al. 1999, David et al. 2004, Martin & Réale 2008a). This concept originates in behavioural ecology and is referred to as phenotypic plasticity (for a review, see West-Eberhard 2003) and the set of phenotypes produced by a genotype over a range of environments is called the reaction norm (Via et al. 1995). While the ideal personality trait is the one that is the most stable across contexts and situations, the plasticity looks for traits that are flexibly adapting to the current environment. However, different individuals may show different levels of plasticity in their behaviour, therefore we can view the plasticity and behavioural flexibility as another aspect of personality (Koolhaas et al. 1999, Coppens et al. 2010). The behavioural reaction norms therefore describe not only the average level of animal behaviour, but also the differences in individual responses to changes in external conditions (Dingemanse et al. 2010).

Table 1. Coefficients of the gls marginal model, analysing the effect of method, number of repetitions and axis of personality on repeatability estimates
Tab. 1. Koeficienty gls marginálního modelu, analyzujícího efekt metody, počtu opakování a osy personality na hodnotu opakovatelnosti

	value hodnota	standard error směrodatná chyba	t-value hodnota t	p-value pravděpodobnost
(Intercept) / (Intercept)	0.408	0.053	7.630	<0.001
method: correlation / metoda: korelační koeficient	0.086	0.053	1.637	0.103
method: LMM repeatability / metoda: LMM opakovatelnost	–0.181	0.607	–2.982	0.003
numer of repetitions / počet opakování	–0.004	0.003	–1.631	0.104
axis: activity / osa: aktivita	0.175	0.051	3.397	0.001
axis: boldness / osa: odvážnost	0.128	0.048	2.665	0.008
axis: exploration / osa: exploračn chování	0.018	0.054	0.333	0.739
axis: sociability / osa: sociabilita	0.087	0.044	2.007	0.046

Behavioural syndrome and correlations between contexts

Another requirement of personality definition is the consistency across contexts. However, the literature is somewhat ambiguous as to what is considered as different contexts. Stamps and Groothuis defined a context as "all of external stimuli surrounding an individual when it expresses a given behaviour" (Stamps & Groothuis 2010b). This can be interpreted in two different ways. For example, some studies consider two different tests of exploratory behaviour as two different contexts and accept the behavioural trait as a personality trait if it is individually consistent across different tests (Dammhahn & Almeling 2012, Žampachová et al. 2017). Some researchers though understand the correlation between contexts as a correlation between different personality axes, e.g., bold individuals are also more aggressive (Edenbrow & Croft 2012, Rudin & Briffa 2012). These correlations are unfortunately usually studied separately from temporal consistency and very few studies focus on both the temporal and contextual stability of candidate personality traits (Guenther & Trillmich 2012, Petelle et al. 2013, Guenther et al. 2014, Hudson et al. 2015). Some theories state that some behavioural traits, while consistent when studied alone, might not stay correlated during the ontogeny because of different selection pressures on adults and juveniles (Groothuis & Trillmich 2011).

Meta-analysis of personality trait repeatability

Given the wide variety of methods used for personality tests demonstrated above, it is not sufficient to merely describe the variability of studies, but one should also present a simple

indicative meta-analysis to provide some insight as to what differences are truly relevant. We restricted our search to the period of classical studies of personality, which focused more on the description of personality rather than ontogenetical changes or an evolutionary context, therefore we used studies from 2000 to 2013. The source studies were searched for in the Web of Science database with the key words "boldness", "neophobia", "risk taking", "activity", "explor*", "aggressi*" and "repeat*". Out of 104 studies containing estimates of repeatability, we chose only 30 articles studying mammals, which yielded 327 estimates of repeatability. To each repeatability estimate we also noted the id of the study (variable "study"), the method used to calculate the repeatability estimate ("method"), the number of times the test was repeated ("number of replications"), the number of subjects ("number of individuals"), the mammalian clade the subjects belonged to ("clade"), and the axis of personality the behaviour belongs to ("axis of personality"). We categorized three methods of the repeatability estimation (methods reviewed in Nakagawa & Schielzeth 2010): a correlation coefficient (further referred to as correlation), an intraclass correlation coefficient using variance components extracted from ANOVA (further referred to as ANOVA repeatability; Lessells & Boag 1987), and an intraclass correlation coefficient using variance components extracted from linear mixed-effects models or generalized linear mixed-effects models (further referred to as LMM repeatability). "Clade" contained three categories, as we found studies of only three clades: Euarchonta (data from the order Primates), Glires (data from the order Rodentia), and Laurasitheria (data from the orders Artiodactyla, Perissodactyla, and Carnivora). The "axis of personality" factor contained five categories of behaviours, corresponding to five axes of personality sensu Réale et al. (2007): activity, aggressiveness, boldness, exploration, and sociability.

We used a marginal linear model (gls function) as implemented in the nlme package under the R-environment (R version 3.2.2, package version 3.1-121; Pinheiro, Bates, DebRoy, Sarkar & R Core Team 2015). The dependent variable was the estimate of repeatability and the fixed factors in the full model were the "method", "number of replications", "number of individuals", "superorder", and "axis of personality". First, we compared the full model with the same model including the "study" as a random factor, using the maximum likelihood (ML) method. Then we reduced the model by excluding the non-significant fixed factors one at a time, while comparing the reduced model to the full model to make sure that the reduced model explains the same variance as the full one.

The model including the "study" as a random factor was significantly different from the same model without a random factor (L-ratio=15.831; $p<0.0001$), therefore the "study" was included as a random factor in all subsequent models.

In the full model, the "number of individuals" and "clade" factors were not significant ($p>0.05$). The reduced model was not significantly different from the full one (L-ratio=2.753; $p=0.413$) and included the "method" ($F=10.683$; $p<0.0001$), "number of replications" ($F=5.274$; $p=0.022$), and "axis of personality" factors ($F=4.503$; $p=0.002$). For a summary of the reduced model, see Table 1. We employed the Tukey HSD test to assess the significance of differences between factor levels. While the correlation and ANOVA repeatability do not differ significantly ($p=0.232$) in the "method" factor, the LMM repeatability produces lower estimates than both the correlation and ANOVA repeatability ($p=0.0002$ and $p=0.009$, respectively). This might mean that the LMM repeatability is either a stricter method, or is applied to less balanced designs, as recommended in Nakagawa & Schielzeth (2010). The Tukey HSD test for the factor "axis of personality" showed that the repeatability of activity was significantly higher than the repeatability of aggressiveness ($p=0.007$) and exploration ($p=0.012$) (see Fig. 4).

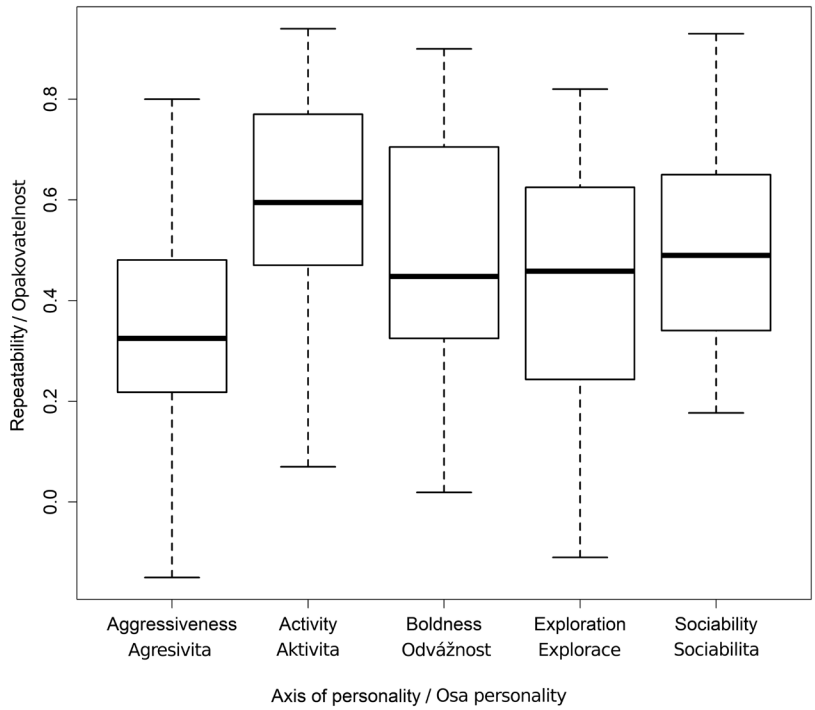

Fig. 4. Repeatability estimates for the five personality axes (Aggressiveness, Activity, Boldness, Exploration, Sociability) from our meta-analysis data.
Obr. 4. Hodnoty opakovatelnosti (Repeatability) pro pět personalitních os (Agresivita, Aktivita, Odvážnost, Explorační chování, Sociabilita).

We have found that the estimate of repeatability is significantly influenced by the choice of the statistical procedure, but also by the methodology of behavioural testing (there was a significant effect of study). We revealed a strong effect of the method used to calculate repeatability, however we believe this effect to be an artefact of the overall methodology. Studies using LMMs to estimate repeatability are usually sophisticated with more rigorous methodology, often with longer inter-test intervals, therefore the repeatability estimates are more sensitive to a random error.

We found that repeatability is lower in studies with a higher number of the test repetitions, which is in contrast with the findings of Biro (2012), who found repeatability improving with a higher number of repetitions. However, it is difficult to compare a meta-analysis and experimental data. The source studies were also all published before 2014, while the first comprehensive article stressing the importance of correcting for habituation and other consistent temporal trends was published in 2015 (Biro & Stamps 2015). None of the studies included in our meta-analysis used a correction for systematic temporal trends, which might have led to undervalued repeatability estimates (Biro & Stamps 2015, Žampachová et al. 2017).

The effect of the behaviour type ("axis of personality") might reflect the issues of measuring the respective behaviours. While activity is usually measured without any experimental disturbance and is therefore rather easy to estimate, aggressiveness and exploratory behaviour are more complicated. Moreover, activity can influence measurements of other behaviours and distort the results, which does not work reversely (Barnett & Cowan 1976, Réale et al. 2007). In both of these behaviours, the experience of tested animals with experimental conditions plays a large role. In the tests of aggressiveness, it is impossible to replicate the exact situation with the same opponents meeting for the first time. Sociability, on the contrary, does not show lower repeatability. Exploratory behaviour is also difficult to test repeatedly, as it relies on a reaction to "novelty". The question remains whether the novel stimuli still function as novel during the test repetitions. An alternative is to use a slightly different stimulus in every repeat of the test, but it is questionable, whether such repetitions are comparable.

We expected to find some effect of the clade, due to the differences in approaches described above and the inherent differences between the animals, but the effect was not significant. This could mean that personality in mammals appears in similar behaviours and we are able to quantify them with comparable accuracy. Moreover, we did not find any effect of the number of subjects. This shows that the studies all use a sufficient number of individuals and that the number of subjects does not create any systematic effect in the studies.

Our meta-analysis therefore shows that experimental studies of personality in different mammals are comparable, it is however advisable to pay close attention not only to the methodology of behavioural testing, but also to the details of statistical analysis.

Concluding remarks

Animal personality is a wide field of study with a variety of terms and various methodological approaches. Some of the terms are synonymous, some of them overlapping, some of them complementary. The definition of personality itself is simple and clear: personality traits are behavioural traits consistent both across time and context. Despite this, researchers often focus on either temporal or contextual consistency only.

In the literature review, we demonstrated on the example of two mammalian orders with the highest number of personality studies that even though the studies try to describe the same concept, their approaches can be so different that a comparison is complicated. The studies on primates are usually more comparable with the studies of human psychology, as the primate studies were originally inspired by psychological methods and therefore have a similar methodology. Rodent studies have a more detailed methodology, focused on quantifying observable behaviour, however, they often describe only one aspect of personality and somehow lack more comprehensive characteristics of personality. However, when we compared experimental studies of personality in two superorders (Euarchontoglires, Laurasiatheria), we found comparable results with no significant differences in the values of repeatability estimates.

It is tempting to study laboratory or domesticated animals, but in these animals the variability of their behaviour is reduced by the domestication process, as shown by de Boer and colleagues (de Boer et al. 2003). Considering that personality is in its core a study of variability, we believe that studies focusing on wild or wild-derived animals may bring many interesting results and can reveal new angles for future studies of personality, e.g., mechanisms in ontogeny of personality, where theoretical models do not always match the observations (Šimková et al. 2017) or specific selection pressures maintaining different personality types in the population (as demonstrated in the black rat, *Rattus rattus*; Žampachová et al. 2017). Wild and wild-deri-

ved animals can answer many questions about the ecological relevance of personality and the interaction of different personality types with the environment that studies on domesticated or laboratory animals cannot.

The most widely used personality tests (the open field test for measuring exploratory behaviour and/or boldness, the resident intruder test for measuring aggressiveness) are also the most universal ones, however, we recommend a careful consideration should be paid when choosing tests suitable for the studied species and in some cases, an adaptation of the standard tests might be required. Based on our previous results (Žampachová et al. 2017), we suggest three to four repetitions of each test to assess the temporal stability and whether the behaviour is subject to habituation. However, our meta-analysis presented above suggests that repeatability is lower with more repetitions of the testing procedure. Therefore, we advise to correct for habituation during a statistical analysis, especially in the tests of aggressiveness and exploratory behaviour, as it might provide more accurate estimates of repeatability. We also recommend testing the same behaviour in different contexts (e.g., different tests of exploratory behaviour) to assess the context generality of behaviour.

We would like to point out that personality studies in mammals focus either on "traditional" orders with specific methodology (primates, rodents) or on large attractive animals (carnivores, ungulates). There are only a few studies concerning other orders, whose unique ecology (e.g., bats) might provide a new insight into the development and structure of animal personality. We encourage zoologists working with many interesting and ecologically unique species to consider personality in their studies, as personality framework may provide more information about the link between the individual variability in behaviour and autecology of many mammalian species.

SOUHRN

Studie zabývající se personalitou a individuálními rozdíly v chování v posledních letech nabývají na popularitě. V tomto přehledu a meta-analýze chceme představit koncept personality a podmínky, které musí prvek chování splňovat, aby mohl být považován za osobnostní rys. Z definice personality to je konzistence (1) v čase a (2) v různých kontextech. U savců je personalita studovaná především u dvou řádů, a to u primátů (Primates) a hlodavců (Rodentia). Popisujeme zde různé přístupy k měření personality u zástupců těchto dvou řádů. Primatologické studie personality vycházejí z lidské psychologie a často spoléhají na skórování subjektů pozorovatelem, zatímco studie na hlodavcích vycházejí spíše z fyziologického výzkumu a používají více experimentální přístup. Dále poskytujeme náhled do metodických úskalí použití koeficientu opakovatelnosti jako statistického nástroje pro měření konzistence chování v čase. Klasické metody výpočtů opakovatelnosti neuvažují habituaci a další trendy, které mohou zkreslit výsledky a vést k příliš nízkým hodnotám opakovatelnosti. Také diskutujeme konzistenci chování v různých kontextech a různé pohledy na definici kontextu. Variabilitu ve studiích savčí personality ilustrujeme meta-analýzou opakovatelnosti. Našli jsme závislost opakovatelnosti na metodice behaviorálních testů a statistické analýze, ale také jsme našli signifikantní efekt počtu opakování daného testu a rozdíly mezi různými okruhy chování. Opakovatelnost klesala s rostoucím počtem opakování a testy agresivity a exploračního chování se ukázaly méně opakovatelné než testy aktivity.

Acknowledgements

The project was funded by Charles University, GA UK No. 900217. This study is a result of research funded by the project Nr. LO1611 with a financial support from the MEYS under the NPU I program. We also thank Silvie Rádlová for her help with figures formatting, František Sedláček and Jakub Polák for their critical reading of the manuscript, and Barbora Kaftanová for providing the tests photographs.

REFERENCES

Adams M. J., Majolo B., Ostner J., Schülke O., De Marco A., Thierry B., Engelhardt A., Widdig A., Gerald M. S. & Weiss A., 2015: Personality structure and social style in macaques. *Journal of Personality and Social Psychology*, **109**: 338–353.

Archer J., 1973: Tests for emotionality in rats and mice: a review. *Animal Behaviour*, **21**: 205–235.

Aulich D., 1976: Escape versus exploratory activity: An interpretation of rats' behaviour in the open field and a light-dark preference test. *Behavioural Processes*, **1**: 153–164.

Barnett S. A., 1958: Experiments on neophobia in wild and laboratory rats. *British Journal of Psychology*, **49**: 195–201.

Barnett S. A. & Cowan P. E., 1976: Activity, exploration, curiosity and fear: an ethological study. *Interdisciplinary Science Reviews*, **1**: 43–62.

Bell A. M., Hankison S. J. & Laskowski K. L., 2009: The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **77**: 771–783.

Benus R., Koolhaas J. M. & Van Oortmerssen G. A., 1987: Individual differences in behavioural reaction to a changing environment in mice and rats. *Behaviour*, **100**: 105–122.

Benus R., Bohus B., Koolhaas J. M. & Van Oortmerssen G. A., 1989: Behavioural strategies of aggressive and non-aggressive male mice in active shock avoidance. *Behavioural Processes*, **20**: 1–12.

Benus R., Bohus B., Koolhaas J. M. & Van Oortmerssen G. A., 1991: Heritable variation for aggression as a reflection of individual coping strategies. *Experientia*, **47**: 1008–1019.

Biro P. A., 2012: Do rapid assays predict repeatability in labile (behavioural) traits? *Animal Behaviour*, **83**: 1295–1300.

Biro P. A. & Stamps J. A., 2015: Using repeatability to study physiological and behavioural traits: ignore time-related change at your peril. *Animal Behaviour*, **105**: 223–230.

de Boer S. F., van der Vegt B. J. & Koolhaas J. M., 2003: Individual variation in aggression of feral rodent strains: A standard for the genetics of aggression and violence? *Behavior Genetics*, **33**: 485–501.

Boissier J. & Simon P., 1962: La réaction d'exploration chez la souris (note Préliminaire). *Thérapie*, **17**: 1226–1232.

Careau V. & Garland T., 2012: Performance, personality, and energetics: correlation, causation, and mechanism. *Physiological and Biochemical Zoology*, **85**: 543–571.

Careau V., Montiglio P. O., Garant D., Pelletier F., Speakman J. R., Humphries M. M. & Réale D., 2015: Energy expenditure and personality in wild chipmunks. *Behavioral Ecology and Sociobiology*, **69**: 653–661.

Carere C. & Eens M., 2005: Unravelling animal personalities: how and why individuals consistently differ. *Behaviour*, **142**: 1155–1163.

Čiháková J. & Frynta D., 1996: Intraspecific and interspecific behavioural interaction in the wood mouse (*Apodemus sylvaticus*) and the yellow-necked mouse (*Apodemus flavicollis*) in a neutral cage. *Folia Zoologica*, **45**: 105–113.

Coppens C. M., de Boer S. F. & Koolhaas J. M., 2010: Coping styles and behavioural flexibility: towards underlying mechanisms. *Philosophical Transactions of the Royal Society B*: *Biological Sciences*, **365**: 4021–4028.

Cowan P. E., 1976: The new object reaction of *Rattus rattus* L: The relative importance of various cues. *Behavioral Biology*, **16**: 31–44.

Cowan P. E., 1977: Neophobia and neophilia: New-object and new-place reactions of three *Rattus* species. *Journal of Comparative and Physiological Psychology*, **91**: 63–71.

Dall S. R. X., Houston A. I. & McNamara J. M., 2004: The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters*, **7**: 734–739.

Dammhahn M. & Almeling L., 2012: Is risk taking during foraging a personality trait? A field test for cross-context consistency in boldness. *Animal Behaviour*, **84**: 1131–1139.

David J. T., Cervantes M. C., Trosky K. A., Salinas J. A. & Delville Y., 2004: A neural network underlying individual differences in emotion and aggression in male golden hamsters. *Neuroscience*, **126**: 567–578.

Deacon R. M. J., 2013: The successive alleys test of anxiety in mice and rats. *Journal of Visualized Experiments*, **76**(e2705): 1–7.
DeFries J. C., Gervais M. C. & Thomas E. A., 1978: Response to 30 generations of selection for open-field activity in laboratory mice. *Behavior Genetics*, **8**: 3–13.
Dewsbury D. A., 1980: Wheel-running behavior in 12 species of muroid rodents. *Behavioural Processes*, **5**: 271–280.
Digman J. M., 1990: Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, **41**: 417–440.
Dingemanse N. J. & Réale D., 2013: What is the evidence that natural selection maintains variation in animal personalities? Pp.: 201–220. In: Carere C. & Maestripieri D. (eds.): *Animal Personalities*: *Behavior*, *Physiology*, *and Evolution 2013*. University of Chicago Press, Chicago, 520 pp.
Dingemanse N. J., Kazem A. J. N., Réale D. & Wright J., 2010: Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology and Evolution*, **25**: 81–89.
Edenbrow M. & Croft D. P., 2012: Sequential hermaphroditism and personality in a clonal vertebrate: the mangrove killifish. *Behavioural Processes*, **90**: 229–237.
File S. E., 2001: Factors controlling measures of anxiety and responses to novelty in the mouse. *Behavioural Brain Research*, **125**: 151–157.
Frynta D., Slábová M., Váchová H., Volfová R. & Munclinger P., 2005: Aggression and commensalism in house mouse: A comparative study across Europe and the Near East. *Aggressive Behavior*, **31**: 283–293.
Gartner M. & Weiss A., 2013: Personality in felids: A review. *Applied Animal Behaviour Science*, **144**: 1–13.
Glowa J. R. & Hansen C. T., 1994: Differences in response to an acoustic startle stimulus among forty-six rat strains. *Behavior Genetics*, **24**: 79–80.
Gosling S. D., 1998: Personality dimensions in spotted hyenas (*Crocuta crocuta*). *Journal of Comparative Psychology*, **112**: 107–118.
Gosling S. D. & John O. P., 1999: Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, **8**: 69–75.
Groothuis T. G. G. & Trillmich F., 2011: Unfolding personalities: The importance of studying ontogeny. *Developmental Psychobiology*, **53**: 641–655.
Guenther A., Finkemeier M. & Trillmich F., 2014: The ontogeny of personality in the wild guinea pig. *Animal Behaviour*, **90**: 131–139.
Guenther A. & Trillmich F., 2012: Photoperiod influences the behavioral and physiological phenotype during ontogeny. *Behavioral Ecology*, **24**: 402–411.
Hall C. S., 1934: Emotional behavior in the rat. *Journal of Comparative Psychology*, **18**: 385–403.
Hayes J. P. & Jenkins S. H., 1997: Individual variation in mammals. *Journal of Mammalogy*, **78**: 274–293.
Heyser C. J. & Chemero A., 2012: Novel object exploration in mice: Not all objects are created equal. *Behavioural Processes*, **89**: 232–238.
Hjelle L. A. & Ziegler D. J., 1992: *Personality Theories*: *Basic Assupmtions*, *Research*, *and Applications*. *Third Edition*. McGraw-Hill, New York, 603 pp.
Hudson R., Rangassamy M., Saldaña A., Bánszegi O. & Rödel H. G., 2015: Stable individual differences in separation calls during early development in cats and mice. *Frontiers in Zoology*, **12**(Suppl. 1) (S12): 1–12.
Knab A. M., Bowen R. S., Moore-Harrison T., Hamilton A. T., Turner M. J. & Lightfoot J. T., 2009: Repeatability of exercise behaviors in mice. *Physiology & Behavior*, **98**: 433–440.
Konečná M., Weiss A., Lhota S. & Wallner B., 2012: Personality in Barbary macaques (*Macaca sylvanus*): Temporal stability and social rank. *Journal of Research in Personality*, **46**: 581–590.
Koolhaas J. M., Korte S. M., De Boer S. F., Van Der Vegt B. J., Van Reenen C. G., Hopster H., De Jong I. C., Ruis M. & Blokhuis H. J., 1999: Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, **23**: 925–35.
Koolhaas J. M., de Boer S. F., Coppens C. M. & Buwalda B., 2010: Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. *Frontiers in Neuroendocrinology*, **31**: 307–321.

Koolhaas J. M., Coppens C. M., de Boer S. F., Buwalda B., Meerlo P. & Timmermans P. J. A., 2013: The resident-intruder paradigm: A standardized test for aggression, violence and social stress. *Journal of Visualized Experiments*, **77**(e4367): 1–7.
Korpela K., 2011: Sex influences rat personality more than geographical origin. *Applied Animal Behaviour Science*, **133**: 95–100.
Koski S. E., 2011: Social personality traits in chimpanzees: temporal stability and structure of behaviourally assessed personality traits in three captive populations. *Behavioral Ecology and Sociobiology*, **65**: 2161–2174.
Koski S. E. & Burkart J. M., 2015: Common marmosets show social plasticity and group-level similarity in personality. *Scientific Reports*, **5**(8878): 1–7.
Lantová P., Šíchová K., Sedláček F. & Lanta V., 2011: Determining behavioural syndromes in voles – the effects of social environment. *Ethology*, **117**: 124–132.
Lessells C. & Boag P., 1987: Unrepeatable repeatabilities: a common mistake. *The Auk*, **104**: 116–121.
Martin J. G. A. & Réale D., 2008a: Temperament, risk assessment and habituation to novelty in eastern chipmunks, *Tamias striatus*. *Animal Behaviour*, **75**: 309–318.
Martin J. G. A. & Réale D., 2008b: Animal temperament and human disturbance: Implications for the response of wildlife to tourism. *Behavioural Processes*, **77**: 66–72.
Massen J. J. M. & Koski S. E., 2014: Chimps of a feather sit together: chimpanzee friendships are based on homophily in personality. *Evolution and Human Behavior*, **35**: 1–8.
McCrae R. R. & John O. P., 1992: An introduction to the five-factor model and its applications. *Journal of Personality*, **60**: 175–215.
McGraw K. O. & Wong S. P., 1996: Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, **1**: 30–46.
Meek T. H., Lonquich B. P., Hannon R. M. & Garland T., 2009: Endurance capacity of mice selectively bred for high voluntary wheel running. *Journal of Experimental Biology*, **212**: 2908–2917.
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**: 282–298.
Munclinger P. & Frynta D., 2000: Social interactions within and between two distant populations of house mouse. *Folia Zoologica*, **49**: 1–6.
Nakagawa S. & Schielzeth H., 2010: Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society*, **85**: 935–956.
Van Oortmerssen G. A. & Bakker T. C. M., 1981: Artificial selection for short and long attack latencies in wild *Mus musculus domesticus*. *Behavior Genetics*, **11**: 115–116.
Pellow S., Chopin P., File S. E. & Briley M., 1985: Validation of open: closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods*, **14**: 149–167.
Petelle M. B., Mccoy D. E., Alejandro V., Martin J. G. A. & Blumstein D. T., 2013: Development of boldness and docility in yellow-bellied marmots. *Animal Behaviour*, **86**: 1147–1154.
Pinheiro J., Bates D., DebRoy S., Sarkar D. & Core Team, 2015: *_nlme*: *Linear and Nonlinear Mixed Effects Models*. *R package Version 3.1–121*. URL: CRAN.R-project.org/package?nlme.
Réale D. & Dingemanse N. J., 2012: *Animal Personality*. *eLS*. John Wiley & Sons, Ltd, Chichester.
Réale D., Reader S. M., Sol D., McDougall P. T. & Dingemanse N. J., 2007: Integrating animal temperament within ecology and evolution. *Biological Reviews of the Cambridge Philosophical Society*, **82**: 291–318.
Réale D., Dingemanse N. J., Kazem A. J. N. & Wright J., 2010: Evolutionary and ecological approaches to the study of personality. *Philosophical Transactions of the Royal Society of London*, *Series B*, *Biological Sciences*, **365**: 3937–3946.
Robinson L. M., Morton F. B., Gartner M. C., Widness J., Paukner A., Essler J. L., Brosnan S. F. & Weiss A., 2016: Divergent personality structures of brown (*Sapajus apella*) and white-faced capuchins (*Cebus capucinus*). *Journal of Comparative Psychology*, **130**: 305–312.

Robinson L. M., Altschul D. M., Wallace E. K., Úbeda Y., Llorente M., Machanda Z., Slocombe K. E., Leach M. C., Waran N. K. & Weiss A., 2017: Chimpanzees with positive welfare are happier, extraverted, and emotionally stable. *Applied Animal Behaviour Science*, **191**: 90–97.
Rudin F. S. & Briffa M., 2012: Is boldness a resource-holding potential trait? Fighting prowess and changes in startle response in the sea anemone, *Actinia equina*. *Proceedings of the Royal Society*, *Biological Sciences*, **279**: 1904–10.
Sherwin C., 1998: Voluntary wheel running: a review and novel interpretation. *Animal Behaviour*, **56**: 11–27.
Sih A. & Bell A. M., 2008: Insights for behavioral ecology from behavioral syndromes. *Advances in the Study of Behavior*, **38**: 227–281.
Sih A., Bell A., Johnson J. & Ziemba R., 2004a: Behavioral syndromes: An integrative overview. *Quarterly Review of Biology*, **79**: 241–277.
Sih A., Bell A. & Johnson J. C., 2004b: Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology and Evolution*, **19**: 372–378.
Staes N., Koski S. E., Helsen P., Fransen E., Eens M. & Stevens J. M. G., 2015: Chimpanzee sociability is associated with vasopressin (Avpr1a) but not oxytocin receptor gene (OXTR) variation. *Hormones and Behavior*, **75**: 84–90.
Staes N., Weiss A., Helsen P., Korody M., Eens M. & Stevens J. M. G., 2016: Bonobo personality traits are heritable and associated with vasopressin receptor gene 1a variation. *Scientific Reports*, **6**: 1–8.
Stamps J. & Groothuis T. G. G., 2010a: Developmental perspectives on personality: implications for ecological and evolutionary studies of individual differences. *Philosophical Transactions of the Royal Society of London*, *Series B*, *Biological Sciences*, **365**: 4029–4041.
Stamps J. & Groothuis T. G. G., 2010b: The development of animal personality: Relevance, concepts and perspectives. *Biological Reviews*, **85**: 301–325.
Suchomelova E. & Frynta D., 2000: Intraspecific behavioural interactions in *Apodemus microps*: a peaceful mouse? *Acta Theriologica*, **45**: 201–209.
Suchomelová E., Munclinger P. & Frynta D., 1998: New evidence of pseudosexual behaviour and female aggression in mice: neutral cage interactions in *Mus spicilegus* and *Mus spretus* (Rodentia: Muridae). *Folia Zoologica*, **47**: 241–247.
Svartberg K., 2005: A comparison of behaviour in test and in everyday life: evidence of three consistent boldness-related personality traits in dogs. *Applied Animal Behaviour Science*, **91**: 103–128.
Šimková O., Frýdlová P., Žampachová B., Frynta D. & Landová E., 2017: Development of behavioural profile in the Northern common boa (*Boa imperator*): Repeatable independent traits or personality? *Public Library of Science One*, **12**(5) (e0177911): 1–35.
Via S., Gomulkiewicz R., De Jong G., Scheiner S. M., Schlichting C. D. & Van Tienderen P. H., 1995: Adaptive phenotypic plasticity: consensus and controversy. *Trends in Ecology and Evolution*, **10**: 212–217.
Walsh R. N. & Cummins R. A., 1976: The open-field test: a critical review. *Psychological Bulletin*, **83**: 482–504.
Weiss A., Adams M. J. & Johnson W., 2011: The big none: No evidence for a general factor of personality in chimpanzees, orangutans, or rhesus macaques. *Journal of Research in Personality*, **45**: 393–397.
Weiss A., King J. E. & Hopkins W. D., 2007: A cross-setting study of chimpanzee (*Pan troglodytes*) personality structure and development. *American Journal of Primatology*, **69**: 1264–1277.
Weiss A., Miho I., Hong K., 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*, **292**: 283–292.
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**: 1355–1365.
West-Eberhard M. J., 2003: *Developmental Plasticity and Evolution*. Oxford University Press, Oxford, 794 pp.
Wolf M., Van Doorn G. S., Leimar O. & Weissing F. J., 2007: Life-history trade-offs favour the evolution of animal personalities. *Nature*, **447**: 581–584.
Žampachová B., Kaftanová B., Šimánková H., Landová E. & Frynta D., 2017: Consistent individual differences in standard exploration tasks in the black rat (*Rattus rattus*). *Journal of Comparative Psychology*, **131**: 150–162.