MASTERARBEIT | MASTER'S THESIS

Titel | Title

Comparison of short-term and long-term repeatability in behaviour of guinea pigs

verfasst von | submitted by

Moritz Eckart Albrecht Wiehler B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 878
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Cognition, Behavior and Neurobiology
Betreut von | Supervisor:	ao. Univ.-Prof. Dr. Eva Millesi
Mitbetreut von | Co-Supervisor:	Mag. Matthias Nemeth PhD

Acknowledgements

First, I want to thank Matthias, who made all this possible in the first place. The way we were able to communicate, discuss ideas, and build upon each other's made it very enjoyable for me. Secondly, I want to thank Eva Millesi, who had a critical eye on things and supported me on the academic side of things. Next, I want to thank the people who directly aided me in data collection, Elena, Eva, and Kiwi. I want to thank the animal caretakers who looked after and fed the guinea pigs. I want to thank my parents for giving me both financial and emotional support all through-out my life and this journey. And most of all I want to thank the person who was always by my side, gave me the motivation to sit down and work on this thesis, and had calming words for me whenever I felt like I could not deal with all of this. Thank you, Eva.

Table of contents

1. Abstract

Repeatability estimation is the proportion of variance explained by between-individual variation, as opposed to within-individual variation, and allows us to quantify the reliability of a measurement. With regard to behaviour, this allows finding consistent personality traits, as well as looking into the importance of behavioural flexibility throughout an animal's life. It has been shown that behavioural repeatability changes depending on life stages and that repeatability decreases with longer inter-test intervals. In the present study, male guinea pigs were tested in three behavioural tests (light-dark test, sociality test, open-field test) four times, including a short (2-week), a long (8-week), and another short inter-test interval. The resulting repeatability estimations were compared between time intervals and behaviours. Repeatability estimates were very high for risk-taking and anxiety-like behaviour as well as for general activity. Repeatability increased over time, showing lower repeatability for the first short-term interval and the long-term interval, and high repeatability for the second short-term interval. However, the sociality test to determine social interest in an unfamiliar conspecific showed very low repeatability estimates throughout, which is discussed in regard to the experimental setup. Overall, the performed tests other than the sociability test repeatably and reliably showed clear consistency within individuals in behavioural traits in the short- and long-term.

2. Introduction and theoretical background

"When individuals behave consistently through time and when individuals behave differently from each other, then the behaviour is repeatable", is how Bell *et al.* (2009) described repeatability in behavioural research. In other words, behavioural consistency is an animal's trait showing repeatability in behavioural responses. In a broader sense, repeatability estimation is a statistical tool to explain variance in data. It allows us to quantify the reliability of a measurement by comparing it to measurements at other points in time (Stoffel *et al.* 2017). Repeatability is an important metric in determining flexibility of behavioural responses, rating its ecological relevance and evolutionary significance (Brust & Guenther 2017), as flexibility or consistency of behaviour can play an important role in an animal's fitness (Dingemanse *et al.* 2002). However, repeatability is not to be confused with animal personality, which includes behavioural consistency over contexts (functional categories of behaviour, like feeding, mating, or parental care; van Oers *et al.* 2005), situations (sets of given conditions at a certain point in time; van Oers *et al.* 2005), and measures, in addition to the aforementioned temporal component (Bergmüller 2010, Briffa & Weiß 2010). Repeatability is an important tool to measure personality (Briffa & Weiß 2010), but analysing personality encompasses more requirements (Spoolder *et al.* 1996). Note that this thesis is looking only into repeatability, as the connections to personality are more complex and would require further and more detailed analyses.

Repeatability can differ between taxa, types of behaviours, contexts, sexes, and the time interval in between measurements (Bell *et al.* 2009). The repeatability is expected to decrease with increasing time intervals between observations. This is due to changing environmental effects, shifting genetic expressions, or differing conditions or states of an animal in comparison to shorter intervals, where these are more likely to stay stable. Hunger-levels, body size, age, health conditions, and dominance hierarchy are some examples for such variables that might or will change after a longer interval. Higher repeatability has been found in behaviours related to mating (Smith & Hunter 2005), habitat selection (Leclerc *et al.* 2016), and aggression (Brust & Guenther 2017). In contrast, activity (Biro 2012), mate preference (Forstmeier & Birkhead 2004), and migration (van Wijk *et al.* 2016) were shown to display very low repeatability in previous studies. Repeatability has been studied in a lot of different species including insects, birds, and amphibians. Interestingly, there are comparably few studies looking into the repeatability of behaviour in mammals (Bell *et al.* 2009), and repeatability and personality in guinea pigs has only started being researched in the last ten to 15 years (see e.g., Zipser *et al.* 2013).

Guinea pigs (*Cavia porcellus*) are social rodents descending from a wild form (*Cavia aperea*) originating from grass lands of South America. Guinea pigs live in complex social hierarchies that hardly ever change (Sachser 1986) and are studied regarding stress responses, hierarchical dominance structures, and behaviour (Wagner & Manning 1976). In previous studies, personality and behavioural repeatability in guinea pigs as well as repeatability of endocrine traits and dominance rank over different time frames were tested. High repeatability over an 8-week interval was found for male-female interactions (though not for sexual behaviour), cortisol reactivity, as well as dominance status, while emotional behaviour and baseline cortisol concentrations were not found to be repeatable (Zipser *et al.* 2013). This can be explained by the fact that baseline measurements of cortisol must be considered to reflect previous situations. Depending on situations within the home enclosure prior to the measurement, the animals might have been experiencing social interactions or other stressors, leading to increased cortisol levels (Masis-Calvo *et al.* 2018). Additionally, time of day (Garris 1979) or prior food intake (Nemeth *et al.* 2014) can influence cortisol concentrations. Therefore, baseline measurements of glucocorticoids in general show higher fluctuation rates. Nonetheless, Mutwill *et al.* (2021, 2023) found repeatability in both baseline cortisol and cortisol reactivity in male guinea pigs. Neither the low repeatability of baseline cortisol nor the higher repeatability of cortisol reactivity differed between pair and group housing conditions (Mutwill *et al.* 2023). The stability in rank and cortisol reactivity is also supported in females (Rystrom *et al.* 2022), though a lot of studies focus on males. Repeatability in behaviour was found in learning behaviour, as well as sociopositive behaviour, aggressive behaviour, and boldness determined in a novel object task (Brust & Guenther 2017). The link between endocrine traits and behavioural repeatability is still not clear though. And while endocrine traits and social hierarchies are well investigated regarding repeatability and results of different studies are mostly consistent, behavioural repeatability in guinea pigs is less well studied and results remain controversial.

The aim of the present study is to fill this gap in current knowledge by investigating individual repeatability in behavioural responses of male guinea pigs over shorter two-week and longer two-month intervals. The behavioural assays chosen for this purpose include (1) a light-dark test focusing on anxiety, activity, and exploration, (2) a sociality test to investigate social interest and activity in a social context, and (3) an open field test as another measurement of activity, exploration, and anxiety. Additionally, baseline saliva cortisol concentrations, body mass, and individual hierarchy ranks/indices were analysed to gain insight in behavioural repeatability in the context of potential stress responses and/or changes in hierarchies. As Bell et al. (2009) found higher repeatability in shorter compared to longer inter-test intervals, this master's thesis will also address different time intervals between measurements. For this, differences between short- term and long-term repeatability will be compared to gain insight into the degree of personality and/or stability of behavioural phenotypes. It is anticipated that behavioural consistency in the form of repeatability will be measured as described in prior work and the repeatability is expected to drop over the longer break. Finding repeatability in certain variables within the behaviour of male guinea pigs will also lead to a better understanding of the meaningfulness and significance of the behavioural tests used. Repeatability in behaviours related to anxiety in the light-dark test, for example, would support the light-dark test as a robust measure for anxiety in male guinea pigs.

3. Methods

2.1 Animals and Husbandry

The animals used for this study were 14 intact, adult (age 18 ± 1 months), male guinea pigs (*Cavia porcellus*), who were housed in the animal care facility of the UBB. They were kept in small single-sex social constellations of two to five individuals in enclosures (1.4 m^2 – 3.2 m^2 depending on group size) made of chip or hardboard. Each enclosure was equipped with shelters (40 cm x 60 cm) and litter covering the floor. Ambient conditions, including a temperature of 22 ± 2 °C, a humidity of 50 ± 5 %, and a 12h/12h light-dark-cycle (lights on at 06:00), were maintained throughout the study. Food pellets (ssniff V-2233; ssniff Spezialdiäten GmbH, Soest, Germany), hay, and water were available *ad libitum*. The animals used in this study were bred within the facility. The males are separated from their mothers shortly before reaching sexual maturity at around 25 days of age and kept in stable single-sex social groups, while the females stay in the larger female colony.

2.2 Experimental procedure

Three behavioural tests were performed on consecutive days within a week: first a light-dark test (Fig. 1A), then a sociality test (Fig. 1B), and finally an open field test (Fig. 1C). This set-up was then repeated two weeks later to determine short-term repeatability and the whole block was carried out again after two months to evaluate long-term repeatability and to compare short-term repeatability before and after the longer interval. The four weeks of testing will be referred to as weeks 1, 3, 12 and 14. For behavioural testing, an animal was taken from its home enclosure in the housing room, transferred to the respective test enclosure set up in an adjacent testing room, and the experimenter(s) left the room. Each test was carried out for ten minutes and filmed by a webcam overhead streaming the footage onto a computer in the housing room, from where the video could be viewed live and was recorded for later behavioural analyses using BORIS (v. 8.21.8, 2023, Friard & Gamba 2016). The animal was returned to its home enclosure after the test, and the test setup was cleaned for the next test. The animals were always tested at the same time (between 10:00 and 15:00) and in the same order in each round of testing to eliminate time-of-day effects. In addition to the behavioural assays, the animals' body mass was measured using a regular kitchen scale and baseline saliva cortisol concentrations were recorded once a week before performing the open-field tests. Additionally, event samplings and *ad libitum* samplings were performed to evaluate social dominance hierarchies.

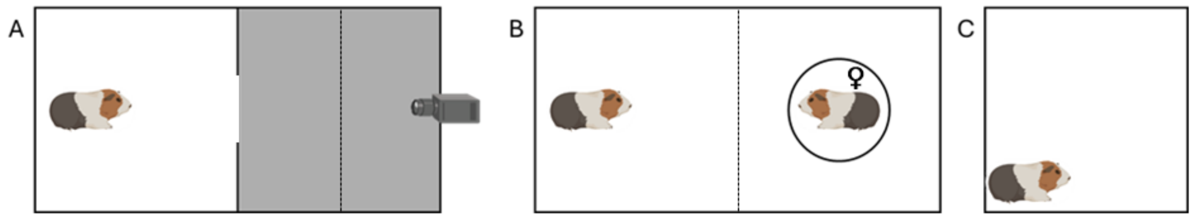

Fig. 1: Schematic illustrations of the three behavioural tests. All tests were filmed by a camera overhead. A: Light-dark test with an additional camera filming the inside of the dark compartment. The dashed line indicates the separation between the front and the back half of the dark compartment. B: Sociality test with female in separate circular wire

mesh enclosure. The dashed line indicates the separation between the empty and the social area of the arena. C: Open field test.

2.3 Light-dark test

The light-dark test was performed in a 120 cm x 60 cm test enclosure made from 40 cm tall white hardboards split into two even compartments, divided by a wall containing a 20cm x 20 cm opening. One of the compartments was covered by a board to darken the inside with an opening to film the inside. The animals were placed along the far wall of the light compartment (Fig. 1A). Video analysis started as soon as the focal animal was placed in the test enclosure and the experimenter stepped back. From then on, 10 minutes were analysed. The following variables were analysed: latency to enter the dark compartment for the first time, total time in the dark compartment, duration of locomotion, and total time spent in the back area of the dark compartment (= the 30 cm in the dark compartment further from the opening).

2.4 Sociality test

For the sociality test, a rectangular white hardboard arena (120 cm x 60 cm, 40 cm height) was set up with a circular wire mesh enclosure (25 cm diameter, 30 cm height) in the centre of one half of the arena. That side was therefore determined to be the “social side” while the other half was deemed the “empty side”. Markings on the top of the walls visible on camera were used to separate the halves of the arena for analysis. A female guinea pig randomly selected from the colony was placed within the wire mesh enclosure shortly before the test began and returned to its home colony after the test was over. The focal animal was placed along the far wall of the empty side of the arena (Fig. 1B). Video analysis started as soon as the focal animal was placed in the test enclosure and the experimenter stepped back. From then on, 10 minutes were analysed. The following variables were analysed: total time on the social side, duration of locomotion, and the time of social interest, which was defined as the focal animal being within the social side with the snout facing the cage.

2.5 Open-field test

Two animals were sampled simultaneously by two experimenters and then placed into a corner each of two open-field arenas made of brown hardboard (100 cm x 100 cm, 40 cm height; Fig. 1C). The two animals were analysed independently of each other, with their

10 minutes of video analysis starting as soon as the focal animal was placed in the test enclosure and the experimenter stepped back. The following variables were analysed: duration of locomotion, time spent in the centre of the arena, which was defined as any time an animal stepped more than 10 cm away from the walls, and the number of times the animals entered the centre.

2.6 Saliva samples and cortisol analysis

To collect saliva samples, standard cotton buds were inserted into the animal's mouth and moved around gently for 1 min. The buds were centrifuged (10,000 rpm, 5 min) in Eppendorf tubes and the resulting saliva was stored at -20 °C. Cortisol concentrations in saliva were analysed by enzyme-linked immunosorbent assays (IBL Hamburg GmbH; Hamburg, Germany) after 1:10 dilution following the manufacturer's instructions. All analyses were run in duplicates. Intra and inter coefficients of variance of control samples were 5.7% and 9.1%, respectively. The biological validity of saliva cortisol analyses in guinea pigs was demonstrated previously (Nemeth et al. 2016).

2.7 Social dominance hierarchy

To determine social dominance hierarchies within the groups, event/*ad libitum* samplings were performed to identify agonistic behaviours. The individuals initiating and receiving these behaviours were noted to determine a rank order within the groups. The Coulon-Index was calculated based on the ratio between displacements and retreats (Coulon 1975). Due to the low number of interactions, the hierarchy could only be calculated for test block one (weeks 1 and 3) versus test block two (weeks 12 and 14), instead of once per week of testing.

2.8 Statistical analysis

Statistics were performed in R (v 4.4.3, R Core Team 2025). Data were first analysed using linear mixed models (LMMs; lme4_1.1-37, Bates *et al.* 2015, glmmTMB_1.1.11, Brooks *et al.* 2017, emmeans_1.11.0, Lenth 2025). Model residuals were inspected for normality and homoscedasticity, and, if necessary, the data was log-transformed prior to analysis. Repeatability for all variables was then calculated over all four measurement points in time, as well as separated by intervals (comparing only week 1 and 3 or 12 and 14 for short-term or only week 3 and 12 for long-term repeatability) using the rptR package (rptR_0.9.23, Stoffel *et al.* 2017). For the repeatability analysis, normality of data was assumed (datatype = Gaussian), so log-transformed data were used whenever necessary. To compare the repeatability results of different behaviours over different time intervals the p- and R-values as well as the confidence intervals were compiled and compared using a linear mixed model. In total, more behavioural parameters were analysed for the study than presented here, as presenting all parameters would have exceeded the framework of this work. Nonetheless all behaviours measured were used within the comparison. All results were visualized in ggplot2 (ggplot2_3.5.2, Wickham 2016).

4. Results

<u>3.1 Light-Dark Test</u>

The animal's latency to first enter the dark compartment of the light-dark test (Fig. 2A) changed only marginally ($F(3, 39) = 2.415$, $p = 0.081$) and slightly decreased over time. The animals displayed significantly lower latency to enter the dark in week 14 compared to week 1 ($p = 0.016$; weeks 1-12: $p = 0.007$; all other comparisons: $p > 0.165$). The latency to enter the dark compartment was repeatable over the four weeks of testing ($R = 0.395$, $p < 0.001$, CI = [0.058, 0.645]). However, no repeatability was found for the first short-term interval (weeks 1-3: $R = 0.079$, $p = 0.384$, CI = [0, 0.566]), but the long-term and the second short-term interval showed repeatability in the latency to enter the dark compartment (weeks 3-12: $R = 0.477$, $p = 0.029$, CI = [0, 0.78]; weeks 12-14: $R = 0.664$, $p = 0.002$, CI = [0.234, 0.88]).

The time the animals spent in the light compartment of the arena (Fig. 2B) did not change significantly over time ($F(3,39) = 0.572$, $p = 0.637$). The time in the light compartment was repeatable over the four weeks ($R = 0.358$, $p = 0.002$, CI = [0.061, 0.616]) and over the long interval (weeks 3-12: $R = 0.463$, $p = 0.033$, CI = [0, 0.782]), but it was not repeatable over the short intervals (weeks 1-3: $R = 0.118$, $p = 0.329$, CI = [0, 0.595]; weeks 12-14: $R = 0.328$, $p = 0.103$, CI = [0, 0.782]). Note that the time the animal spent in the light compartment, as analysed here, represents reciprocal results for the time spent in the dark compartment.

The time the animals spent in the half of the dark compartment further from the entrance (Fig. 2C), the back, did not change significantly over time ($F(3, 39) = 0.317$, $p = 0.813$). The duration was repeatable over time (R = 0.375, p = 0.00123, CI = [0.077, 0.629]). The first short-term interval was marginally repeatable (weeks 1-3: R = 0.342, p = 0.094, CI = [0, 0.741]), the long-term interval showed no repeatability (weeks 3-12: R = 0.207, p = 0.217, CI = [0, 0.654]), and the second short-term interval showed repeatability (weeks 12-14: R = 0.511, p = 0.02, CI = [0, 0.805]).

The animals' time spent in locomotion during the light-dark test (Fig. 2D) decreased significantly over time (F(3, 38) = 14.404, p < 0.001), dropping off after week 1 (p < 0.001) and staying consistent afterwards (p = 1). The time of locomotion was repeatable overall (R = 0.461, p < 0.001, CI = [0.113, 0.686]), only marginally repeatable when comparing weeks 1 and 3 (R = 0.396, p = 0.078, CI = [0, 0.769]), but significantly repeatable over both the long interval from week 3 to 12 (R = 0.498, p = 0.023, CI = [0, 0.814]), and the second short interval from week 12 to 14 (R = 0.619, p = 0.005, CI = [0.101, 0.861]).

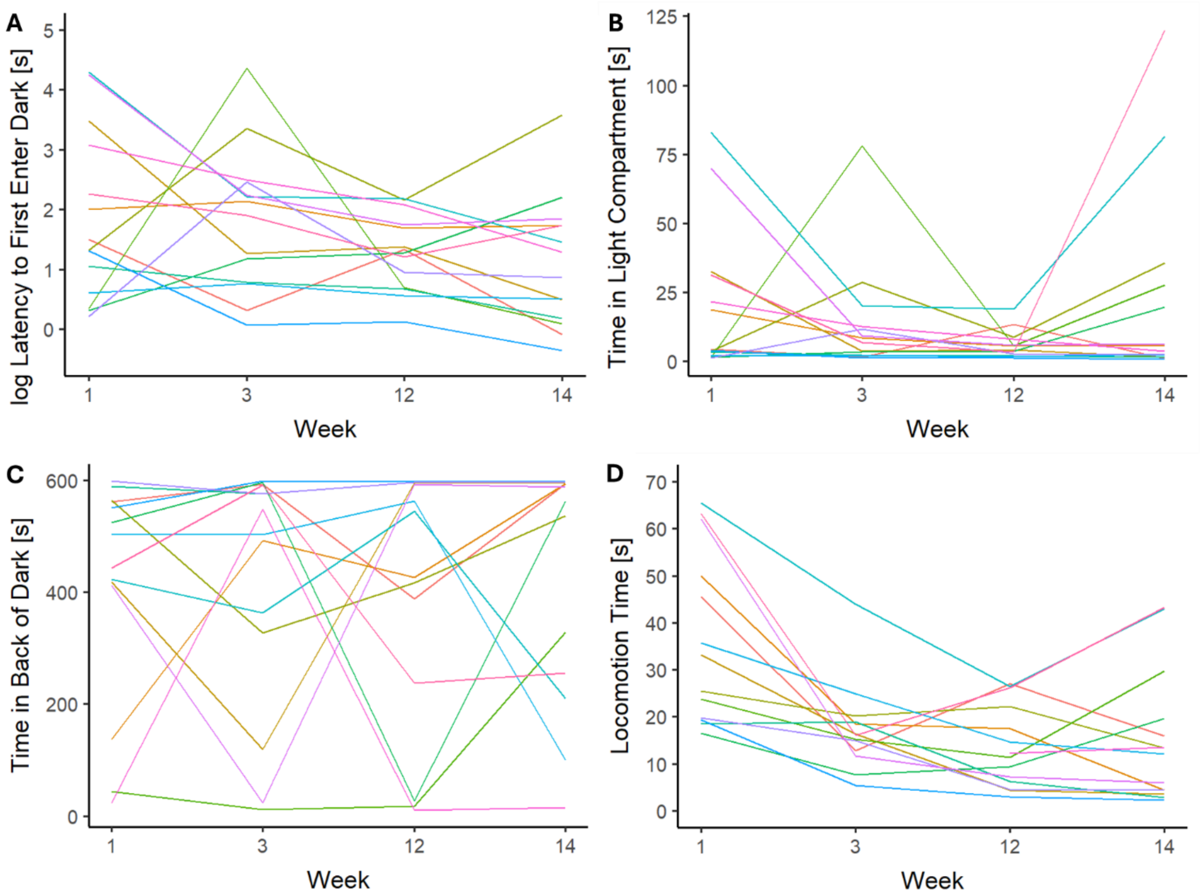

Fig. 2: Behavioural responses of the animals in the light-dark test. (A) Latency to first enter the dark compartment. (B) Total time spent in the light compartment. (C) Time spent in the back half of the dark compartment. (D) Time spent in locomotion. The y-axes show

the total time a behaviour was shown separated by animals and differences over time on the x-axis.

3.2 Sociality test

The time the animals spent in the empty area of the sociality test (Fig. 3A) changed significantly from week to week ($F(3, 39) = 6.696$, $p < 0.001$), dropping off significantly after week 1 ($p = 0.042$) and not changing significantly afterwards ($p = 1$). The time spent in the empty area was repeatable over time ($R = 0.41$, $p < 0.001$, CI = [0.082, 0.645]) and over the short intervals (weeks 1-3: $R = 0.533$, $p = 0.015$, CI = [0.044, 0.834]; weeks 12-14: $R = 0.828$, $p < 0.001$, CI = [0.551, 0.94]), but no repeatability was found over the longer time interval (weeks 3-12: $R = 0.23$, $p = 0.192$, CI = [0, 0.657]). Note that the time the animal spent in the empty area, as analysed here, represents reciprocal results for the time spent in the social area.

The time the animals spent interacting with the conspecific in the wire mesh cage (Fig. 3B) changed significantly over the weeks ($F(3, 39) = 3.9132$, $p = 0.0155$), increasing from week 1 to week 12 ($p = 0.048$; weeks 1-14: $p = 0.021$, all other comparisons $p > 0.529$), and no repeatability was found in any timeframe (overall: $R = 0.141$, $p = 0.115$, CI = [0, 0.406]; weeks 1-3: $R = 0.015$, $p = 0.478$, CI = [0, 0.538]; weeks 3-12: $R = 0.004$, $p = 0.494$, CI = [0, 0.511]; weeks 12-14: $R = 0.151$, $p = 0.285$, CI = [0, 0.579]).

The time of locomotion in the sociality test (Fig. 3C) remained stable over time ($F(3, 39) = 1.051$, $p = 0.381$). The data showed repeatability over the 14 weeks ($R = 0.265$, $p = 0.015$, CI = [0, 0.53]), as well as increasing repeatability estimations over time, leading to significant repeatability in comparison of weeks 3 and 12 as well as weeks 12 and 14 (weeks 1-3: $R = 0$, $p = 1$, CI = [0, 0.504]; weeks 3-12: $R = 0.567$, $p = 0.01$, CI = [0.087, 0.843]; weeks 12-14: $R = 0.842$, $p < 0.001$, CI = [0.584, 0.943]).

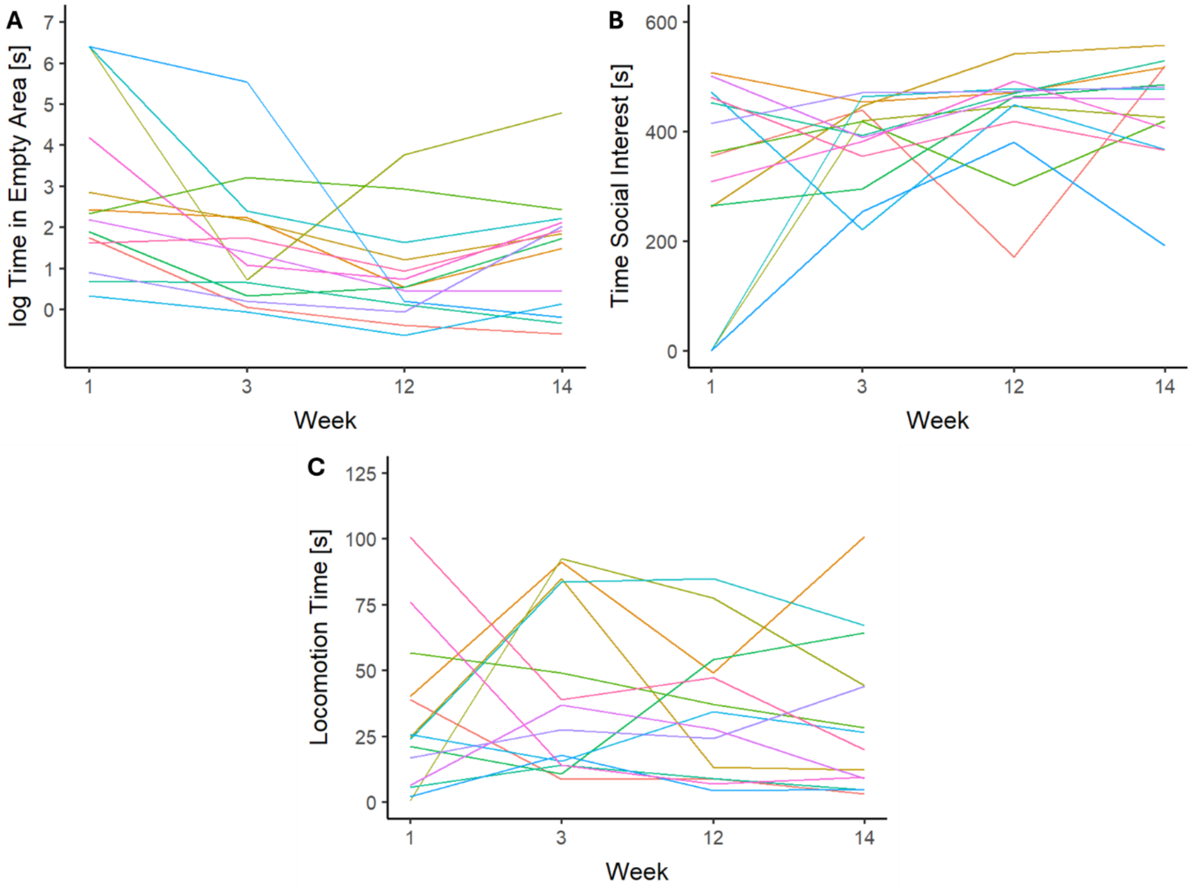

Fig. 3: Behavioural responses of the animals in the sociality test. (A) Time spent in the empty area. (B) Time of social interest in the female. (C) Time spent in locomotion. The y-axes show the total time a behaviour was shown separated by animals and differences over time on the x-axis.

<u>3.3 Open-field test</u>

The time the animals spent in the centre of the open-field (Fig. 4A) did not show a significant effect of week ($F(3, 39) = 0.671$, $p = 0.575$). Significant repeatability was found overall ($R = 0.55$, $p < 0.001$, CI = [0.245, 0.746]), and in the short-term intervals (weeks 1-3: $R = 0.678$, $p = 0.002$, CI = [0.25, 0.881]; weeks 12-14: $R = 0.46$, $p = 0.034$, CI = [0, 0.776]), while the long-term interval revealed a tendency in repeatability (weeks 3-12: $R = 0.384$, $p = 0.068$, CI = [0, 0.737]).

The animals entered the centre of the arena an average of 1.098 times per experiment (Fig. 4B), showing no changes over the weeks ($F(3, 39) = 0.695$, $p = 0.561$). Entering the centre showed repeatability over time, though a tendency was found for the second short-term interval from week 12 to 14 (overall: $R = 0.694$, $p < 0.001$, CI = [0.184, 0.907]; weeks

1-3: R = 0.665, p = 0.012, CI = [0, 0.882]; weeks 3-12: R = 0.575, p = 0.011, CI = [0, 0.829]; weeks 12-14: R = 0.496, P = 0.051, CI = [0, 0.829]).

The time the animals spent in locomotion in the open-field test (Fig. 4C) stayed consistent over time ($F(3, 39) = 1.7$, $p = 0.183$) and showed repeatability overall ($R = 0.413$, $p < 0.001$, CI = [0.106, 0.671]). When comparing the separate intervals, the repeatability estimate was higher for the short-term intervals than the long-term interval, though only showing significant repeatability from week 12 to 14 (weeks 1-3: R = 0.283, p = 0.14, CI = [0, 0.7]; weeks 3-12: R = 0.069, p = 0.397, CI = [0, 0.563]; weeks 12-14: R = 0.521, p= 0.018, CI = [0, 0.805]).

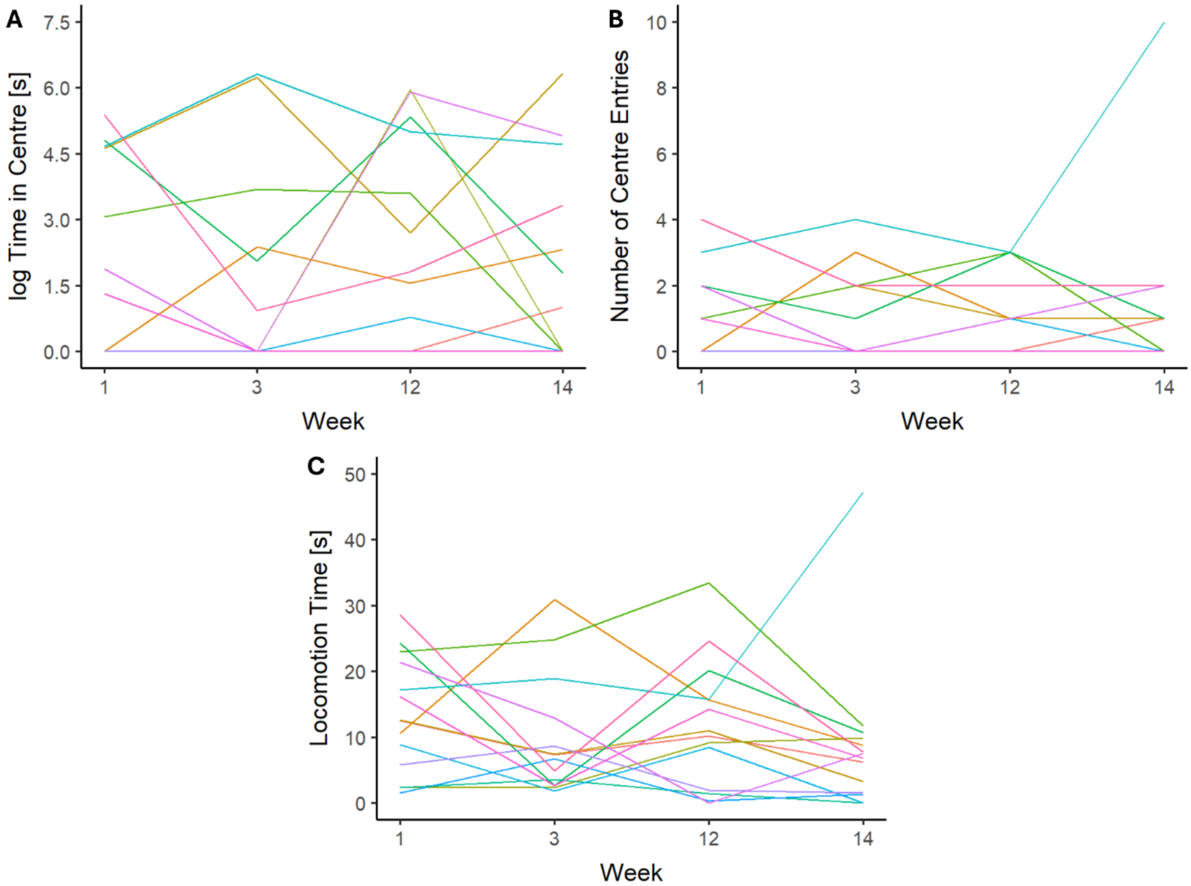

Fig. 4: Behavioural responses of the animals in the open-field test. (A) Time spent in the centre of the arena. (B) Number of centre entries. (C) Time spent in locomotion. The y-axes show the total time or total number of times a behaviour was shown separated by animals and differences over time on the x-axis. (A, B) Note that curves representing several individuals are superimposed on 0 throughout.

3.4 Body mass

The body mass of the animals (Fig. 5A) did not change significantly over the study period ($F(3, 39) = 2.093$, $p = 0.117$), and was highly repeatable throughout (overall: $R = 0.877$, $p < 0.001$, CI = [0.691, 0.942]; weeks 1-3: $R = 0.982$, $p < 0.001$, CI = [0.944, 0.994]; weeks 3-12: $R = 0.816$, $p < 0.001$, CI = [0.544, 0.941]; weeks 12-14: $R = 0.888$, $p < 0.001$, CI = [0.696, 0.966]).

3.5 Cortisol concentrations

Cortisol concentrations (Fig. 5B) changed over time ($\chi^2(3) = 16.043$, $p = 0.001$), reaching low levels in week 3, and increasing afterwards ($p = 0.003$). The cortisol concentrations showed no repeatability in any time frame, except for a tendency from week 12 to 14 (overall: $R = 0.066$, $p = 0.286$, CI = [0, 0.343]; weeks 1-3: $R = 0$, $p = 0.5$, CI = [0, 0.549]; weeks 3-12: $R = 0$, $p = 1$, CI = [0, 0.575]; weeks 12-14: $R = 0.353$, $p = 0.09$, CI = [0, 0.746]).

3.6 Social dominance

The Coulon-Index measuring the social dominance hierarchy (Fig. 5C) did not change significantly from the first test block to the second ($F(1, 13) = 0.014$, $p = 0.907$). The animal's hierarchical position stayed repeatable within the measured time frame ($R = 0.6$, $p = 0.006$, CI = [0.194, 0.843]).

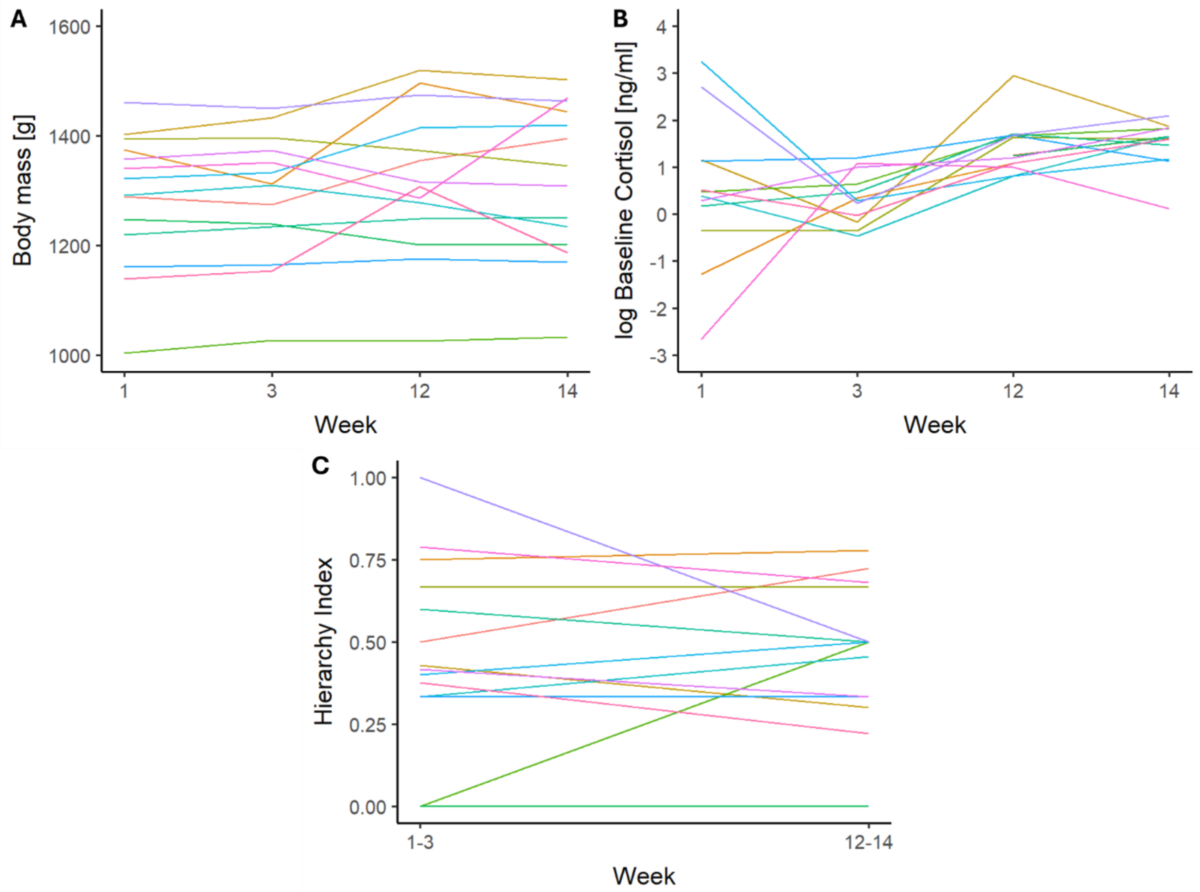

Fig. 5: General (body) conditions. (A) Cortisol concentration. (B) Body mass. (C) Coulon (hierarchy) index.

3.7 Meta analysis

The R-values (repeatability estimations) found in this study (Fig. 6) showed significant differences between the investigated time intervals ($F(2, 63) = 5.088$, $p = 0.009$). Repeatability was significantly higher for the second short-term interval compared to the first short-term and the long-term interval (Short1-Short2: $p = 0.024$; Long-Short2: $p = 0.021$), while no difference was found between the repeatability scores of the first short-term and the long-term interval ($p = 1$).

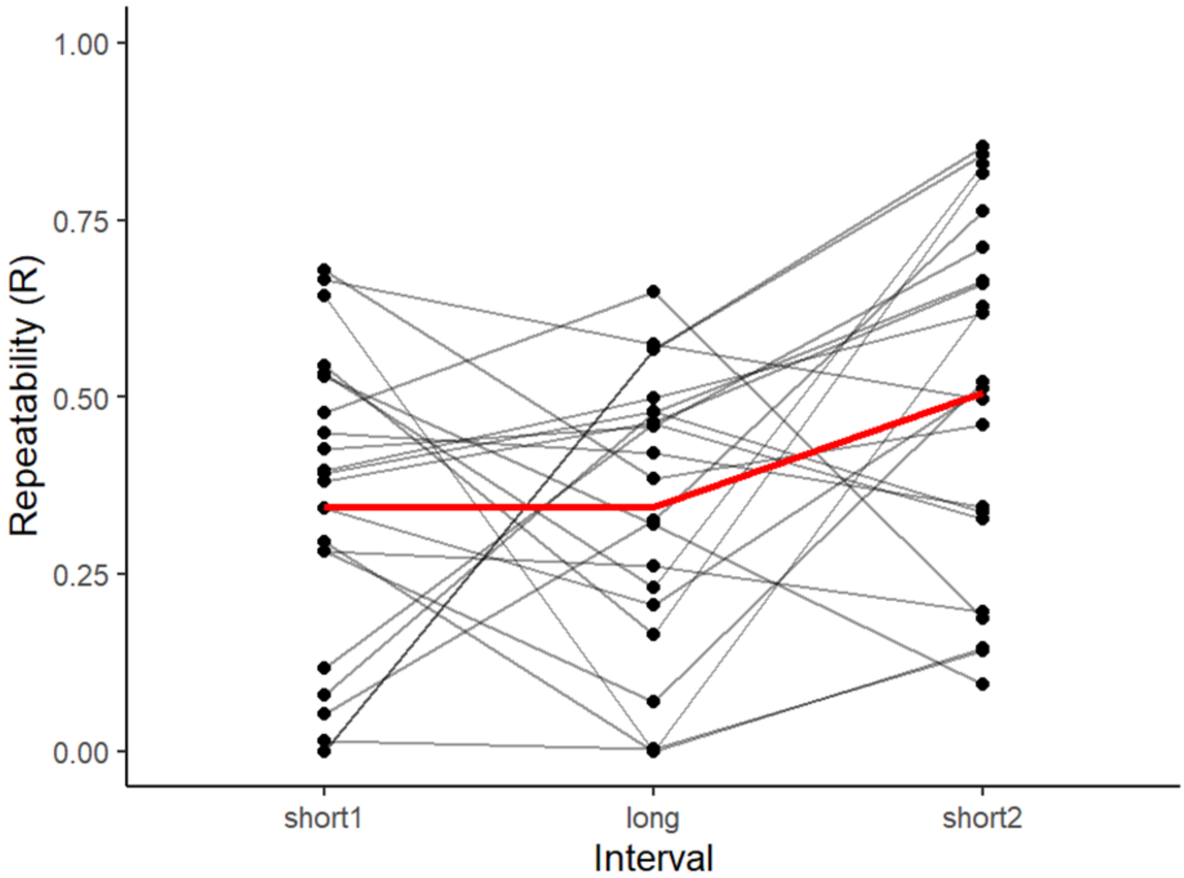

Fig 6: Comparison of repeatability estimations of the measured behaviours between the first short-term interval (weeks 1-3), the long-term interval (weeks 3-12) and the second short-term interval (weeks 12-14) based on the respective R-values

5. Discussion

Guinea pigs are commonly used in behavioural physiological research, but knowledge on individual behavioural stability over time is surprisingly scarce. This study investigated short- and long-term repeatability of behavioural responses in male guinea pigs to gain insight into the degree of personality and/or stability of behavioural phenotypes. Moreover, this should test the reliability of behavioural tests to indicate stable behavioural differences. In the present study, 90% of behaviours measured were repeatable overall, 30% were repeatable in the first short-term interval, 50% in the long-term interval, and 70% in the second short-term interval. Despite the fact that most behaviours were repeatable overall, repeatability estimations increased with time and with the number of tests, respectively.

Bell *et al.* (2009) described in their meta-analysis of behavioural repeatability that it is expected to drop with increasing inter-test intervals. Possible explanations for this include different genetic expression patterns throughout different life stages, environmental effects or changes in the animal's current state, which might refer to its hunger, size, age, health condition, or dominance (Bell *et al.* 2009). The environmental conditions did not change over the course of this study, and as the male guinea pigs were all adult when beginning the tests, the explanations based on changes in genetic expression over life stages are also less likely in the framework of this study. As male guinea pigs reach ages of up to 5 years in our lab, an age of 18 months at the beginning of the study could not be considered "old" either, and no considerable aging effect might have come to bear. Furthermore, hunger cannot differentiate between short- or long-term repeatability, as the availability of food did not change over time and the timing of eating is shifting moment to moment, instead of a longer-term change. This timing of food intake (and the energy availability derived from that) may lead to changing behaviour from trial to trial (Mathot & Dingemanse 2015), possibly also based on changing cortisol levels (Nemeth *et al.* 2014) but is not tied to the inter-test interval. The factors of general body conditions and social dominance are more relevant when looking specifically at the differences between short- and long-term repeatability in this study. As the body mass did not change significantly over time it might seem obvious to also disregard this parameter, but as individual masses shifted around relatively stable means, it is not possible to say if or how these changes in weight might have been related to individual food intake and/or internal energy reserves that might have affected the animals' stress loads or behaviour. Cortisol values fluctuated over time showing no repeatability. This result matches with what was previously found (Zipser *et al.* 2013) and what was outlined in the introduction. As cortisol levels stayed instable though, behaviour became more repeatable over time. This leads to the conclusion that there was no direct relationship between the two. On the other hand, the extreme individual changes in cortisol levels, especially between weeks 1 and 3, might correlate with the large variation in behaviour in that timeframe. A direct correlation of cortisol stability and behavioural repeatability would have to be analysed within a separate study aimed at that specifically. As individual hierarchy indices stayed repeatable within the study period, meaning the animals' rank didn't change, changing dominance structures can be excluded from being the reason for changing behavioural repeatability. The two individuals for whom the Coulon-Index changed most drastically were pair-housed individuals who showed barely any interactions in either of the timeframes, meaning that no clear dominance relationships could be determined in pair-housed males at all. Nevertheless, also group-housed males showed very low frequency of agonistic interactions. This general low number of interactions is a sign of stable hierarchies (Tibbetts *et al.* 2022).

The results of the present study, finding differences between the long-term and the second short-term, and none between the first short-term and the long-term, allow for two possible interpretations. Either the repeatability increases over time, regardless of interval, or the animals needed the first trial(s) to adjust to the setups and experiments, then leading to increased repeatability after habituation. Whenever repeatability is measured and behavioural tests are repeated, sensitization or habituation are factors to be taken into account (Rudeck *et al.* 2020). Habituation, as a process of adapting to a situation previously experienced, affects the outcome of repeated measures (van Oers *et al.* 2005). The conditions of capturing and handling prior to the test might also affect the behaviour of an animal within a test (Martin & Réale 2008), which could have played a role in the present study. This increase in stress might affect the behaviour displayed within the tests (Walker *et al.* 2017, Toates 2000) but will probably also be something the animals get used to over time. The behaviours displayed in this study in week 1 deviate from the results of the other weeks. This is supported by the fact that the behaviour in weeks 12 and 14 revealed higher repeatability than in week 1 and 3. Another long-term interval after the second short-term interval or trial tests for habituation would have helped to differentiate between the hypothesis of habituation and lower repeatability for longer inter-test intervals after that, and the hypothesis of increasing repeatability with higher number of tests. According to Bell *et al.* (2009) though, no evidence exists that repeatability estimates are affected by the number of observations per individual.

High repeatability was found in the present study in risk-taking and anxiety-like behaviour, as indicated by repeatability of behaviours determined in the light-dark and open-field test. When measuring personality using similar test setups, Zipser *et al.* (2013) did not find repeatability. Possible reasons for this might include the necessity for a habituation phase, or that the 8-week inter-test intervals used in that study were too long to determine repeatability in guinea pigs. Based on the results in the present study, personality in the emotional domain of guinea pigs is likely. To confirm this, more analyses, for example comparing behaviours across contexts, would be necessary. Contrary to the results of this study, a study by Žampachová *et al.* (2017) on rats found high repeatability in behaviours relating to exploration and lower repeatability in boldness-related behaviours. In a different study (Hakataya *et al.* 2023), shyness-boldness was discovered as a personality trait in laboratory rats, though no consistency over time was measured. There are still only few studies looking into repeatability of behaviour and personality in guinea pigs, and further research would be necessary to make an informed conclusion. Nevertheless, the high repeatability found within the present study allows us to confirm the light-dark test, and the open-field test as adequate measures for these behavioural traits, showing stable and robust results.

Measured variables referring to the animals' activity showed repeatability in the light-dark and the open-field test, only marginally so in the sociality test. This strong difference in repeatability measures of the same behaviour between different tests means that no stability across measures could be found. According to the definition, activity could not be determined as a personality trait in guinea pigs within the present study. This also indicates that especially the sociality test is not adequate to measure general activity and that different tests should also not be used interchangeably. Measuring repeatability in activity in one behavioural test does not transfer or apply to the others, as they are not equivalent and measure different aspects of behaviour (Rudolfová *et al.* 2022). Hakataya *et al.* (2023) did discover activity as a personality trait in laboratory rats, though no consistency over time was measured. The general activity was probably additionally skewed by factors outlined above, as food intake and cortisol, leading to less stable individual differences.

Repeatability of behaviours in the social domain however, which is highly supported in previous literature (Zipser *et al.* 2013, Brust & Guenther 2017), was not found in the present study. The duration of social interaction increased slightly over time, but no repeatability was observed. In the sociality test, the animals were presented with multiple stimuli that are not considered as part of a social context. In addition to the social and sex-driven stimuli produced by the female sitting in the cage, the wire mesh cage itself was a novel object to the animals, which might have been of interest, though guinea pigs are said to be neophobic (Kleven 2021). In addition to that, the cage made the "social" side of the arena narrower, which made the space feel more walled off, as guinea pigs prefer staying close to the cage walls (White *et al.* 1989). These findings were supported by trial runs comparing the empty arena with the arena and the wire mesh cage without a conspecific sitting within the cage showing the animal's interest in the cage regardless of the conspecific (personal observations, Wiehler & Koller). Since this study focuses on repeatability of behaviour, it is not relevant which of these stimuli triggered the animal to behave the way it did, as long as the stimuli remained consistent over time. It is, however, important to note, that it is likely that the sociality test presented here probably did not measure the actual social interest.

As the animals used in this study were kept in different housing conditions (in pairs or small groups), this might be considered to lead to differences in behaviour. Mutwill *et al.* (2023) showed differences in cortisol responsiveness and baseline cortisol (though effect of this decreased with age) between housing conditions. Since cortisol levels might affect an animal's behaviour (Walker *et al.* 2017, Toates 2000), the housing conditions might have played a role in this study. Since group sizes were comparably low (meaning no big difference between pair- and group-housing conditions) and sample size in this study would not have been robust after splitting into subgroups, housing conditions were not tested separately.

In conclusion, the study shows a difference in behavioural repeatability between short, and long inter-test intervals. A follow-up study controlling for habituation could support this claim. The light-dark test is supported as a measure of risk-taking and anxiety-like behaviour, with animals showing high repeatability in behaviours linked to those traits. The open-field test, as a very established measure of both activity and risk-taking and anxiety-like behaviour, is backed by the results shown in this study. The sociality test as performed here should be reassessed and repeated in a way that more clearly isolates the animals' social interest and behaviour. Overall, the performed tests repeatably and reliably show clear individual differences in behavioural traits.

6. References

D. Bates, M. Maechler, B. Bolker, S. Walker, (2015), Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48

A.M. Bell, S.J. Hankison, K. L. Laskowski, (2009), The repeatability of behaviour: A meta-analysis. Animal Behaviour, 77(4), 771–783

R. Bergmüller, (2010), Animal personality and behavioural syndromes. In: P. Kappeler, Animal Behaviour: Evolution and Mechanisms. Springer, Berlin, Heidelberg

P.A. Biro, (2012), Do rapid assays predict repeatability in labile (behavioural) traits? Animal Behaviour, 83(5), 1295-1300

M. Briffa, A. Weiss, (2010), Animal Personality. Current Biology, 20(21), 912-914

M.E. Brooks, K. Kristensen, K.J. van Benthem, A. Magnusson, C.W. Berg, A. Nielsen, H.J. Skaug, M. Maechler, B.M. Bolker, (2017), glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal, 9(2), 378-400

V. Brust, A. Guenther, (2017), Stability of the guinea pigs personality–cognition–linkage over time. Behavioural Processes, 134, 4-11

J. Coulon, (1975), Les relations sociales chez le Cobaye domestique male. I. Etude de la hierarchie sociale (Social Relationships in the Male Domestic Guinea Pig. I. Study of Social Hierarchy). Behaviour, 53(3/4), 183-199

N.J. Dingemanse, C. Both, P.J. Drent, K. Van Oers, A.J. Van Noordwijk, (2002), Repeatability and heritability of exploratory behaviour in great tits from the wild. Animal Behaviour, 64, 929-938

W. Forstmeier, T.R. Birkhead, (2004), Repeatability of mate choice in the zebra finch: consistency within and between females. Animal Behaviour, 68(5), 1017-1028

J. Fox, S. Weisberg, (2019), An R Companion to Applied Regression (3rd Edition). Thousand Oaks, CA

O. Friard, M. Gamba, (2016), BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution, 7(11), 1325-1330

D.R. Garris, (1979), Diurnal fluctuation of plasma cortisol levels in the guinea pig. Acta Endocrinologica, 90(4), 692–695

S. Hakataya, N. Katsu, K. Okanoya, G. Toya, (2023), An exploratory study of behavioral traits and the establishment of social relationships in female laboratory rats. PLOS ONE, 18(12), e0295280

G.A. Kleven, (2021), Behavioral Biology of Guinea Pigs. Behavioral Biology of Laboratory Animals (1st Edition)

M. Leclerc, E. Vander Wal, A. Zedrosser, J.E. Swenson, J. Kindberg, F. Pelletier, (2016), Quantifying consistent individual differences in habitat selection. Oecologia, 180, 697–705

R.V. Lenth, (2025), emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.11.0

J.G.A. Martin, D. Réale, (2008), Temperament, risk assessment and habituation to novelty in eastern chipmunks, Tamias striatus. Animal Behaviour, 75(1), 309-318

M. Masis-Calvo, A.K. Schmidtner, V.E. de Moura Oliveira, C.P. Grossmann, T.R. de Jong, I.D. Neumann, (2018). Animal models of social stress: the dark side of social interactions. Stress, 21(5), 417–432

K.J. Mathot, N.J. Dingemanse, (2015), Energetics and behavior: unrequited needs and new directions. Trends in Ecology & Evolution, 30(4), 199 - 206

M. McGillycuddy, D.I. Warton, G. Popovic, B.M. Bolker, (2025), Parsimoniously Fitting Large Multivariate Random Effects in glmmTMB. Journal of Statistical Software, 112(1), 1-19

A.M. Mutwill, H. Schielzeth, T.D. Zimmermann, S.H. Richter, S. Kaiser, N. Sachser, (2021), Individuality meets plasticity: Endocrine phenotypes across male dominance rank acquisition in guinea pigs living in a complex social environment. Hormones and Behavior, 131, 104967

A.M. Mutwill, H. Schielzeth, S.H. Richter, S. Kaiser, N. Sachser, (2023), Conditional on the social environment? Roots of repeatability in hormone concentrations of male guinea pigs. Hormones and Behavior, 155, 105423

M. Nemeth, E. Millesi, K-H. Wagner, B. Wallner, (2014), Effects of Diets High in Unsaturated Fatty Acids on Socially Induced Stress Responses in Guinea Pigs. PLoS ONE, 9(12), e116292

M. Nemeth, E. Pschernig, B. Wallner, E. Millesi, (2016), Non-invasive cortisol measurements as indicators of physiological stress responses in guinea pigs. PeerJ, 4, e1590

M. Nemeth, I. Eisenschenk, A. Engelmann, F.M. Esser, M. Kokodynska, V.F. Szewczak, E. Barnreiter, B. Wallner, E. Millesi, (2021), Flaxseed oil as omega-3 polyunsaturated fatty acid source modulates cortisol concentrations and social dominance in male and female guinea pigs. Hormones and Behavior, 134, 105025

K. can Oers, M. Klunder, P.J. Drent, (2005), Context dependence of personalities: risk-taking behavior in a social and a nonsocial situation. Behavioral Ecology, 16(4), 716-723

J.C. Pinheiro, D.M. Bates, R Core Team (2025). _nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-167

R Core Team, (2025), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

J. Rudeck,S. Vogl,S. Banneke,G. Schönfelder,L. Lewejohann, (2020), Repeatability analysis improves the reliability of behavioral data. PLOS ONE, 15(4), e0230900

V. Rudolfová, T. Petrásek, E. Antošová, D. Frynta, E. Landová, K. Valeš, T. Nekovářová, (2022), Inter-individual differences in laboratory rats as revealed by three behavioural tasks. Scientific reports, 12, 9361

T.L. Rystrom, R.C. Prawitt, S.H. Richter, N. Sachser, S. Kaiser, (2022), Repeatability of endocrine traits and dominance rank in female guinea pigs. Frontiers in Zoology, 19(1), 4

N. Sachser, (1986), Different Forms of Social Organization At High and Low Population Densities in Guinea Pigs. Behaviour, 97(3–4), 253–272

A.C. Schuster, T. Carl, K. Foerster, (2017), Repeatability and consistency of individual behaviour in juvenile and adult Eurasian harvest mice. The Science of Nature, 104(3), 10

M.J. Smith, D. Hunter, (2005), Temporal and Geographic Variation in the Advertisement Call of the Booroolong Frog (Litoria booroolongensis: Anura: Hylidae). Ethology, 111(12), 1103-1115

H.A.M. Spoolder, J.A. Burbidge, A.B. Lawrence, P. Howard Simmins, S.A. Edwards, (1996), Individual behavioural differences in pigs: intra-and inter-test consistency. Applied Animal Behaviour Science, 49(2), 185-198,

M.A. Stoffel, S. Nakagawa, H. Schielzeth, (2017), rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol Evol, 8(11), 1639-1644

E.A. Tibbetts, J. Pardo-Sanchez, C. Weise, (2022), The establishment and maintenance of dominance hierarchies. Philosophical Transactions of the Royal Society B, 377(1845), 20200450

F. Toates, (2000), Multiple factors controlling behaviour: implications for stress and welfare. In: G.P. Moberg, J.A. Mench, The biology of animal stress: basic principles and implications for animal welfare. CABI Publishing, 199–226

J.E. Wagner, P.J. Manning (Ed.), (1976), The biology of the guinea pig. Elsevier

S.E. Walker, O. Zanoletti, I.G. de Suduiraut, C. Sandi, (2017), Constitutive differences in glucocorticoid responsiveness to stress are related to variation in aggression and anxiety-like behaviors. Psychoneuroendocrinology, 84, 1-10

W.J. White, M.W. Balk, C.M. Lang, (1989), Use of cage space by guinea pigs. Laboratory Animals, 23(3), 208-214

H. Wickham, (2016), ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

R.E. van Wijk, S. Bauer, M. Schaub, (2016), Repeatability of individual migration routes, wintering sites, and timing in a long-distance migrant bird. Ecology and Evolution, 6(24), 8679-8685

B. Žampachová, B. Kaftanová, H. Šimánková, E. Landová, D. Frynta, (2017), Consistent individual differences in standard exploration tasks in the black rat (Rattus rattus). Journal of Comparative Psychology, 131(2), 150-162

B. Zipser, S. Kaiser, N. Sachser, (2013), Dimensions of Animal Personalities in Guinea Pigs. Ethology, 119(11), 970-982

Deutsche Zusammenfassung

Repeatability (Wiederholbarkeit) beschreibt den Grad der individuellen Konsistenz einer Messung (z.B. eines Verhaltens) und ermöglicht es uns, die Verlässlichkeit einer Messung zu quantifizieren. In Bezug auf Verhalten erlaubt dies, stabile Persönlichkeitsmerkmale zu finden. Die Flexibilität von Verhalten im Verlauf des Lebens eines Tieres im Anbetracht sich ändernder Lebensbedingungen ermöglicht Anpassung und hat einen Einfluss auf das Überleben des Individuums. Es wurde gezeigt, dass Wiederholbarkeit von Verhalten in längeren Intervallen zwischen Tests abnimmt. In der vorliegenden Arbeit wurden die Verhaltensreaktionen von Meerschweinchen in drei Verhaltenstests (light-dark test, sociality test, open-field test) über kürzere (zwei Wochen) und längere (8 Wochen) Zeiträume verglichen, um die Repeatability verschiedener Aspekte des Meerschweinchenverhaltens festzustellen. Dies gibt Einblicke in die Stabilität von Verhalten, Persönlichkeit von Meerschweinchen, und erlaubt es die verwendeten Tests als Methode zum Messen bestimmter Verhaltensweisen zu bestätigen. Verhaltens-weisen, die im Zusammenhang mit Mut und Risikobereitschaft standen, zeigten hohe Repeatability, ebenso Verhaltensweisen, die auf hohe Aktivität hinweisen. Die gemessene Repeatability nahm im Laufe der Zeit zu und zeigte geringere Repeatability für das erste kurze Intervall so wie das lange Intervall, mit hoher Repeatability im zweiten kurzen Intervall. Sozialverhalten zeigte sehr geringe Repeatability, sollte jedoch im Kontext des Experiments diskutiert werden. Die durchgeführten Tests zeigten insgesamt über Zeit wiederholbare und stabile Verhaltensreaktionen.