



Exploration behaviour in a different light: testing cross-context consistency of a common personality trait



Lisa K. Arvidsson^{*}, Frank Adriaensen, Stefan van Dongen, Nils De Stobbeleere, Erik Matthysen

Evolutionary Ecology Group, Universiteit Antwerpen, Antwerpen, Belgium

ARTICLE INFO

Article history:

Received 9 March 2016

Initial acceptance 25 April 2016

Final acceptance 5 August 2016

Available online 29 November 2016

MS. number: 16-00208R

Keywords:

activity
animal personality
context validation
cross-context consistency
exploration behaviour
novel environment

Laboratory measures of personality traits are often used to answer both evolutionary and ecological questions involving behavioural variations in the wild within and between populations. However, little is known about the actual behavioural mechanisms behind any correlation with behaviours in the wild, and traits often lack validation in a different context. We examined whether the commonly used exploration behaviour trait constitutes an active exploration strategy, by testing whether the activity in the exploration test could also be captured in a different exploration context. We subjected great tits, *Parus major*, to two different tests, one being the standardized exploration behaviour test and one a newly constructed test. The new test arena contained eight large rooms connected by corridors, where we scored the activity of individuals and the number of rooms visited as a proxy for exploration. We found that our new exploration test captured repeatable behaviour in activity and exploration of rooms both within and across years. We found no correlations between the two tests, suggesting that they may not capture the same behaviour, in terms of exploration of rooms or activity. We conclude that in our study population, the classic exploration behaviour test seems context specific, rather than constituting a general exploration strategy.

© 2016 The Authors. Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Consistent between-individual variation in behaviour, most commonly defined as animal personality (Gosling, 2001; Sih, Bell, & Johnson, 2004) or temperament (Gosling, 2001), is a widely studied topic from both an evolutionary and ecological perspective (Sih, Cote, Evans, Fogarty, & Pruitt, 2012; Wolf & Weissing, 2012). Animal personality is normally considered a set of several underlying personality traits such as aggression (Araya-Ajoy & Dingemans, 2014; Huntingford, 1976), activity (Gosling, 2001) and exploration (Verbeek, Drent, & Wiepkema, 1994). These traits are typically quantified by their expression under standardized laboratory conditions and are commonly related to a number of behaviours and life history traits in the wild (Reale, Dingemans, Kazem, & Wright, 2010). Findings of such relationships support the idea that animal personality plays a central role in population dynamics and micro-evolution (Clobert, Baguette, Benton, & Bullock, 2012; Dochtermann & Dingemans, 2013; Wolf & Weissing, 2012).

Currently, there is a gap of knowledge in how behaviours quantified in an artificial laboratory setting, often in small test environments with minimal complexity such as test cages and test rooms, translate to behaviour in larger and more complex environments as an intermediate step to the wild. There are several important reasons to study this question. One is that selection ultimately acts on the expression of personality traits in the wild, and not on behaviour measured in a laboratory setting. Another equally important reason is that there is little understanding of what behaviours are actually measured in captivity. The importance of validating the meaning of behavioural traits has been emphasized recently (Carter, Feeney, Marshall, Cowlshaw, & Heinsohn, 2013; Niemela & Dingemans, 2014). This is particularly important in scenarios in which a behavioural trait measured in a single captive context is used to draw conclusions about different behaviours in the wild, or when an empirical correlation is found between a laboratory trait and a presumed different behaviour in the wild. The difficulty in understanding a personality trait's functional meaning makes cross-context validations of behavioural measures necessary, which has been referred to as convergent validity (Carter et al., 2013). One way to validate a laboratory measure is to correlate measures of behaviour that are presumed to measure the same

^{*} Correspondence: L. K. Arvidsson, Groenenborgercampus, G.V. 330, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium.

E-mail address: Elisabeth.Arvidsson@uantwerpen.be (L. K. Arvidsson).

trait, or closely linked traits, in captivity and the wild. In great tits, *Parus major*, for example, Morand-Ferron and Quinn (2011) found that individual problem-solving capacity in captivity was correlated with problem solving at a feeder study during winter. Herborn et al. (2010) assayed both exploration behaviour and neophobia in blue tits, *Cyanistes caeruleus*, and found correlations between these two traits in captivity and their analogues in the wild. Many laboratory-measured traits still need cross-context validation, however, and more work is needed to understand how behaviours expressed in standardized environments correlate with behaviour in other environments (Niemela & Dingemanse, 2014).

A common measure of personality used in a wide range of animal species is the rate at which an individual moves through a novel space, normally referred to as 'exploratory behaviour' or 'exploration' (Dingemanse, Both, Drent, Van Oers, & Van Noordwijk, 2002; Reale, Reader, Sol, McDougall, & Dingemanse, 2007; Verbeek et al., 1994). It is often measured in a single non-subdivided space with few structures and is assumed to reflect an underlying 'exploration trait' which can only be expressed in a novel environment (Reale et al., 2007). Correlations have been found between individual variation in exploration behaviour and other behaviours including foraging routines (Verbeek et al., 1994), risk-taking behaviour (van Oers, Drent, de Goede, & van Noordwijk, 2004), antipredator behaviour (Jones & Godin, 2010) and aggression towards conspecifics (Carere, Drent, Privitera, Koolhaas, & Groothuis, 2005; Wilson et al., 2009). Exploration behaviour has been shown to be heritable within several taxa such as fish (Dingemanse et al., 2009), mammals (Careau et al., 2011; Kanda, Louon, & Straley, 2012) and birds (Dingemanse et al., 2002; Korsten et al., 2010; Quinn, Patrick, Bouwhuis, Wilkin, & Sheldon, 2009). As it is measured as a response to a novel environment, exploration behaviour has often been predicted to play a role in spatial movement in the wild (Cote, Clobert, Brodin, Fogarty, & Sih, 2010), and multiple studies have found correlations between exploration behaviour and expressions of spatial behaviour such as dispersal distance (Dingemanse, Both, van Noordwijk, Rutten, & Drent, 2003; Fraser, Gilliam, Daley, Le, & Skalshi, 2001; Hoset et al., 2010; but see Cote, Fogarty, Weinersmith, Brodin, & Sih, 2010), seasonal dispersal movements (Chapman et al., 2011; Thijs van Overveld, Careau, Adriaensen, & Matthysen, 2014), home range sizes (Minderman, Reid, Evans, & Whittingham, 2009; van Overveld, Adriaensen, & Matthysen, 2011) and family movements (van Overveld et al., 2011). A recent study found evidence for a genetic integration of heritable variation in both exploration behaviour and dispersal distance (Korsten, van Overveld, Adriaensen, & Matthysen, 2013). Exploration behaviour thus reflects a personality trait that can be a target of selection associated with spatial behaviours, and is therefore an informative trait for examining individual and population level processes in behavioural and movement ecology (Cote, Clobert, et al., 2010; Sih et al., 2012).

Despite the accumulating evidence that exploration behaviour correlates with spatial behaviours, the way in which variation in exploration behaviour translates into these behaviours is still not clear. An obvious candidate behavioural mechanism is that individual variation in exploration behaviour directly translates into larger-scale spatial movements in the wild through variation in exploration strategies. Differences in the intensity, willingness or thoroughness of exploration may result in significant between-individual variation in spatial displacement over time. If exploration behaviour indeed reflects a general strategy of exploring unknown environments, in the laboratory as well as in the wild, then we would expect variation in exploration behaviour to be highly consistent across contexts that vary strongly in spatial scale and complexity. In this study we tested this hypothesis by measuring individual exploration behaviour in two different laboratory based

contexts: a standardized exploration behaviour test that is widely used in our focal species by several different research groups and a newly constructed test involving a larger space and more complex environment. The alternative hypothesis is that the link between spatial behaviours and exploration behaviour is caused by covariation between exploration behaviour and other behaviours that do not directly involve exploration, such as aggressiveness and dominance (Adriaenssens & Johnsson, 2010b; Dingemanse, 2004) or differences in use of social information (Aplin, Farine, Mann, & Sheldon, 2014), which we did not address in this study. Additionally, underlying differences in physiology such as metabolic rate could also cause differences in spatial movements through differences in general activity levels (Biro & Stamps, 2008).

Few previous studies have tried to relate exploration behaviour measures across contexts. A study in crickets (Dochtermann & Nelson, 2014) found no evidence for cross-context consistency. Herborn et al. (2010) tested relationships between discovery of new feeders in the wild and activity in a captive novel environment, but found no relationship. When they instead used a residual measure for exploration behaviour by subtracting activity in a known environment from the activity in the novel environment, they found a significant relationship with exploration behaviour in the wild. It has already been noted that individual variation in activity might interfere with the measure of exploration (Reale et al., 2007). Therefore we aimed at quantifying different behaviours that might reflect either activity or exploration behaviour expressed in activity and additionally exploration behaviour in terms of forward movement. We used two different tests for this: the standardized exploration behaviour test which quantifies the activity in a novel but limited environment, where the measurement is referred to as the exploration behaviour score, and a newly constructed test which quantifies both the activity and forward movement in an equally novel, but more complex environment.

We used the great tit as a model species. In this species, variation in exploration behaviour in a small novel room has been studied in great detail and has been replicated in many populations since the early 1990s (Dingemanse et al., 2002; Hollander, Van Overveld, Tokka, & Matthysen, 2008; Quinn et al., 2009; Verbeek et al., 1994). Several studies have confirmed the repeatability of the exploration behaviour score, as well as showing similar responses to test conditions such as test sequence or time elapsed between tests (Dingemanse et al., 2012; Korsten et al., 2013; Quinn et al., 2009).

We designed a new laboratory based exploration test, named the arena test, in which we challenged great tits to explore a more complex novel environment composed of several connected rooms. In addition to quantifying activity within a room, the arena also gives the option of forward movement through the set of rooms as opposed to the exploration behaviour test. By this design we hoped to be able to differentiate between two behavioural responses that may or may not reflect variation from the same assumed latent 'exploration trait': activity as measured by number of movements and forward movement as the number of rooms visited. We refer to these proxies as 'Arena Activity' (total number of movements within all rooms) and 'Rooms Visited' (forward exploration of rooms).

In this paper, we use 'exploration behaviour' as an overall term for behaviours observed in novel room tests such as the exploration behaviour test. The quantified behaviour from the standardized exploration test is referred to as the exploration behaviour score. From the newly constructed exploration test, the arena, we quantified two variables: Arena Activity and Rooms Visited. Arena Activity was measured similarly to the exploration behaviour score with total number of movements within the tests, whereas Rooms Visited was the number of rooms a bird visited, only possible through forward movement.

Our main aims were to test whether (1) the measurements we collected from the arena test captured repeatable individual variation in behaviour and (2) the variation in behaviour expressed in the exploration behaviour test correlated positively with either or both of the behaviours measured in the arena.

METHODS

Study Area and Bird Handling

The study area consists of a set of 12 forest fragments scattered across a countryside area in the north of Belgium (51°80'80"N, 4°83'20"E), where every forest fragment contains a set of nestboxes (for a detailed description of the study area see [Matthysen, 2002](#)). The great tit population is closely monitored through inspection of the nestboxes during the breeding season every year, as well during regular roost checks twice a winter. During the breeding seasons of 2012–2014 and all capture events within these years all full-grown birds (including 15-day-old chicks) were ringed and equipped with an additional ring containing an individual passive integrated transponder tag (PIT tag) weighing 0.2 g corresponding to at most 1.3% of the body weight of the smallest individuals. Scanning PIT tags through nestboxes allowed a quick selection of individuals during roost checks in the field for behavioural testing, with minimal disturbance of nonselected individuals. Birds were collected and transported in fabric bird bags. Individuals were brought to the laboratory for one of the two behavioural tests and released immediately afterwards. Thus, no individuals were subjected to the two tests on the same day. Individuals were selected in such a way as to maximize the number of birds subjected to both tests, regardless of the year of testing. Thus, individuals could be subjected to both tests in different years, or in the same year, with a minimum of 2 weeks between respective tests. Whenever possible, individuals were subjected to both the arena and exploration behaviour test in more than 1 year. Since more individuals could be tested per day in the exploration behaviour test than in the arena test, birds caught for the first time were subjected preferentially to the exploration behaviour test, and to the arena at a later capture date. This enabled us to build up a larger sample size of individuals with both tests. Individuals were caught by mist netting in the afternoon (1400–1800 hours) or roost checks (1700–0000 hours) during winter between October and February. Morphological measures during capture and/or the morning after testing included weight (upon capture, upon arriving at the laboratory and immediately after testing), tarsus length, fat score and sex and age based on plumage coloration. All birds were housed individually in cages overnight before being tested and provided with ad libitum food (peanuts, mealworms and fat ball) and water. Testing was initiated at least 1 h after sunrise (arena test) or after 1 h of artificial light (exploration behaviour test), and conducted between 0830 and 1330 hours. Artificial light was regulated according to the natural light cycle. All birds were released near the site of capture before 1400 hours the day following capture and kept in captivity for a maximum of 20 h. Both the arena test and the exploration behaviour test were conducted over 4 years, yielding respectively 197 tests on 125 individuals (arena) and 405 tests on 313 individuals (exploration behaviour room) with total repeats within each test ranging from zero to four.

Exploration Behaviour Test

The exploration behaviour test was based on a study by [Verbeek et al. \(1994\)](#) and was performed following a highly standardized procedure described in detail in [Dingemanse et al. \(2002\)](#). The test refers to spatial exploration and does not include object

exploration, but in line with previous literature we simply refer to this test as exploration behaviour ([Arvidsson & Matthysen, 2016](#); [Dingemanse et al., 2012](#)). Overnight cages in the exploration behaviour room measured 0.8 × 0.4 m and 0.5 m high. The test room was 4 × 2.4 m and 2.3 m high and contained five artificial trees 1.5 m high each with four branches 20 cm long ([Fig. 1](#)). We introduced each bird into the room without handling by opening the sliding door from the inside, turning on the light in the test room and if necessary approaching the cage, after which all birds flew into the room. Scoring started immediately. All hops within trees and movements between trees were counted as well as landings on the floor, ceiling or walls. Movements within a branch were not counted. All movements were added up to an exploration behaviour score over 2 min of testing. Because exploration behaviour scores have previously been shown to be repeatable ($N = 224$, $r = 0.42$, [Dingemanse et al. 2012](#)) and heritable ($h = 0.30 \pm 0.11$, [Korsten et al. 2013](#)) in our study population, predominantly only one measure per individual was taken each winter resulting in 405 tests on 313 individuals in total (tests: 2012: $N = 114$; 2013: $N = 123$; 2014: $N = 141$; 2015: $N = 27$). The number of individuals tested in two different winters was 21 from 2012–2013, 18 from 2013–2014, 22 from 2014–2015 and five from 2012–2014. Two individuals were excluded from subsequent analyses because of human error during testing or visible illness of the bird.

Arena Construction

We constructed a 5 × 20 m and 2.4 m high 'arena' out of oriented strand board panels. We deliberately chose different materials and appearance from the exploration behaviour test room to minimize habituation effects to the general test environment. The arena contained eight connected rooms of 2.5 × 5 m each. Each room contained a narrow corridor of 0.7 × 3.6 m leading to the next room ([Fig. 2](#)). The corridors were at opposite ends such that movement through the full length of a room was required to reach the next corridor. Each room contained three (entry room) or six (all other rooms) perches 0.3 m long arranged perpendicular to the walls at 2 m height. The arena roof consisted of netting 'Heavy Duty BirdNet' of ½ inch polypropylene, permitting near-natural light conditions. The entire arena was constructed within a large greenhouse with a glass ceiling which had closed semi-transparent blinds which prevented direct visibility of the sky, but

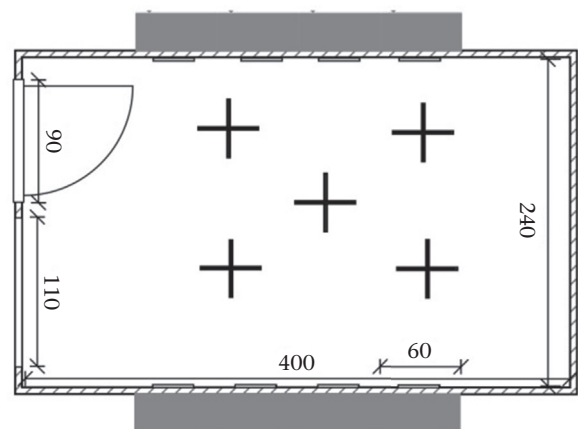


Figure 1. Schematic overview of exploration behaviour set-up. Crosses indicate five artificial trees inside the room surrounded by walls. On the left is a door and a one-way see-through observational window. Overnight cages are indicated in grey located on the outside of the top and bottom walls, cages are in two layers on top of each other. All measurements are in cm.

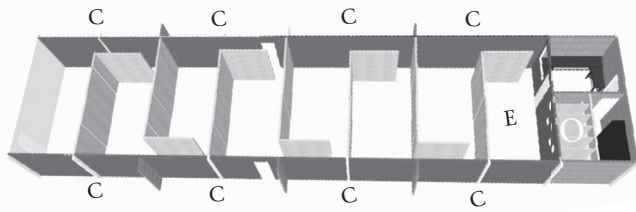


Figure 2. Schematic 3D overview of the arena set-up. Birds were housed in one of six overnight cages (O) and released in the entry room (E), the first of eight rooms connected with L-shaped corridors. Camera positions are indicated with C. A control room is in the upper right corner.

allowed enough natural light before and during testing. The ground was covered by white plastic to provide a neutral background and enhance the contrasts for the camera monitoring. Eight cameras connected to two recorders and mounted on opposite walls were used to monitor the behaviour of the birds, with each camera covering two adjacent rooms and each room covered by two cameras. Individuals were housed overnight in small (0.3×1 m and 2.4 m high) compartments adjacent to the arena. Each housing cage had a perch, ad libitum food and water, solid OSB walls and roof netting similar to the arena itself. There were six overnight cages which was the maximal number of individuals tested per day. Temperature in the greenhouse was regulated during winter so that it never fell below 4°C . Individuals were released into the arena by pulling a string from the control room which opened a 0.3×0.3 m hatch at 1.50 – 1.80 m height to the arena, thus avoiding handling of the bird before the test. After the test, the observer entered the arena from the opposite end, chasing the focal individual backwards towards the first compartment while trying to catch it manually with a handheld net. In this way individuals were prevented from fleeing into parts of the arena they had not yet explored. In addition, a window was opened into an adjacent room holding a 1×1 m mist net, which facilitated catching the bird. After testing, individuals were housed in smaller cages in the same housing room adjacent to the arena as before.

Arena Testing

Camera recordings were used to score each individual during 15 min after entering the arena using the free behavioural scoring program Jwatcher v.0.9 (Blumstein, Daniel, & Evans, 2007). We counted the landings on the net roof, floor, branches and walls and recorded the moment of entry to all corridors and rooms. A switch in position on the net was counted if the individual clearly released the net with its feet and landed more than 25 cm away from the original position, or if it ‘walked’ over 1 m of the net. Similarly, if an individual walked more than 1 m on top of a wall edge this was also added as another move. Movements of more than 1 m hopping or flying on the floor were also counted as a move. Movements within a perch were not counted. If the bird was lost from sight, for example on the floor or in a corridor, the position of the bird was approximated from the angle at which it arrived into view. All landings on the various structures were added up to an Arena Activity score, similar to the exploration behaviour score. For forward movement, we counted the different rooms visited by each individual (‘Rooms Visited’). Individuals typically travelled repeatedly between rooms, and nearly all individuals returned to previously visited rooms multiple times during the trial, sometimes all the visited rooms. For analyses, individuals that ‘froze’ and stopped moving for more than 5 min during the 15 min observation period

were excluded ($N = 16$). See the [Supplementary Material](#) for a video example of scoring.

In total, 122 individuals were tested in the arena test. In 2012 and 2013 we aimed to test all individuals twice per winter to assess within-year repeatability, resulting in a final sample size for 2012 of $N = 33$ and for 2013 of $N = 11$, with a total of 44. During 2012 and 2013, two individuals were tested three times in the arena by mistake, but data for these were retained for analyses. During 2014 and 2015 individuals were brought in only once. In total, 26 individuals were repeatedly tested between years (2012–2013; $N = 13$; 2013–2014; $N = 11$; 2014–2015; $N = 2$). Overall, 55 individuals in the arena and 236 individuals in the exploration behaviour test were tested only once, and 95 individuals were tested at least once in both the arena and the exploration behaviour test.

Analysis: Within- and Between-Individual Variation

For all analyses we used the free statistical environment R v. 2.15.1 (RCoreTeam, 2012) using the mcmcglmm-package v.2.21 (Hadfield, 2013). All estimates were made using a Bayesian approach with Markov Chain Monte Carlo generalized linear mixed-effects models (MCMC glmms; Hadfield, 2010), predominantly following the approach outlined in Dingemans and Dochtermann (2013). For our hypothesis it is important to partition phenotypic-level variance into between- and within-individual variance components, as we aimed to capture differences in between-individual repeatability and correlations (Dingemans & Dochtermann, 2013). Looking at both the within- and between-individual level of variation in behaviours makes it possible to distinguish between variation that is due to the individual’s current state rather than the innate underpinnings of the different test variables. We estimated both across- and within-year repeatability of Arena Activity and Rooms Visited, while for the exploration behaviour score, we only estimated an across-year repeatability as we only collected one score per year for this test. We further tested whether the behaviours covaried by studying between- and within-individual correlations across years. We also correlated the exploration behaviour score and the arena variables within years to detect possible influences of between-year variation in the state of individuals influencing both tests. Because we had only single measurements of exploration behaviour scores within a year we could not separate between- and within-individual variation for these correlations. Additionally, due to possible seasonal changes in state we included within-year effects of July day, which was the number of days elapsed since the previous 1 July until the test date (see Dingemans et al., 2002). Owing to habituation effects within tests we included sequence (first, second or third test within a season) for exploration behaviour scores and arena test separately. Finally, we included sex in our statistical models. We also tested whether there was an effect of seasonal state contributing to a potential correlation between scores by including year as a fixed variable and ID*year as random effect.

To answer our questions, we constructed a single model for estimating across- and within-year repeatabilities and correlations of all behaviours. One MCMC chain was generated with 1 000 000 iterations with a burn-in of 100 000 iterations. To avoid strong autocorrelations, a thinning of 200 was applied.

Repeatability and Correlations

The multivariate model used Arena Activity, Rooms Visited and exploration behaviour score as response variables. Fitted effects included sex (female/male), year (2012, 2013, 2014, 2015), sequence

within a year (1–3) and July day. The data set was structured around unique combinations of individual, year and sequence, where both exploration behaviour scores and the arena scores for these combinations could fit on the same row. Therefore, some combinations yielded empty slots for either the arena measures or exploration behaviour scores, as no such test for a specific year*sequence*individual combination was available. Because of this, we had a number of missing July day values for those empty slots. In that case, due to model requirements, we imputed the median of July day for arena scores and exploration behaviour scores separately (a sensitivity analysis replacing the missing data by more extreme values did not influence the outcome, details not shown).

Fixed-effect interactions were tested between year and sequence, and sex, year and sequence. Random effects included ID and ID*year to estimate within- and across-year repeatabilities. All response variables used a Poisson distribution. Priors were uninformative ($\nu = 0.003$) because of no prior expectations for the arena data. From the final models we obtained a posterior distribution from which mode and 95% credible intervals (CI) were calculated for repeatabilities and correlations as well as the fixed-effect parameters. All results are presented as (posterior) point estimates with $\pm 95\%$ CIs to test for effects. Across- and within-year repeatabilities were obtained from the posterior distribution by dividing the within-individual (residual) variance by the within-individual variance plus the between-individual variance for each behavioural measure. Correlations across years were obtained by dividing the covariances by the variances of the posterior distribution, using the random effect ID matrix. Correlations within years on the phenotypic level were similarly obtained by dividing the covariances by the variances of the posterior distribution using the random effect ID:year interaction matrix.

Ethical Note

All researchers involved were licensed for bird ringing, handling and measuring by the Belgian Ringing Scheme. Behavioural experiments on wild birds were licensed by the Environment, Nature and Energy Department of the Flemish Government. All bird handling was done over as little time as possible and with care, and catching was only done with consideration to weather.

RESULTS

Distribution and Variation in Variables

Exploration behaviour scores ranged from 1 to 46 (mean \pm SD = 17.9 ± 9.5) and Arena Activity scores from 41 to 256 (mean \pm SD = 136.5 ± 44.7). The frequency distribution of Rooms Visited is shown in Fig. 3. There was no effect of sex (in the sense that zero was included in the 95% CI) in slope estimates for any of the variables (Table 1). There were similarly no effects of year, July day and sequence for any of the behaviours, nor was there a year*sequence interaction for the arena scores (Table 1). Thus we conclude that there were no habituation effects in the arena behavioural measures.

Repeatability of Behaviours

Exploration behaviour scores had no across-year repeatability, but had a within-year repeatability (Table 2). Arena Activity proved to be highly repeatable within years as well as across years (Table 2). Rooms Visited was also repeatable within years, but had a slightly lower repeatability across years (Table 2).

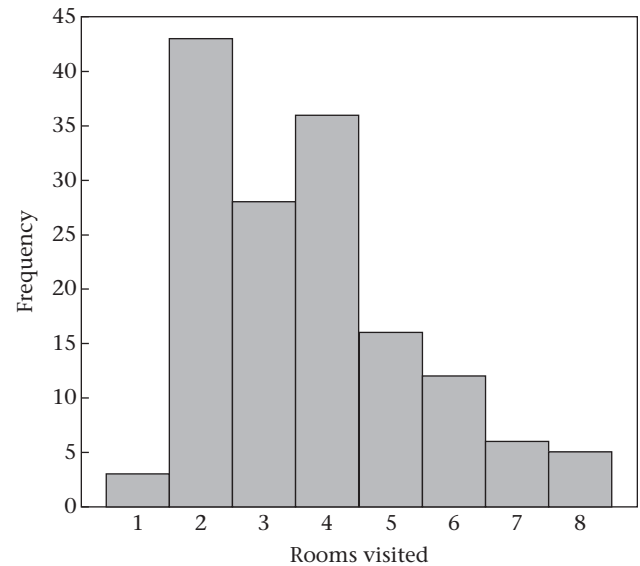


Figure 3. Frequency distribution of Rooms Visited for all arena tests. Rooms are numbered 1 to 8, with 8 being the furthest room. Birds with a score of 1 never left the Entry Room (see Fig. 2).

Correlations Between Arena Scores and Exploration Behaviour Scores

We found neither within- nor between-individual correlations across years between most of the behavioural variables, and all posterior distributions were wide. We only found a within-individual correlation between Arena Activity and Rooms Visited (Table 3). Similarly, we did not find a within-year correlation between exploration behaviour scores and the arena variables. See Table 3 for a summary of correlations and Fig. 4 illustrating the between-individual correlations using mean individual scores.

DISCUSSION

In several species exploration behaviour has been established as a repeatable personality trait, but cross-context consistency of this trait has not been thoroughly examined (but see [Dochtermann & Nelson, 2014](#)). Despite this the trait has received considerable weight in evolutionary and ecological studies ([Adriaenssens & Johnsson, 2010a, 2010b](#); [Boon, Reale, & Boutin, 2007](#); [Dingemans, Both, Drent, & Tinbergen, 2004](#); [Nicolaus et al., 2015](#)), and only recently has the cross-context validity of this trait been brought up for discussion ([Carter et al., 2013](#)). In this study we attempted to test this and we found no obvious relationship between a newly devised exploration test in a more complex environment (arena test) and exploration behaviour scores. We did not find any obvious relationship between exploration behaviour scores and Arena Activity either within years on a phenotypic level or across years on a between-individual level. However, our conclusions are restricted due to an unexpected lack of repeatability in exploration behaviour scores across years. It seems, therefore, that only a slight difference in context brings out a different behaviour of individuals even if scored in a highly similar way. This emphasizes that extrapolations of behaviour interpretations from one context to another should be done with care.

By definition, interindividual differences in expression of personality traits should be time-and/or context-independent. Arena Activity showed strong within- and across-year repeatabilities for behavioural observations, indicating that we

Table 1
Sources of variation in exploration behaviour scores and arena scores

	Exploration behaviour β (95% CIs)	Arena Activity β (95% CIs)	Rooms Visited β (95% CIs)
July day	0.002 (–0.001–0.005)	0.000 (–0.002–0.001)	0.002 (–0.001–0.005)
Sex	0.067 (–0.063–0.19)	0.17 (–0.037–0.38)	–0.10 (–0.45–0.25)
Year	0.16 (–0.40–0.75)	0.14 (–0.058–0.35)	0.18 (–0.28–0.61)
Sequence	0.33 (–0.91–1.67)	0.12 (–0.11–0.36)	0.16 (–0.34–0.68)
Year*Sequence	–0.06 (–0.66–0.48)	–0.07 (–0.27–0.12)	–0.15 (–0.55–0.28)
Sex*Year*Sequence		0.022 (–0.071–0.11)	–0.034 (–0.20–0.12)

Estimates are derived from a multivariate mixed-effect model. July day exploration behaviour, July day Arena, sex, year and sequence were fitted as fixed effects. Interaction terms included year*sequence for all behavioural variables and sex*year*sequence for the arena behaviours. Point estimates for parameters (β) are presented with 95% Bayesian credibility intervals (CIs) in parentheses.

Table 2
Within- and across year repeatabilities for exploration behaviour scores

	Repeatability within-year β (95% CIs)	Repeatability across-year β (95% CIs)
Exploration behaviour	0.63 (0.10–0.83), $N=26$	0.04 (0.003–0.32), $N=67$
Arena Activity	0.68 (0.50–0.81), $N=44$	0.50 (0.23–0.73), $N=26$
Rooms Visited	0.73 (0.21–0.92), $N=44$	0.23 (0.023–0.64), $N=26$

Exploration behaviour = exploratory behaviour measured in the small room; the other variables were measured in the arena test. Point estimates for parameters (β) are presented with 95% Bayesian credibility intervals (CIs) in parentheses.

Table 3
Correlations between exploration behaviour, Arena Activity and Rooms Visited

	Correlation within-individual β (95% CIs)	Correlation between-individual β (95% CIs)	Within-year correlation β (95% CIs)
Exploration behaviour: Arena Activity	0.91 (–0.90–0.96), $N=17$	–0.93 (–0.99–0.22), $N=95$	0.90 (–0.90–0.97), $N=92$
Exploration behaviour: Rooms Visited	0.93 (–0.79–0.97), $N=17$	–0.89 (–0.99–0.66), $N=95$	0.92 (–0.80–0.97), $N=92$
Arena Activity: Rooms Visited	0.91 (0.27–0.99) , $N=58$	0.92 (–0.30–0.99), $N=122$	

Point estimates for parameters (β) are presented with 95% Bayesian credibility intervals (CIs) in parentheses. Bold font indicates correlation where CIs do not overlap zero.

captured the expression of a personality trait in the new test. Therefore, we would expect to have captured some amount of behavioural variation found in the exploration behaviour test. Consequently, the lack of correlations between Arena Activity and exploration behaviour score was surprising. Both tests measure activity in a novel environment and following the

definition of exploration behaviour (Reale et al. 2007), these variables should reflect the same trait. Instead, it seems that we captured the expression of an independently varying trait in a novel environment, assuming no other extrinsic factors influenced the arena test. This is in line with what has been found for other study species such as crickets (Dochtermann & Nelson, 2014). However, although we found significant repeatabilities both within and across years for both arena measures, we could only find a within-year but not between-year repeatability of exploration behaviour scores. This limits the certainty of our

conclusions regarding the cross-context correlations across years.

Like Arena Activity, we also could not detect any relationship between exploration behaviour scores and Rooms Visited. If the exploration trait was expressed as a strategy to explore the environment, it could constitute a behaviourally mechanistic

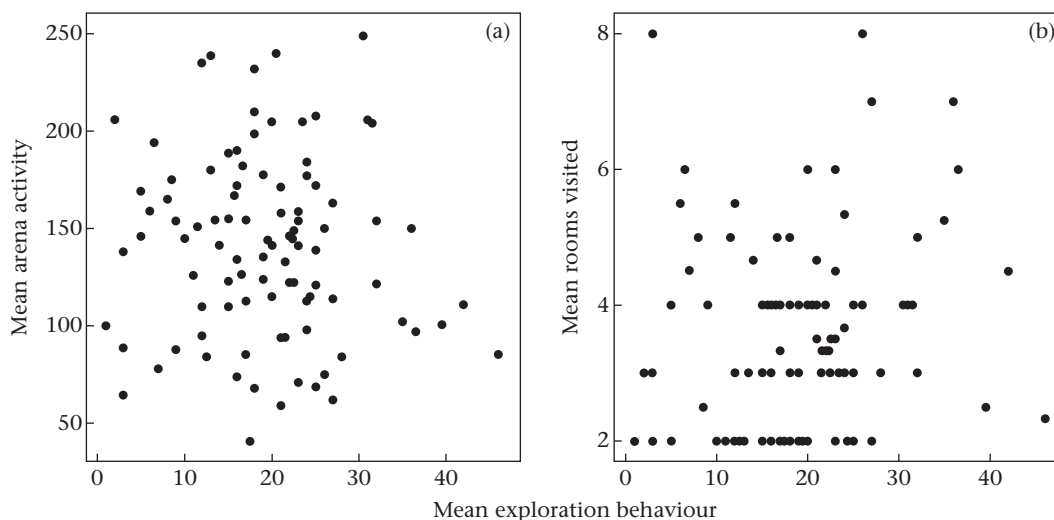


Figure 4. Between-individual correlations using the mean individual scores across years involving no other effects for (a) exploration behaviour scores and Arena Activity and (b) exploration behaviour scores and Rooms Visited. Each point represents an individual.

pathway to produce the observed correlations between exploration behaviour and spatial movement in the wild such as natal dispersal. How far an individual would choose to move in the arena test could be interpreted as an exploratory strategy, because of the need for an active choice of moving from one room to another. With the nonexistent correlation between a choice of exploring other rooms and exploration behaviour scores, questioning the convergent validity of exploration behaviour as an 'exploration trait', we are inclined to reject this idea. On a cautionary note, the lack of across-year repeatability of exploration behaviour scores means that conclusions about the relationship are necessarily limited. Nevertheless, our results support the idea that the relationship between exploration behaviour scores and spatial movement in the wild might be caused by other behavioural mechanistic pathways than an underlying exploratory strategy, such as stress reaction (Baugh et al., 2013).

Our measure of forward movement in the arena, Rooms Visited, was correlated with the Arena Activity within tests (within-individual correlation), but not on a between-individual level. This suggests that how far an individual moved in the arena may have been driven by state-dependent activity, but it could also be that how far an individual moved in the arena reflects a state-dependent motivation to explore and thus individuals became more active. Both behaviours were repeatable, meaning that they are relatively stable within and between individuals over time, but the lack of a between-individual correlation between the two behaviours suggests they measure two distinct traits. Alternatively, the correlated change in the two behaviours within individuals may suggest the influence of external factors or state of the individuals on both scores. With both Arena Activity and Rooms Visited, it is possible that we have captured repeatable variation in behaviour that is specific to the context of the arena (Niemela & Dingemans, 2014).

We found that exploration behaviour score was repeatable within years but not across years in our data set. At first sight, this would indicate that exploration behaviour is state-dependent but not an inherent trait in our sample. This is highly surprising, given that we know from earlier studies that there is a significant repeatability and heritability of this trait in our study population across multiple years. We could not test for any interobserver differences, but this seems highly unlikely, since scoring had little room for subjective interpretation and observers were taught carefully before testing. Repeatability of exploration behaviour scores in our population was earlier estimated at 0.42 which is the average repeatability reported for personality studies (Bell, Hankison, & Laskowski, 2009), but it is possible that environmental factors have influenced the expression of exploration behaviour during one or several of our years. For example, 2013 was a very bad reproductive year with many failed broods and second broods, which could have influenced future fitness expectations and in turn exploration behaviour score (Marion Nicolaus et al., 2012).

Another noticeable difference between these exploration behaviour scores and previously published data on this and other populations (Dingemans et al., 2012) is that we could not find an influence of July day, sequence or interval in our model. In previous studies, exploration behaviour score increased with July day and sequence within a year, and decreased with interval between tests, but we could not find any such effect in our model.

Conclusions

There is an acknowledgment that exploration is a complex trait, but the notion of its importance in natural selection is growing

stronger (Dochtermann & Dingemans, 2013; Wolf & Weissing, 2012). Because of its influence in evolutionary ecology, we wanted to know whether exploration behaviour could be extrapolated to a larger and more complex context, testing whether the exploration trait is expressed as an exploration strategy. Our experimental set-up did not generate proof of this but instead hinted at the opposite, that there are no generalized individual differences in exploration choices, and that there is a lack of convergent validity between different contexts. Owing to a lack of across-year repeatability we could not say this with certainty, however. Other behavioural mechanistic pathways may cause relationships between exploration behaviour and variables in the wild.

Acknowledgments

We thank Joris Elst, Indra Jacobs and Sophie Philtjens for planning and construction of the arena, bird catching and handling, Jolien Verhelst, Kaisa Telve and Emily Vanhosbeek for help with the catching and behavioural testing, Michalis Vardakis for statistical suggestions and help with catching birds. This research was funded by a BOF-TOP grant by the University of Antwerp to E.M.

Supplementary Material

Supplementary material associated with this article is available, in the online version, at <http://dx.doi.org/10.1016/j.anbehav.2016.09.005>.

References

- Adriaenssens, B., & Johnsson, J. I. (2010a). Personality and life-history productivity: Consistent or variable association. *Behavioral Ecology*, 22(1), 135–143. <http://dx.doi.org/10.1016/j.tree.2008.12.003>.
- Adriaenssens, B., & Johnsson, J. I. (2010b). Shy trout grow faster: Exploring links between personality and fitness-related traits in the wild. *Behavioral Ecology*, 22(1), 135–143. <http://dx.doi.org/10.1093/beheco/arq185>.
- Aplin, L. M., Farine, D. R., Mann, R. P., & Sheldon, B. C. (2014). Individual-level personality influences social foraging and collective behaviour in wild birds. *Proceedings of the Royal Society B: Biological Sciences*, 281(1789), 20141016. <http://dx.doi.org/10.1098/rspb.2014.1016>.
- Araya-Ajoy, Y. G., & Dingemans, N. J. (2014). Characterizing behavioural 'characters': An evolutionary framework. *Proceedings of the Royal Society B: Biological Sciences*, 281(1776), 20132645. <http://dx.doi.org/10.1098/rspb.2013.2645>.
- Arvidsson, L. K., & Matthysen, E. (2016). Individual differences in foraging decisions: Information gathering strategies or flexibility? *Behavioral Ecology*. <http://dx.doi.org/10.1093/beheco/arw054>.
- Baugh, A. T., van Oers, K., Naguib, M., & Hau, M. (2013). Initial reactivity and magnitude of the acute stress response associated with personality in wild great tits (*Parus major*). *General and Comparative Endocrinology*, 189, 96–104. <http://dx.doi.org/10.1016/j.ygcen.2013.04.030>.
- Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: A meta-analysis. *Animal Behaviour*, 77(4), 771–783. <http://dx.doi.org/10.1016/j.anbehav.2008.12.022>.
- Biro, P. A., & Stamps, J. A. (2008). Are animal personality traits linked to life-history productivity? *Trends in Ecology & Evolution*, 23(7), 361–368. <http://dx.doi.org/10.1016/j.tree.2008.04.003>.
- Blumstein, D. T., Daniel, J. C., & Evans, C. S. (Producer). (2007). JWatcher™ 0.9. Retrieved from: <http://www.jwatcher.ucla.edu/>.
- Boon, A. K., Reale, D., & Boutin, S. (2007). The interaction between personality, offspring fitness and food abundance in North American red squirrels. *Ecology Letters*, 10(11), 1094–1104. <http://dx.doi.org/10.1111/j.1461-0248.2007.01106.x>.
- Careau, V., Thomas, D., Pelletier, F., Turki, L., Landry, F., Garant, D., et al. (2011). Genetic correlation between resting metabolic rate and exploratory behaviour in deer mice (*Peromyscus maniculatus*). *Journal of Evolutionary Biology*, 24(10), 2153–2163. <http://dx.doi.org/10.1111/j.1420-9101.2011.02344.x>.
- Carere, C., Drent, P. J., Privitera, L., Koolhaas, J. M., & Groothuis, T. G. G. (2005). Personalities in great tits, *Parus major*: Stability and consistency. *Animal Behaviour*, 70(4), 795–805. <http://dx.doi.org/10.1016/j.anbehav.2005.01.003>.
- Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlshaw, G., & Heinsohn, R. (2013). Animal personality: What are behavioural ecologists measuring? *Biological Reviews of the Cambridge Philosophical Society*, 88(2), 465–475. <http://dx.doi.org/10.1111/brv.12007>.
- Chapman, B. B., Hulthen, K., Blomqvist, D. R., Hansson, L. A., Nilsson, J. A., Brodersen, J., et al. (2011). To boldly go: Individual differences in boldness

- influence migratory tendency. *Ecology Letters*, 14(9), 871–876. <http://dx.doi.org/10.1111/j.1461-0248.2011.01648.x>.
- Clobert, J., Baguette, M., Benton, T. G., & Bullock, J. M. (2012). *Dispersal ecology and evolution*. Oxford, U.K.: Oxford University Press.
- Cote, J., Clobert, J., Brodin, T., Fogarty, S., & Sih, A. (2010). Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1560), 4065–4076. <http://dx.doi.org/10.1098/rstb.2010.0176>.
- Cote, J., Fogarty, S., Weinersmith, K., Brodin, T., & Sih, A. (2010). Personality traits and dispersal tendency in the invasive mosquitofish (*Gambusia affinis*). *Proceedings of the Royal Society B: Biological Sciences*, 277(1687), 1571–1579. <http://dx.doi.org/10.1098/rspb.2009.2128>.
- Dingemanse, N. J. (2004). The relation between dominance and exploratory behavior is context-dependent in wild great tits. *Behavioral Ecology*, 15(6), 1023–1030. <http://dx.doi.org/10.1093/beheco/arp115>.
- Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society B: Biological Sciences*, 271(1541), 847–852. <http://dx.doi.org/10.1098/rspb.2004.2680>.
- Dingemanse, N. J., Both, C., Drent, P. J., Van Oers, K., & Van Noordwijk, A. J. (2002). Repeatability and heritability of exploratory behaviour in great tits from the wild. *Animal Behaviour*, 64(6), 929–938. <http://dx.doi.org/10.1006/anbe.2002.2006>.
- Dingemanse, N. J., Both, C., van Noordwijk, A. J., Rutten, A. L., & Drent, P. J. (2003). Natal dispersal and personalities in great tits (*Parus major*). *Proceedings of the Royal Society B: Biological Sciences*, 270(October 2002), 741–747. <http://dx.doi.org/10.1098/rspb.2002.2300>.
- Dingemanse, N. J., Bouwman, K. M., van de Pol, M., van Overveld, T., Patrick, S. C., Matthysen, E., et al. (2012). Variation in personality and behavioural plasticity across four populations of the great tit (*Parus major*). *Journal of Animal Ecology*, 81(1), 116–126. <http://dx.doi.org/10.1111/j.1365-2656.2011.01877.x>.
- Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: Mixed-effect modelling approaches. *Journal of Animal Ecology*, 82(1), 39–54. <http://dx.doi.org/10.1111/1365-2656.12013>.
- Dingemanse, N. J., Van der Plas, F., Wright, J., Reale, D., Schrama, M., Roff, D. A., & Barber, I. (2009). Individual experience and evolutionary history of predation affect expression of heritable variation in fish personality and morphology. *Proceedings of the Royal Society B: Biological Sciences*, 276(1660), 1285–1293. <http://dx.doi.org/10.1098/rspb.2008.1555>.
- Dochtermann, N. A., & Dingemanse, N. J. (2013). Behavioral syndromes as evolutionary constraints. *Behavioral Ecology*, 24(4), 806–811. <http://dx.doi.org/10.1093/beheco/art002>.
- Dochtermann, N. A., & Nelson, A. B. (2014). Multiple facets of exploratory behavior in house crickets (*Acheta domesticus*): Split personalities or simply different behaviors? *Ethology*, 120(11), 1110–1117. <http://dx.doi.org/10.1111/eth.12284>.
- Fraser, D. F., Gilliam, J. F., Daley, M. J., Le, A. N., & Skalski, G. T. (2001). Explaining leptokurtic movement distributions in intrapopulation variation in boldness and exploration. *American Naturalist*, 158(2), 124–135. <http://dx.doi.org/10.1086/321307>.
- Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, 127, 45–86. <http://dx.doi.org/10.1037/0033-2909.127.1.45>.
- Hadfield, J. D. (2010). MCMC methods for multi-response generalized mixed models the MCMCglmm R package. *Journal of Statistical Software*, 33(2), 1–36.
- Hadfield, J. D. (Producer). (2013). MCMCglmm-package. Retrieved from: <http://www.jstatsoft.org/v33/i02/>.
- Herborn, K. a., Macleod, R., Miles, W. T. S., Schofield, A. N. B., Alexander, L., & Arnold, K. E. (2010). Personality in captivity reflects personality in the wild. *Animal Behaviour*, 79(4), 835–843. <http://dx.doi.org/10.1016/j.anbehav.2009.12.026>.
- Hollander, F. a., Van Overveld, T., Tokka, I., & Matthysen, E. (2008). Personality and nest defence in the great tit (*Parus major*). *Ethology*, 114(4), 405–412. <http://dx.doi.org/10.1111/j.1439-0310.2008.01488.x>.
- Hoset, K. S., Ferchaud, a. L., Dufour, F., Mersch, D., Cote, J., & Le Galliard, J. F. (2010). Natal dispersal correlates with behavioral traits that are not consistent across early life stages. *Behavioral Ecology*, 22(1), 176–183. <http://dx.doi.org/10.1093/beheco/arp188>.
- Huntingford, F. A. (1976). The relationship between anti-predator behaviour and aggression among conspecifics in three-spined stickleback. *Animal Behaviour*, 24, 245–260.
- Jones, K. A., & Godin, J. G. (2010). Are fast explorers slow reactors? Linking personality type and anti-predator behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 277(1681), 625–632. <http://dx.doi.org/10.1098/rspb.2009.1607>.
- Kanda, L. L., Louon, L., & Straley, K. (2012). Stability inactivity and boldness across time and context in captive Siberian dwarf hamsters. *Ethology*, 118(6), 518–533. <http://dx.doi.org/10.1111/j.1439-0310.2012.02038.x>.
- Korsten, P., Mueller, J. C., Hermannstädter, C., Bouwman, K. M., Dingemanse, N. J., Drent, P. J., et al. (2010). Association between DRD4 gene polymorphism and personality variation in great tits: A test across four wild populations. *Molecular Ecology*, 19(4), 832–843. <http://dx.doi.org/10.1111/j.1365-294X.2009.04518.x>.
- Korsten, P., van Overveld, T., Adriaensen, F., & Matthysen, E. (2013). Genetic integration of local dispersal and exploratory behaviour in a wild bird. *Nature Communications*, 4(May), 2362. <http://dx.doi.org/10.1038/ncomms3362>.
- Matthysen, E. (2002). Boundary effects on dispersal between habitat patches by forest birds (*Parus major*, *P. caeruleus*). *Landscape Ecology*, 509–515. <http://dx.doi.org/10.1023/A:1021483915033>.
- Minderman, J., Reid, J. M., Evans, P. G. H., & Whittingham, M. J. (2009). Personality traits in wild starlings: Exploration behavior and environmental sensitivity. *Behavioral Ecology*, 20(4), 830–837. <http://dx.doi.org/10.1093/beheco/arp067>.
- Morand-Ferron, J., & Quinn, J. L. (2011). Larger groups of passerines are most efficient problem solvers in the wild. *Proceedings of the National Academy of Sciences United States of America*, 108(38), 15898–15903. <http://dx.doi.org/10.1073/pnas.1111560108>.
- Nicolaus, M., Mathot, K. J., Araya-Ajoy, Y. G., Mutzel, A., Wijmenga, J. J., Kempenaers, B., et al. (2015). Does coping style predict optimization? An experimental test in a wild passerine bird. *Proceedings of the Royal Society B: Biological Sciences*, 282(1799), 20142405. <http://dx.doi.org/10.1098/rspb.2014.2405>.
- Nicolaus, M., Tinbergen, J. M., Bouwman, K. M., Michler, S. P. M., Ubels, R., Both, C., et al. (2012). Experimental evidence for adaptive personalities in a wild passerine bird. *Proceedings of the Royal Society B: Biological Sciences*. <http://dx.doi.org/10.1098/rspb.2012.1936>.
- Niemela, P. T., & Dingemanse, N. J. (2014). Artificial environments and the study of 'adaptive' personalities. *Trends in Ecology & Evolution*, 29(5), 245–247. <http://dx.doi.org/10.1016/j.tree.2014.02.007>.
- van Oers, K., Drent, P. J., de Goede, P., & van Noordwijk, A. J. (2004). Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society B: Biological Sciences*, 271(1534), 65–73. <http://dx.doi.org/10.1098/rspb.2003.2518>.
- van Overveld, T., Adriaensen, F., & Matthysen, E. (2011). Postfledging family space use in great tits in relation to environmental and parental characteristics. *Behavioral Ecology*, 22(4), 899–907. <http://dx.doi.org/10.1093/beheco/arr063>.
- van Overveld, T., Careau, V., Adriaensen, F., & Matthysen, E. (2014). Seasonal- and sex-specific correlations between dispersal and exploratory behaviour in the great tit. *Oecologia*, 174, 109–120. <http://dx.doi.org/10.1007/s00442-013-2762-0>.
- Quinn, J. L., Patrick, S. C., Bouwhuis, S., Wilkin, T. A., & Sheldon, B. C. (2009). Heterogeneous selection on a heritable temperament trait in a variable environment. *Journal of Animal Ecology*, 78(6), 1203–1215. <http://dx.doi.org/10.1111/j.1365-2656.2009.01585.x>.
- RCoreTeam. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reale, D., Dingemanse, N. J., Kazem, A. J., & Wright, J. (2010). Evolutionary and ecological approaches to the study of personality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1560), 3937–3946. <http://dx.doi.org/10.1098/rstb.2010.0222>.
- Reale, 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(2), 291–318. <http://dx.doi.org/10.1111/j.1469-185X.2007.00010.x>.
- Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology & Evolution*, 19(7), 372–378. <http://dx.doi.org/10.1016/j.tree.2004.04.009>.
- Sih, A., Cote, J., Evans, M., Fogarty, S., & Pruitt, J. (2012). Ecological implications of behavioural syndromes. *Ecology Letters*, 15(3), 278–289. <http://dx.doi.org/10.1111/j.1461-0248.2011.01731.x>.
- Verbeek, M., Drent, P. J., & Wiepkema, P. R. (1994). Consistent individual differences in early exploratory behaviour of male great tits. *Animal Behaviour*, 44, 1113–1121.
- Wilson, A. D. M., Whattam, E. M., Bennett, R., Visanuvimol, L., Lauzon, C., & Bertram, S. M. (2009). Behavioral correlations across activity, mating, exploration, aggression, and antipredator contexts in the European house cricket, *Acheta domesticus*. *Behavioral Ecology and Sociobiology*, 64(5), 703–715. <http://dx.doi.org/10.1007/s00265-009-0888-1>.
- Wolf, M., & Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology & Evolution*, 27(8), 452–461. <http://dx.doi.org/10.1016/j.tree.2012.05.001>.