Personality, social environment, and maternal-level effects: insights from a wild kangaroo population

Weliton Menário Costa

A Thesis in the Fields of Behavioural & Evolutionary Ecology

for the Degree of Doctor of Philosophy

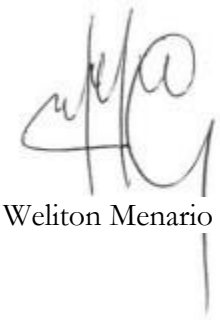
The Australian National University

May 2021

Copyright 2021 • Weliton Menário Costa

Declaration

The research presented in this thesis is my own original work except where otherwise referenced or acknowledged. The data chapters will form manuscripts with co-authors as indicated. No part of this thesis has been submitted for any previous degree at any academic institution.



Weliton Menario Costa

Abstract

The study of animal 'personality', or consistent individual differences in behaviour, has received much attention in the last two decades, but several important questions remain unclear. In particular, how important is individual personality in shaping behaviour, relative to the effects of an individual's immediate social environment? What is the interplay between individual and group effects on behavioural variation: in particular, for social species, is variation between groups in behaviour shaped by the personalities of key individuals, or by all individuals in a group? And how early in life are consistent differences in behaviour between individuals detectable? In this thesis, I address these issues using a study of a population of wild eastern grey kangaroos (*Macropus giganteus* Shaw 1790) in Victoria, Australia, in which marked individuals have been monitored since 2008.

Eastern grey kangaroos are social animals living in fusion-fission groups, and previous work on the species has shown evidence of repeatable individual variation in behaviour, or 'personalities'. However, previous personality analyses in this species did not consider the role of among-group variance in behaviour, and were limited to adult females. There is also very little known – for any species – about the effect of among-mother variance on the behavioural differences found between individuals, and hence of the contribution of maternal effects to the magnitude of individual repeatability. Here, I used a combination of experimental tests and behavioural observations to address three main questions:

(1) What is the relative importance of individual repeatability *versus* among-group variance in behavioural responses to a series of experimental stimuli?

(2) Are group-level responses driven by collective behaviour of all individuals in the group, or by the behaviour of particular individuals?

(3) How early in life is among-individual variance in behaviour apparent, and to what extent is it shaped by maternal-level effects causing differences between offspring of different mothers?

The thesis provides a comprehensive illustration of the importance of considering social context when studying animal behaviour in the wild, building on and contributing to work in the fields of behavioural and evolutionary ecology. It is composed of five chapters: a general introduction, three data chapters, and a general discussion. The data chapters were prepared as manuscripts, so some repetition in the material in each is inevitable.

In Chapter I, I establish the rationale for my research goals by reviewing the literature. I summarise ongoing animal personality research, introduce my study species and show why it is an ideal system for addressing my questions, and present previous results on the species and the study population. Finally, I describe the general methods and a thesis outline.

In Chapters II-IV, I address the main questions listed above. I use behavioural observations on over 300 individuals across three years; the data chapters are therefore interconnected parts of a single large body of work.

In Chapter II, I show (a) among-individual consistent differences in behaviour, indicative of personality, (b) among-group variance in behaviour, (c) individual behavioural change due to both habituation and location, indicative of plasticity in behaviour, but that behaviour was less affected by age or sex, (d) individual differences in habituation rates, and (e) that behavioural traits co-varied across individuals.

In Chapter III, I demonstrate that (a) the mean personality of members of a group is a better predictor of group-level responses than extreme personalities or the personality of older individuals in the group, (b) personality is also a stronger driver of mean group-level responses than other group demographic factors, and (c) individuals form groups based on their personality, over and above spatial-temporal structure, indicating that personality may contribute to social organisation.

In Chapter IV, I use measures of behaviour of very young (in the pouch) and juvenile kangaroos to show that (a) behaviour differed between offspring of different mothers (indicating that offspring of the same mother were relatively similar to each other), even at a very young age, comparable to a fetus at late development stage in eutherians, (b) juvenile survival differed between offspring of different mothers, but was independent of maternal or juvenile personality, and (c) a strong covariance between maternal and juvenile offspring personalities, probably driven by both shared environmental as well as genetic effects.

Finally, in Chapter V, I discuss the general considerations and advances from my work and set out future directions.

Frontispiece

Dedication

I dedicate this thesis to my dearest friend Carolina Mastella Botelho to inspire me to pursue a PhD and be there every second during my applications and beyond. You believed in me more than anyone else, and I thank you for that. You inspire me every day.

Acknowledgements

I did this but because we did this together.

I would like to thank (and send a massive shout-out!) my supervisor, Prof Loeske Kruuk, who has been so generous with her time and knowledge, for believing I could do so much more than I thought I could and for being so understanding when I needed. I would have never made it without your support.

Thanks to my co-supervisor, Professor Marco Festa-Bianchet, who agreed that I joined his team in the field to carry out my experiments using their long-term population, and for giving me access to the long-term data. Thanks also for all your insights and dedication to my project and being a great person to have around.

I would also like to acknowledge the other co-founders of the long-term kangaroo programme at Wilsons' Prom, A/Professor Graeme Coulson, Dr Wendy King, and Dr Dave Forsyth, and all students, research assistants, and park rangers who contributed throughout the years. In particular, I would like to thank Luca, Marc-Antoine, Charly, Becca, Carolina, Emily, Lynden, and Elowen for the fantastic work and time spent together during my fieldwork. Your dedication and hard work were mind-blowing – you are an inspiration.

Thanks to my co-authors in various chapters of this thesis. Wendy King, you were so dedicated and careful with providing comments in the text – also very resourceful with managing the long-term data and providing various forms of guidance. Timothee Bonnet, you were absolutely brilliant with statistical direction and incredibly patient in the (several) times I asked for some help. You were both an essential part of my PhD, and I appreciate that.

Thanks to the other members of my supervisory panel: Prof Rob Magrath and Prof Naomi Langmore. You were fundamental in the first stages of my project – thanks for all your generosity.

Thanks to the people responsible for making E&E function so greatly: Scott Keogh, Adrienne Nicotra, Rob Lanfear, Michael Jennious, Celeste Linde, Jack, Wes, Audra, Allen Rodrigo, Craig Moritz, IT staff, and all students who volunteered in various committees. Thanks also to all department colleagues and friends for providing an energetic research community – and beyond!

Thanks to my lab friends: Piet, Tim, Eve, Rocco, Richard, and Caro for the great conversations, support, and various pieces of advice.

Thanks also to my previous supervisor, Dr Aparecida Madella, for giving me the first tools and encouragement to take on new challenges in research. Thanks for all your advice and friendship.

Thanks to my past work team back in Brazil for the support and understanding during the stages of my PhD applications (Deila Barelli, Luciano Menini, and Valdete Tanure, to name a few).

To my family and friends back home and around the world for sending me so much love. Gloria, Valquiria, Joao, Celso, Tatiane, Emily, Nicholas, and my beautiful expected niece, Luiza. Suely, Credigar, and Nathalia. Carolina and Rafael.

Thanks to all the friends I made here in Canberra – you made everything so much better! Je, Leo, Dani, Rodrigo, Chase, George, Leandro, Alana, Chan, Amanda, Andrea – and so many others I cannot fit in the list. If you are reading this, it's you!

The Holsworth Endowment supported my fieldwork and conference travels, and I received a scholarship and tuition fee waiver from ANU.

Table of Contents

Chapter I.

General Introduction

Overview

Any person who owns at least two dogs will know that different individuals behave differently from each other – these consistent individual differences in behaviour are what is referred to as 'animal personality'. The dog owner may also realise that barking from a more suspicious dog might trigger a second more relaxed individual to also bark. This observation illustrates how individuals grouped together influence each other's behaviour. However, conducting such observations systematically in wild animals can be more challenging. To draw meaningful conclusions, sample sizes will need to be much larger, and the identification of individuals from natural markings could be difficult. In my PhD, I decided to take on this challenge. I repeatedly measured the behaviour of several hundred marked eastern grey kangaroos in their natural habitat, considering the social context, and responses to a series of experimental tests. Statistically, I dissected the relative importance of individual personality versus the social environment on the behavioural traits, and tested for other causes and consequences of behavioural variation at the individual and group levels.

Background

The field of animal 'personality'

The field of animal 'personality' has received much attention in the last two decades (Beekman and Jordan 2017), with evidence emerging for consistent among-individual differences in behaviour in multiple animal species (Réale et al. 2007). These differences have substantial importance for ecology and evolution (Wolf and Weissing 2012) as variation in individual traits is the substrate for selection to act on (Shelton and Martins 2017). By definition, among-individual differences occur within populations (Briffa 2017), and

personality exists if a significant proportion of the total variance in behaviour arises from individual differences (Réale et al. 2007). Behavioural traits that co-vary across individuals may also reflect personality (Sih et al. 2012). This research field used to be highly descriptive (Beekman and Jordan 2017), but recent advances in statistical tools have expanded its horizons into more quantitative analyses (Dingemanse et al. 2010a; Mathot et al. 2012; Dingemanse and Dochtermann 2013; Roche et al. 2016; Allegue et al. 2017).

We can measure individuals' behaviour in the wild to test for 'personality' by assessing responses to controlled stimuli (Dall and Griffith 2014) repeatedly across time (Niemelä and Dingemanse 2018; Hertel et al. 2020). The range of experimental tests used in the past goes from straightforward approaches, such as response to an approaching human (i.e., 'flight initiation distance') (Best et al. 2015; Found and St Clair 2016) or to being trapped (Réale et al. 2009), to more elaborate ones, such as response to novel objects, sounds (Found and St Clair 2016), or the simulation of predator encounters (Carlson et al. 2017). Nevertheless, most experiments tend to be restricted to assessing the immediate responses of a single target individual at a time, which may not entirely capture how individual behaviour is shaped by social influences (Webster and Ward 2011). For example, the flight initiation distance test measures an individual's immediate flee response (Weston et al. 2012), so the data describes a reaction across a few seconds only. Assessing individuals in their natural habitat is considered essential, but to date there are few studies that effectively incorporate the social context into experiments (Dall and Griffith 2014). More attention may need to be given to whether all group members are being stimulated in a setting that enhances the use of 'social information' (Danchin et al. 2004). Behavioural responses triggered by 'surprise' stimuli or recorded under brief periods may limit interaction opportunity or social information use. Also, measuring only a single target individual per

group enables the quantification of the proportion of behaviour variation caused by differences among groups.

Long-term monitoring of individuals in wild populations provides an invaluable way to understand ecological and evolutionary aspects of phenotypic variance (Kruuk and Hill 2008; Clutton-Brock and Sheldon 2010). By studying animals in their natural environment, we can measure the expression of behavioural variance among individuals in the ecological and social context in which they have evolved and potentially identify the selective pressures affecting them (Archard and Braithwaite 2010). Long-term studies allow us to link instantaneous measures of behaviour with lifetime fitness and also potentially work out a genetic basis of the variation – so we can quantify selection pressures and even potential rates of evolution (Kruuk and Hill 2008). There may be substantial difficulties associated with maintaining a long-term study (Archard and Braithwaite 2010) and, as a result, to date, not many studies have used natural populations to investigate links between personality and life-history variation (Špinka 2006; Biro and Stamps 2008; Smith and Blumstein 2008; Réale et al. 2009; Patrick and Weimerskirch 2015), maternal effects on behaviour (Taylor et al. 2012; White and Wilson 2019; Zablocki-Thomas et al. 2019), and heritability of personality (Petelle et al. 2015; Edwards et al. 2017; St-Hilaire et al. 2017).

In this thesis, I conducted observations on a wild population that has been monitored for over a decade, with many marked individuals, some of which were monitored from birth. I measured behavioural responses to standardised tests by individuals of all ages and sexes directly in the field (these data are used in all chapters) and combined these measures with data on dominance (Chapter II), group composition (Chapter III), life-history traits, and relatedness (in Chapter IV). To date, work on personality in wild populations has been dominated by studies on primates (Seyfarth et al. 2012, 2013), ungulates (Conde et al.

2002; Réale and Festa-Bianchet 2003; Réale et al. 2009), and passerine birds (Dingemanse et al. 2002; van Oers et al. 2008; Schuett and Dall 2009). Very little personality research has been done on marsupials. Studies on eastern grey kangaroo personality (Edwards et al. 2013; Best et al. 2015; Menz et al. 2017) have generally been limited to adult females, with little investigation into links between personality traits and life-history variation (e.g., survival and reproductive success) and none into their heritability. An association between individual phenotypic variation and life-history variation shows that the trait distribution may be altered by natural selection (Kruuk and Hill 2008). If there is heritable genetic variance underlying the phenotype distribution, the changes due to selection may be passed on to the next generation, causing an evolutionary change across generations (Kruuk and Hill 2008).

Behaviour may reflect personality, so to understand personality, we need to understand the causes of variation in behaviour among individuals. We can treat behavioural traits like any individual phenotypic trait (Pollak 2001), and quantitative analyses can be used to dissect their different causes of variance. Figure 1 summarises how we can characterise behavioural variation to help us understand behavioural causes and consequences. In particular, I illustrate how we test for personality (Carter et al. 2013a), plasticity (Dingemanse et al. 2010b), individual differences in plasticity (Réale and Dingemanse 2010), and multivariate associations (Sih et al. 2004; Cain et al. 2011).

As a relatively new field, animal personality has experienced some conceptual and methodological difficulties, especially given the lack of consensus on terminologies and conflicts in how the same trait is measured across taxa (Carter et al. 2013b). Because of this, I have tried throughout the thesis to avoid using secondary terms, such as 'boldness' ('willingness to take risks'), 'exploration' ('neophilia'), or 'activity' under risk (i.e., duration of foraging or probability of displacement) (Réale et al. 2007). Instead, I describe analyses of

behavioural responses that may reflect these secondary terms (or possibly a combination of them).

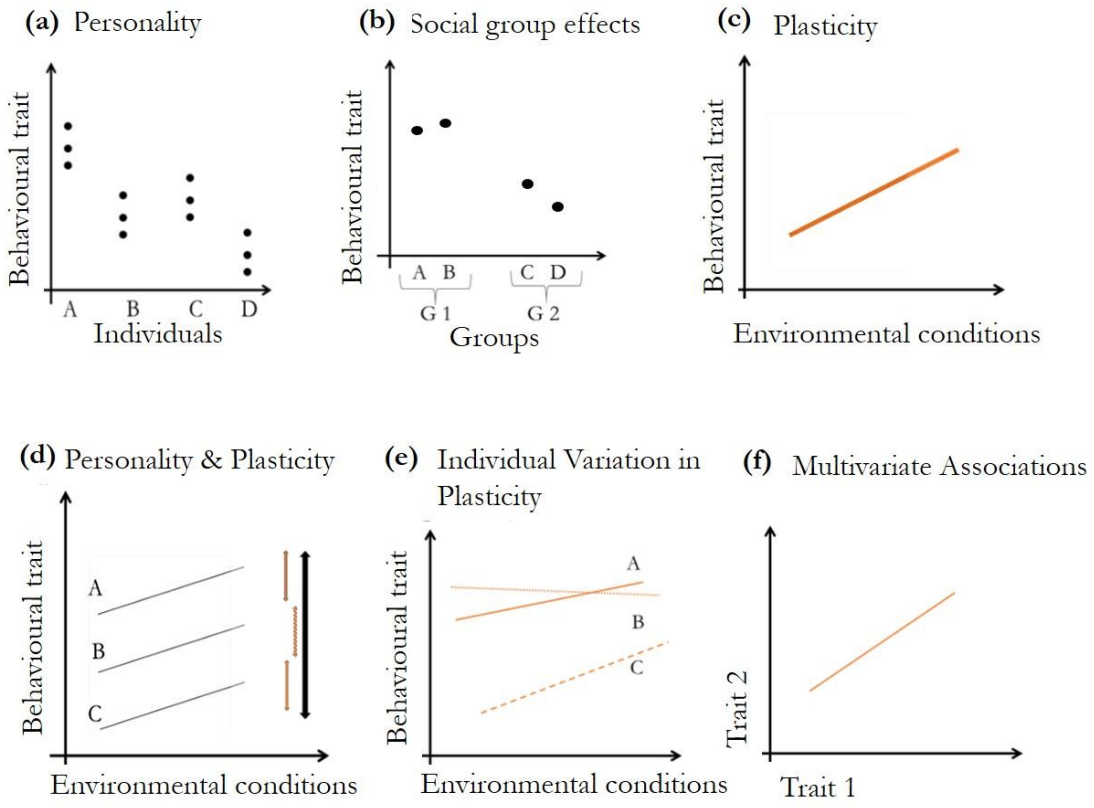


Figure 1. Characterisation of behavioural variation.

In this theoretical example, A, B, C, and D refer to different individuals and G1 and G2 (shown in (b)) to different social groups. By measuring the responses of individuals in different social groups, we can quantify statistically how much of the total variance in behaviour can be attributed to ***(a)*** *differences among individuals (or personalities) or* ***(b)*** *differences among groups. In the natural world, many other factors can also contribute to behaviour. In this example, we show* ***(c)*** *how a behavioural trait can change across an axis of environmental conditions. Despite the behavioural change due to environmental effects,* ***(d)*** *the individuals' relative responses are consistent: individuals have different intercepts of response, with A having consistently highest values for the trait and C consistently lowest. A step further,* ***(e)*** *individuals also differ in their slopes of change across the locations. With data on additional traits measured on the same individuals,* ***(f)*** *we observe a positive association between trait 1 and trait 2.*

Estimating the repeatability of traits

In the analyses here, I assess variance in behaviour, and in particular, the extent of consistent among-individual variation (or personality), and also the proportion of variance in behaviour due to differences among the levels of other biological factors, for example, differences among social groups or among offspring of different mothers (non-siblings) (Table 1), by estimating *repeatability* of measures across the different levels (Bell et al. 2009). The repeatability of a trait estimates the proportion of the total phenotypic variance attributed to differences among the levels of a biological factor (e.g., individual, social group, or mother's identity) by considering among-level and within-level variance (Stoffel et al., 2017; Wilson, 2018). This estimate indicates the importance of certain factors for phenotypic variance, but it provides no information about the mechanisms underpinning that variance (Wilson 2018). I estimated individual repeatability (Chapter II), among-group proportions or 'group repeatability' (Chapter II and III), and among-mother proportions or 'maternal repeatability' (Chapter IV), defined as the proportion of total variance explained by differences among individuals, groups, or offspring of different mothers (i.e., non-siblings), respectively (Table 1). For a behavioural trait to reflect personality, therefore statistical support for individual identity contribution to behavioural variance, the behaviour must be repeatable: there has to be statistical support for repeatability being non-zero (Bell et al. 2009; Réale and Dingemanse 2010). Different intrinsic factors can generate individual-level repeatability of a particular trait. Of these, sex and age are commonly accounted for in statistical analyses before quantifying repeatability of behavioural variance among individuals, and evidence of 'personality' is typically only concluded if there is repeatable variation over and above basic characteristics such as sex and age (Dingemanse et al. 2002; Réale et al. 2007; Dingemanse and Dochtermann 2013). Among-group proportion can be interpreted as

the magnitude of social plasticity and among-mother proportion as the proportion of individual variation due to differences between mothers.

Table 1. Definition of different repeatability calculations.

Term	Definition	Chapters
Individual repeatability	Among-individual variance divided by the total variance (Bell et al. 2009)	II and IV
Among-group proportion	Among-group variance divided by the total variance	II and III
Among-mother proportion	Among-offspring of different mothers variance divided by the total variance	IV

Mixed-effect modelling approaches are an important tool used to partitioning variance into proportions. They allow us to fit factors such as individual identity with a random effect structure, to test for differences between the levels of these factors in their average values (random-intercept models) and also in their response to covariates (random-slope models, testing for example, whether individuals vary in their rates of habituation via multiple tests) (Dingemanse and Dochtermann 2013). Generalised linear mixed models also allow analyses of non-normal data distributions and can be extended to multivariate models to consider covariances between multiple traits (Hadfield 2010). In this thesis, I used random-intercept and random-slope univariate models to address questions in Chapters II and III, and random intercept multivariate models in Chapters II and IV. I worked with Gaussian and non-Gaussian data using Markov Chain Monte Carlo (MCMC) approaches implemented in a Bayesian framework (Hadfield, 2010). This approach provides a tractable means of fitting multivariate generalised linear mixed models and assessing uncertainty on these parameters (Hadfield 2010; Nakagawa and Schielzeth 2010). Behavioural ecologists are increasingly adopting MCMC methods (Dingemanse et al. 2012; Aplin et al. 2015; St-Hilaire et al. 2017; Urszán et al. 2018). However, the practicalities of estimating repeatabilities on

non-Gaussian behavioural traits have not been widely explored in the behavioural ecology literature. My analyses here demonstrate the advantages of these methods for addressing a variety of questions on the causes and consequences of behavioural variation. I also explore some of the inherent complexities of the approach such as the difference between latent-scale versus data-scale repeatabilities from generalised linear models of non-Gaussian data, in which parameter estimates are provided on latent scales (Nakagawa and Schielzeth 2010); for such models, estimation of repeatability on the data scale requires further back-transformation, which to my knowledge has not been applied in the analyses of behavioural data. Throughout, I have provided annotated code for my analyses.

Plasticity

Individual behaviour may change depending on environmental or internal-state conditions, reflecting plasticity, and we can test this using the same models described above. Behaviour may change more rapidly than most morphological and physiological traits do, and this rapidity may be of adaptive value (Dall et al. 2004). For example, avian species with a higher propensity to innovate their food acquisition behaviour in the wild are less likely to have declining populations, and hence risk of global extinction, than less innovative species (Ducatez et al. 2020). More plastic populations may also respond slowly to selection (van Schaik 2013). We can use the same models of behavioural traits to test for plasticity and personality and test whether individuals differ consistently in their plasticity rates (Dingemanse and Araya-Ajoy 2015; Strickland and Frère 2019). In the analyses presented in this thesis, I fitted fixed effects of variables that describe a change in an environmental or temporal context, to test for plastic responses. With a particular focus on habituation (i.e., decrease in response following repeated or prolongated stimulation), I used random slope

models to test whether individuals differed in their habituation rates and if more responsive individuals would have higher or lower habituation rates.

Sociality and fission-fusion dynamics

Studying how social interactions arise, change, or are maintained informs our understanding of 'social evolution', and is therefore a foundation for my thesis. In particular, the key questions I address are how individual behaviour changes under the social context, in the presence of other conspecifics (Chapter II), and how personality contributes to group-level behaviour and group assorting (Chapter III). In Table 2, I describe the main terms related to sociality and social evolution used throughout the thesis, which may be relevant to shaping phenotypic variation (e.g., Aplin et al. 2015).

Species living in fission-fusion dynamics provide an ideal opportunity to assess behavioural variation and dissect the relative importance of individual versus group-level effects on behaviour. Fission-fusion dynamics are observed across various species (Aureli et al. 2008) and are characterised by the high frequency with which group size and composition change over time (Couzin and Laidre 2009). In mammals, fission-fusion societies have some structure, usually categorised as a multi-level organisation (Couzin and Laidre 2009), which reflect groups of different sizes. As an example, a low-level unit is a 'one-male-unit' in hamadryas baboons (*Papio hamadryas*) or a 'mother-calf-unit' in African elephants (*Loxodonta africana*). These low-level units can join or leave subgroups, and these subgroups may even fuse or fragment from each other at larger aggregation levels, depending on the context or conditions (Couzin and Laidre 2009). Individuals dynamically range across these social levels over time by joining or leaving groups. The timescale by which fissions and fusions occur varies between and within species, depending on external or individual internal conditions

(Couzin and Laidre 2009). In eastern grey kangaroos, groups with at least one adult female change every 9.4 minutes on average (King and Goldizen 2016) and group sizes can range from 1 to over 30 individuals (Jarman 1987a).

Table 2. Key terms related used in studies of sociality and social evolution, and their definitions as used in the context of this thesis.

Term	Definition
Social context	The immediate social environment of an individual, defined as the number of conspecifics with whom it is currently interacting. Used in models in this thesis to contrast 'being alone' vs not. Being 'in a social context' means being in the presence of or grouped with at least another conspecific. A social context will be different every time the 'group composition' changes.
Social environment	The physical and social environment created by conspecifics (their presence, attributes, and behaviour) when individuals are grouped together.
Social group	The conspecifics observed together at a location at a given time. I use the 10-m-chain rule (Jarman 1987) to define 'being together' as all individuals within 10 m of at least one other individual in the same group.
Group composition	The identities of the individuals composing the same social group.
Group associates	Defined as the more durable bonds within a social group, typically between a pair of individuals (compared to a general fission-fusion association of larger number of individuals in a whole group). For example, a mother with a young at any stage before weaning, or a male courting a female.
Social influence	The influence pf other group members (through their presence, attributes, or behaviour) on an individual's behaviour, causing it to change in some way.

Groups can change unpredictably or occur based on individuals' preferences. Group fissions are more likely to occur stochastically, resulting in no predictable sub-group composition if there is limited variance in the costs and benefits of interacting with different subsets of individuals (Cameron et al. 2009; Frère et al. 2010; Silk et al. 2014). If two individuals are physically far from each other, which is particularly more frequent as groups get larger, they may not interact directly, so stochasticity increases (Silk et al. 2014).

However, there are specific benefits from associating with a set of individuals in a fission-fusion social system (Couzin 2006; Silk et al. 2014). Frequent associations are usually non-random (Kerth and Konig 1999; Patriquin et al. 2010) and could happen passively or actively. Passive associations typically do not require any individual recognition capability (Couzin and Krause 2003), and this is the case of 'self-sorting': given behavioural or motivational property differences among individuals, we may observe spontaneous persistent-closer associations among individuals that share similarities in some attributes (Couzin and Krause 2003), such as social rank (Smith et al. 2007) or personality (Johnson et al. 2017). Active associations are a product of personal choice, for example, based on familiarity or environmental cues (Couzin 2006). In my study, I considered the '10-metre-chain rule' to define groups (Jarman 1987b; King et al. 2015a), meaning that individuals within 10 metres from each other will be considered grouped together. With the help of long-term marking of individuals (see more below), we could assess group composition by identifying the ID of the marked individuals grouped together. These data were used here (Chapter III) to determine grouping patterns and test whether personality contributed to assortment in our population.

Study system

Study species

Eastern grey kangaroos (Figure 2) are large-bodied grazing marsupials from the Macropod family (Macropodidae; Gray 1821), distributed across Australia's eastern mainland and eastern Tasmania. Kangaroos are crepuscular and nocturnal (Dawson 1995), have indeterminate growth (Gélin et al. 2016), an extended breeding season, and a brief gestation period (Poole 1982). Lifespan in the wild is up to 20 years (Poole 1982), but a high proportion of individuals do not survive to adulthood (Coulson et al. 2014). Females typically conceive during summer (December-January), although they are able to do so at any time of the year, and young typically leave the pouch permanently the following spring (Poole 1982). Gestation takes approximately 36 days, and after birth, the young enters the pouch and nurses while completing its development (Poole 1982). Pouch exit starts at approximately seven months of age (in around September-October) (Poole 1982), and young permanently exit the pouch at about 11 months of age. Then, the young-at-foot continues to suckle for another 7-10 months, until weaning at approximately 18-21 months (King and Goldizen 2016). Individuals reach sexual maturity at 20-36 months of age, but males mature at a later age than females (Poole 1973). Females first conceive between their second and fourth summer (Poole 1973).

There is strong sexual dimorphism in eastern grey kangaroos (Weckerly 1998), with sex differences in body size (Poole 1982), dispersal patterns (King et al. 2015b), and dominance structure (Grant 1973). Males can weigh up to 90 kg (Poole 1982), tend to disperse upon reaching sexual maturity (Stuart-Dick 1987; King et al. 2015b), and form strong dominance hierarchies (Miller et al. 2010). Females typically weigh less than 40 kg

(Poole 1982), tend to be philopatric (Stuart-Dick 1987) but non-matrilineal (Best et al. 2013), and have a weak dominance organisation (Kaufmann 1975; Maguire et al. 2006). Eastern greys are among the most social Macropods (Heathcote 1987), but neither establish territories nor defend harems (Poole 1982). Social groups are non-random 'fission-fusion' aggregations (Carter et al. 2009), meaning that individuals leave or join groups frequently, following their strong gregariousness (Poole 1982). Spatial structure (but not kinship) most strongly determines social organisation (Best et al. 2014; King et al. 2015b).

In our eastern grey kangaroo study population, there is weak fine-scale genetic structure among adult females, in reference to group composition (King et al. 2015b). There is no support for any genetic structure among adult males, as a higher proportion of males disperses (King et al. 2015b). Most groups consist of non-relatives, suggesting that kinship does not strongly affect social interactions (King et al. 2015b). Sons leave their mothers earlier (at about 18-25 months of age) than daughters. Daughters typically forage slightly closer to their mothers and are weaned about two months later than sons (King and Goldizen 2016). Mothers associate more closely with daughters than sons, although only when daughters were aged 10-29 months, before weaning (King and Goldizen 2016). Mothers generally had few interactions with their young other than nursing, which appears to be correlated with the very weak social bonding among adult kin (King and Goldizen 2016). More males than females were found to disperse as subadults or young adults (34% of sons versus 6% of daughters ($p<0.05$)). Still, amongst the females that do disperse, their median distance moved is not different from males, at 2486 m (King et al. 2015b). There have been a few cases of adoption in the population (at a rate of ~3% of all offspring), reinforcing the suggestion of a poorly developed mother-offspring recognition mechanism (King et al. 2015a).

Kangaroos are typically found in open grassy habitats (Heathcote 1987) where gregariousness may have evolved as an anti-predator defence (Jarman and Coulson 1989; Banks 2001). Across Australia, most predation is on young by wedge-tailed eagles (*Aquila audax* Latham 1801) (Olsen et al. 2010), red foxes (*Vulpes vulpes* Linnaeus 1758) (Banks et al. 2000), and dingoes (*Canis lupus dingo* Linnaeus 1758) (Poole 1982; Robertshaw and Harden 1986; Carter et al. 2009). Domestic dogs (*Canis lupus familiaris* Linnaeus 1758) are also potential predators (Cripps et al. 2011). In urbanised areas, the major cause of adult mortality is road kills, which is strongly male-biased (Coulson et al. 2014), and can increase during drought (Coulson 1989). Overall, weather conditions are linked to seasonal mortality rates as they reflect resource availability changes (Owen-Smith 2008).

As eastern grey kangaroos are potential prey to a range of species, vigilance is a critical activity (Edwards et al. 2013), as individuals trade off safety with food acquisition (Favreau et al. 2014). Individuals vary in how they balance this trade-off (Favreau et al. 2014; Gélin et al. 2015). However, social bonding can reduce the magnitude of the trade-off (Carter et al. 2009): females with large numbers of 'associates' spend less time being vigilant, and hence graze for longer than those with fewer (Carter et al. 2009). In our population, social bonding among mother and offspring, particularly before weaning (when young were 18-21 months of age), was more common than among adult females (King 2015; King et al. 2017). Some individual juveniles tended to spend more time close to their mothers when aged 18-21 months, which increased their growth and survival (King et al. 2017). These results suggest benefits of sociality (Cameron et al. 2009).

Figure 2. Eastern grey kangaroos at study site.

Marked individuals at our population at Wilsons Promontory National Park, VIC, Australia. From the left, the photograph shows two adult females, one with a large pouch young, and two subadults (a one-year-old female and a two-year-old male, in that order). The second adult female still has her Leader tags, added when her mother was captured with her in the pouch in 2009; we have not been able to capture her since then. Image credits: Weliton Menário Costa.

Study programme

I studied a wild semi-habituated eastern grey kangaroo population at Wilsons Promontory National Park, VIC (38º57'S, 146º17'E; Figure 3) that has been intensively studied for over a decade, with long-term data available on artificially marked individuals. The programme started in 2008 as a collaboration between the University of Melbourne (A/Prof Graeme Coulson), the University of Sherbrooke (Prof Marco Festa-Bianchet), and University of Queensland and Bishop's University (Dr Wendy J. King). Dr Dave Forsyth, now with the NSW Department of Primary Industries, joined the team later in the same year. At present, Prof Festa-Bianchet leads the programme, supported by the other group

members and since 2016 via a new collaboration with Prof Loeske Kruuk (The Australian National University, Australia). To my knowledge, this is the only ongoing long-term study of a natural population of individually marked macropods in the world. The combination of long-term monitoring and short-term projects on this population has generated a range of insights into the population and behavioural ecology of eastern grey kangaroos. Findings include aspects of individual fitness and sociality (Gélin et al. 2013, 2015; King et al. 2017; Quesnel et al. 2018; Toni et al. 2020; Montana et al. 2020).

Each year, the Wilsons Prom study monitors 120-140 adult females, 40-80 adult males, and 30-80 juveniles, captured with Zoletil injection from a pole syringe, mostly between August and December (King et al. 2011). We consider as adults individuals of at least three years of age. Females of more than 18 kg and males of more than 38 kg are marked with visual plastic collars of unique symbol combinations and Allflex plastic eartags with unique colours (Figure 2). Pouch young are marked when their mother is captured. We attempt to mark all the pouch young of marked females when they are at least seven months of age, weigh at least one kilogram and are fully furred. Up to two years of age, young are marked with smaller Leader ear tags. Juveniles that survive to about two years of age are recaptured in February-March of their third year, at which point their Leader ear tags are replaced with Allflex plastic tags. We attempt to catch all unmarked adult males that enter the study area during the main breeding season, from October to December. If there are few surviving juvenile females (2-year-olds) in a given year, we also capture new adult females to maintain a sample of at least 120 marked adult females. At capture, we measure body size and mass, collect a tissue sample for genetic analysis, and assess lactation status for females.

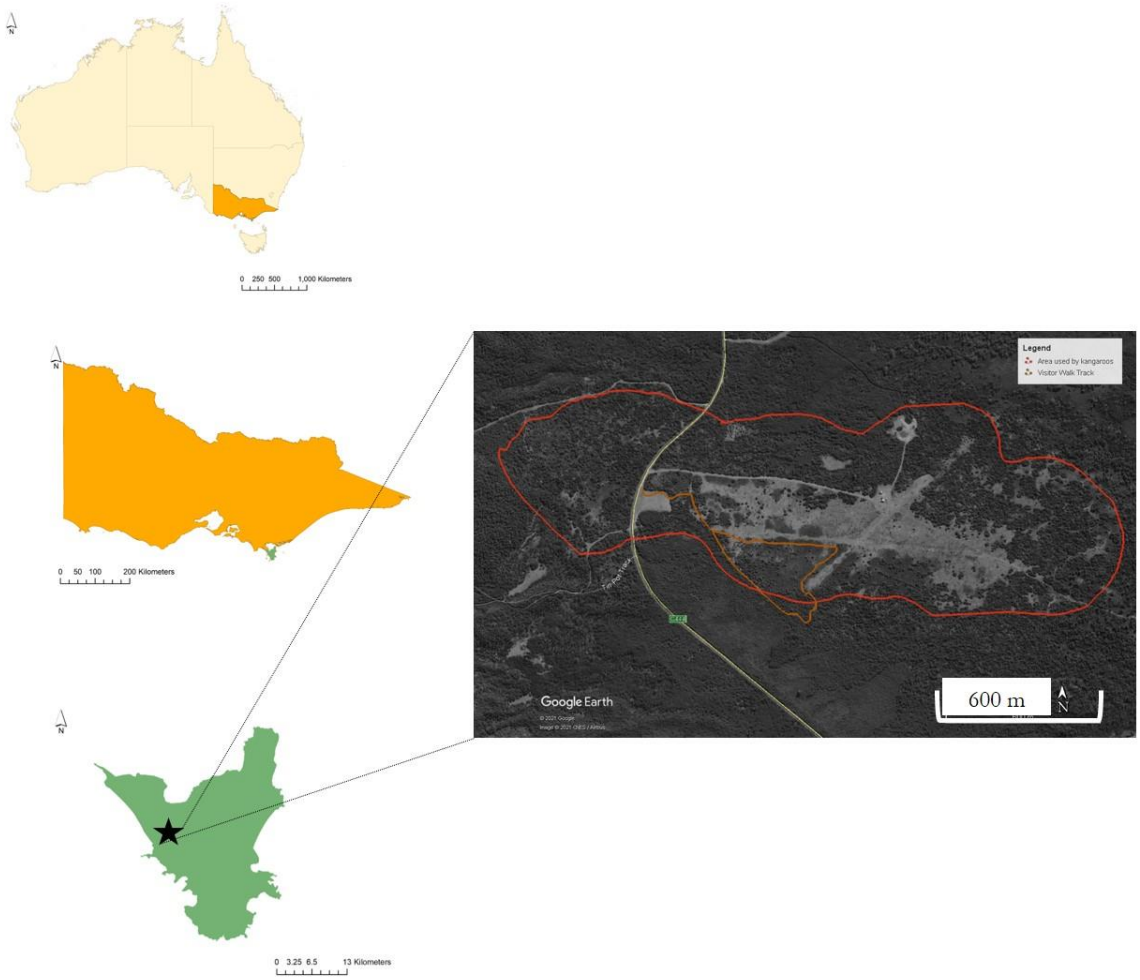


Figure 3. Location and an aerial image of the study site.

On the left, maps show my study site's location relative to the country of Australia (in beige), the state of Victoria (in orange), and the Wilsons Promontory National Park (in green). On the right, there is an aerial image of the study area. Image © 2021 CNES/Airbus © Google. The area most frequently used by kangaroos is shown with a red line. The orange line illustrates the visitor walking track, and the yellow line outlines the main road.

The long-term study also involves the collection of data on survival and reproduction via repeated observations. Marked animals are observed regularly during twice-daily surveys of the study area, approximately 150 times per year, to provide data on survival,

area use through GPS coordinates, social behaviour, and reproductive success of females. The latter is assessed through observation of an extended pouch, noting the approximate pouch size as none, small, medium or large, and the presence of a young-at-foot. Once a year, adult and sub-adult females are classified as per their reproductive cycle as 'successful' (juvenile survives to the large pouch early-stage, i.e. 7 - 8 months of age) or 'unsuccessful' (no evidence of reproduction, or juvenile does not survive to the large pouch early stage). Paternity is assessed through microsatellite analysis of tissue samples taken at capture (Rioux-Paquette et al. 2015; King et al. 2015a), and from this, variance in male reproductive success can be analysed. For more detailed methodologies of the long-term data collection, see (Toni et al. 2020; Montana et al. 2020).

General methods

My thesis chapters are interconnected pieces of a large body of work as all data come from the study population of eastern grey kangaroos in South Victoria, Australia (see above). Individuals in this population have been monitored since 2008, and we have long-term data on social and ranging behaviour, reproduction, and survival (>1000 marked individuals, >300 alive currently). A large part of my PhD was dedicated to fieldwork (9 months across three years). In this period, I conducted thousands of experimental tests and observations of group composition and dominance interactions. All thesis chapters used repeated measures on individual behaviour to quantify the causes and consequences of behavioural variation. Most behavioural responses (Table 3) were measured through standardised experimental tests between 2017 and 2019. For Chapters II and III, I also used social data from population surveys collected by multiple observers, including dominance interactions (2012-2019) and group composition (2017-2019). Life-history data were used in Chapter IV (2017-2019). Table 2 presents a summary of the data I used throughout this thesis. More details of behavioural traits and experimental tests are described in Chapters II-IV.

The study was approved by the Animal Ethics Committee at The Australian National University (Protocols A2017/17 and A2018/02), and the programme has also been approved, in previous years, through committees at the University of Melbourne (0911512.1 and 1312902.1), Université de Sherbrooke (MFB2008-03, MFB2012-02, MFB2016-01), and had research permits from the Victorian Department of Environment, Land, Water & Planning (no.s 10005558, 10007062, 10008630).

Table 3. Summary of behavioural data.

Chapter	Response	Data source	Use of behaviour as a predictor?	Both univariate and multivariate responses?
II	- Flight Initiation Distance-FID (m) - Duration of Feeding-Object (%) - Hopped Away-Object (binary) - Duration of Feeding-Sounds (%) - Hopped Away-Sounds (binary) - Dominance interactions	- Approaching Human Test - Foreign Object Test -Novel Sounds Test - Population Surveys	No	Yes
III	- Mean Group Duration of Feeding (%) - All Hopped Away (binary) -Individual's intrinsic FID	- Foreign Object and Novel Sounds Tests - Approaching Human Test (Non-Target Individuals), Foreign Object, and Novel Sounds Tests - Approaching Human Test	Yes (I used FID as a proxy for individuals' personalities as a fixed effect) Yes (I used Group ID from the group composition from population surveys as a random effect)	No
IV	- FID-Juveniles (m) - FID-Mothers (m) - Pouch Young Movement (binary) -Juvenile Survival (binary)	- Approaching Human Test - Pouch Young Measurements -Population Surveys	No	Yes

Summary of how behavioural data were used throughout the thesis. In some questions, I used the behavioural traits as response variables to quantify their (co)variance. In others, I also used them as predictors—descriptions and sample sizes in the chapters below. Additionally to the traits listed here, in Chapter II I also used principal component summaries as response variables and compared the results against the raw trait results. In Chapter III, Individual's intrinsic FID refers to the FID-'BLUP' (Best Linear Unbiased Prediction) from a model correcting for some factors, which is different from all other cases that I used the raw data.

Thesis outline

This thesis builds on and contributes to work in the areas of behavioural and evolutionary ecology. The original contribution is that I produced data and extended some important ecological and evolutionary questions to social mammal species in their natural habitat.

Relevance

Animal personality studies focus on quantifying the importance of individual differences in behaviour. However, gregarious animals do not act alone - they tend to coordinate behaviour to maintain the group, and we are missing a relative assessment of the importance of the social environment versus personality for behavioural variation in the wild. In the natural world, for any social species – like humans – personality can influence an individual's response, but the group social environment can also alter the response. To my knowledge, no studies are quantifying the relative magnitude of individual versus groups effects on behavioural phenotypes, hence, dissecting the relative importance of animal personality versus social plasticity, respectively, in experimentally measured traits. Eastern grey kangaroos live in fission-fusion dynamics, which is ideal for assessing behavioural variation in the wild and the relative importance of individual versus group-level effects. They are marsupials, readily observable, which allows us to measure the behaviour of juveniles at very early developmental stages while still in the pouch. In my thesis, I quantified individual and group effects on behaviour (Chapter II). I then tested which factors contributed to group-level responses to help us understand the causes of group differences in behaviour (Chapter III). Finally, I tested for the factors underlying individual behavioural repeatability early in life, with a particular focus on the contributions of maternal-level

effects to differences among juveniles (Chapter IV). I present a map linking the main ideas addressed in this thesis in Figure 4. My thesis is an excellent example of incorporating the social context in wildlife experiments and measuring its importance statistically.

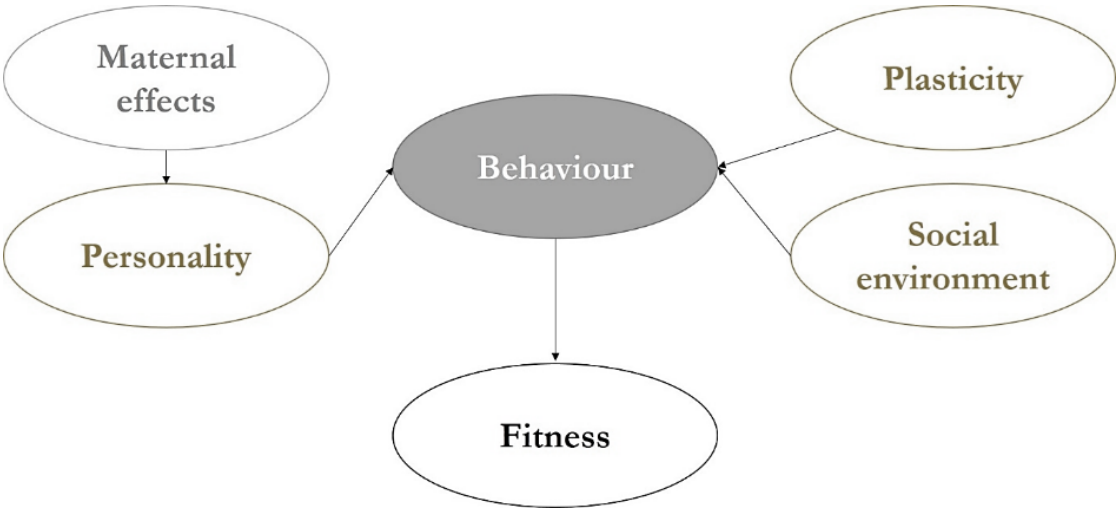


Figure 4. Thesis conceptual map.

In my thesis, I measured individual behaviour in the wild and assessed the importance of individual personality in comparison with the social environment effects for the behavioural phenotype. I also addressed other factors contributing to behavioural plasticity, such as habituation and location effect. I estimated the importance of personality to group-level response and whether personality contributes to group composition. Finally, I assessed the contribution of maternal-level effects to behavioural differences among juveniles and the fitness consequences of behaviour early in life.

Below, I describe the motivation and specific questions addressed within each of my thesis chapters:

Chapter II: Group differences in behaviour are more important than individual 'personality' in eastern grey kangaroos

Motivation

Behavioural traits can be treated like any phenotypic trait that can be measured on individuals, and quantitative analyses can be used to dissect their different causes of variance. For social, group-living species, it is not clear whether behavioural responses are likely to be more driven by individual personality or by the social environment of the group.

Specific questions

1. Is there significant variation in behaviour among individuals (personality)?
2. Is there significant variation across conditions (plasticity)?
3. Is there variation among individuals in their slopes of change across conditions (individual plasticity variation)?
4. Do behavioural traits co-vary across individuals (multivariate associations)?
5. What is more important for the observed behaviour: personality or group social effects?
6. Are there sex or age-class effects on behaviour?
7. Are results of behavioural analyses using raw traits versus principal component summaries the same?

<u>Chapter III: Whose personality determines group-level behaviour in a wild marsupial?</u>

Motivation

Group-level outcomes arise from individual decisions and hence can be shaped by individual trait variation. Previous studies have shown that personality contributes to group-level behaviour of animal social groups – but whose personality matters, and how important is personality compared to other demographic factors?

Specific questions

1. Whose personality matters for group-level responses?

2. Do the elders influence the group more than other individuals?
3. How important are individuals' personalities relative to other characteristics of the group?
4. Are kangaroos grouped based on their personalities?

Chapter IV: Maternal variance in juvenile behavior and survival in a wild marsupial

Motivation

Among-individual differences may be family-specific: offspring of different mothers are likely to differ in their phenotypes (explained by heritable genetic or maternal effects). Thus, if heritable genetic or maternal effects are important sources of individual differences in behaviour, they will play an important role in determining behavioural repeatability.

Specific questions

1. Is offspring behaviour repeatable across mothers?
2. Are offspring behavioural traits associated with survival?
3. Are there associations between mothers and offspring in behaviour?
4. Do maternal effects contribute to the repeatability of offspring behaviour?

References

Allegue H, Araya-Ajoy YG, Dingemanse NJ, et al (2017) Statistical Quantification of Individual Differences (SQuID): an educational and statistical tool for understanding multi-level phenotypic data in linear mixed models. Methods Ecol Evol 8:257–267. https://doi.org/10.1111/2041-210X.12659

Aplin LM, Firth JA, Farine DR, et al (2015) Consistent individual differences in the social phenotypes of wild great tits, Parus major. Anim Behav 108:117–127. https://doi.org/10.1016/j.anbehav.2015.07.016

Archard GA, Braithwaite VA (2010) The importance of wild populations in studies of animal temperament. J Zool 281:149–160. https://doi.org/10.1111/j.1469-7998.2010.00714.x

Aureli F, Schaffner CM, Boesch C, et al (2008) Fission-fusion dynamics: new research frameworks. Curr Anthropol 49:627–654. https://doi.org/10.1086/586708

Banks PB (2001) Predation-sensitive grouping and habitat use by eastern grey kangaroos: a field experiment. Anim Behav 61:1013–1021. https://doi.org/http://dx.doi.org/10.1006/anbe.2001.1686

Banks PB, Newsome AE, Dickman CR (2000) Predation by red foxes limits recruitment in populations of eastern grey kangaroos. Austral Ecol 25:283–291. https://doi.org/10.1046/j.1442-9993.2000.01039.x

Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? Behav Ecol 28:617–623. https://doi.org/10.1093/beheco/arx022

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Best EC, Blomberg SP, Goldizen AW (2015) Shy female kangaroos seek safety in numbers and have fewer preferred friendships. Behav Ecol 26:639–646. https://doi.org/10.1093/beheco/arv003

Best EC, Dwyer RG, Seddon JM, Goldizen AW (2014) Associations are more strongly correlated with space use than kinship in female eastern grey kangaroos. Anim Behav 89:1–10. https://doi.org/10.1016/j.anbehav.2013.12.011

Best EC, Seddon JM, Dwyer RG, Goldizen AW (2013) Social preference influences female community structure in a population of wild eastern grey kangaroos. Anim Behav 86:1031–1040. https://doi.org/10.1016/j.anbehav.2013.09.008

Biro PA, Stamps JA (2008) Are animal personality traits linked to life-history productivity? Trends Ecol Evol 23:361–368. https://doi.org/10.1016/j.tree.2008.04.003

Briffa M (2017) Abandoning animal personality would cause obfuscation: A comment on Beekman and Jordan. Behav Ecol 28:625–626. https://doi.org/10.1093/beheco/arx025

Cain KE, Rich MS, Ainsworth K, Ketterson ED (2011) Two sides of the same coin? Consistency in aggression to conspecifics and predators in a female songbird. Ethology 117:786–795. https://doi.org/10.1111/j.1439-0310.2011.01932.x

Cameron EZ, Setsaas TH, Linklater WL (2009) Social bonds between unrelated females increase reproductive success in feral horses. Proc Natl Acad Sci U S A 106:13850–13853. https://doi.org/10.1073/pnas.0900639106

Carlson N V., Healy SD, Templeton CN (2017) Hoo are you? Tits do not respond to novel predators as threats. Anim Behav 128:79–84. https://doi.org/10.1016/j.anbehav.2017.04.006

Carter AJ, Feeney WE, Marshall HH, et al (2013a) Animal personality: What are behavioural ecologists measuring? Biol Rev 88:465–475. https://doi.org/10.1111/brv.12007

Carter AJ, Feeney WE, Marshall HH, et al (2013b) Animal personality: what are behavioural ecologists measuring? Biol Rev 88:465–475. https://doi.org/10.1111/brv.12007

Carter AJ, Macdonald SL, Thomson VA, Goldizen AW (2009) Structured association patterns and their energetic benefits in female eastern grey kangaroos, *Macropus giganteus*. Anim Behav 77:839–846. https://doi.org/10.1016/j.anbehav.2008.12.007

Clutton-Brock T, Sheldon BC (2010) Individuals and populations: The role of long-term, individual-based studies of animals in ecology and evolutionary biology. Trends Ecol. Evol. 25:562–573

Conde LD, Marún C, Suib SL, Fathi Z (2002) Frequency effects in the catalytic oligomerisation of methane via microwave heating. ACS Div Fuel Chem Prepr 47:273–277. https://doi.org/10.1006

Coulson G (1989) The effect of drought on road mortality of macropods. Aust Wildl Res 16:79–83

Coulson G, Cripps J, Wilson M (2014) Hopping down the main street: eastern grey kangaroos at home in an urban matrix. Animals 4:272–291. https://doi.org/10.3390/ani4020272

Couzin ID (2006) Behavioral ecology: social organisation in fission–fusion societies. Curr Biol 16:R169–R171. https://doi.org/10.1016/J.CUB.2006.02.042

Couzin ID, Krause J (2003) Self-organisation and collective behavior in vertebrates. Adv Study Behav 32:1–75. https://doi.org/10.1016/S0065-3454(03)01001-5

Couzin ID, Laidre ME (2009) Fission–fusion populations. Curr Biol 19:R633–R635. https://doi.org/10.1016/J.CUB.2009.05.034

Cripps JK, Wilson ME, Elgar MA, Coulson G (2011) Experimental manipulation of fertility reveals potential lactation costs in a free- ranging marsupial. 859–862

Dall SRX, Griffith SC (2014) An empiricist guide to animal personality variation in ecology and evolution. Front Ecol Evol 2:1–7. https://doi.org/10.3389/fevo.2014.00003

Dall SRX, Houston AI, McNamara JM (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. Ecol Lett 7:734–739. https://doi.org/10.1111/j.1461-0248.2004.00618.x

Danchin É, Giraldeau LA, Valone TJ, Wagner RH (2004) Public information: From nosy neighbors to cultural evolution. Science (80-.). 305:487–491

Dawson TJ (1995) Kangaroos: biology of the largest marsupials. Comstock Pub. Associates

Dingemanse NJ, Araya-Ajoy YG (2015) Interacting personalities: behavioural ecology meets quantitative genetics. https://doi.org/10.1016/j.tree.2014.12.002

Dingemanse NJ, Both C, Drent PJ, et al (2002) Repeatability and heritability of exploratory behaviour in great tits from the wild. Anim Behav 64:929–938. https://doi.org/10.1006/anbe.2002.2006

Dingemanse NJ, Bouwman KM, van de Pol M, et al (2012) Variation in personality and behavioural plasticity across four populations of the great tit Parus major. J Anim Ecol 81:116–126. https://doi.org/10.1111/j.1365-2656.2011.01877.x

Dingemanse NJ, Dochtermann N, Wright J (2010a) A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets. Anim Behav 79:439–450. https://doi.org/10.1016/j.anbehav.2009.11.024

Dingemanse NJ, Dochtermann NA (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. J Anim Ecol 82:39–54. https://doi.org/10.1111/1365-2656.12013

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010b) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Ducatez S, Sol D, Sayol F, Lefebvre L (2020) Behavioural plasticity is associated with reduced extinction risk in birds. Nat Ecol Evol 4:788–793. https://doi.org/10.1038/s41559-020-1168-8

Edwards AM, Best EC, Blomberg SP, Goldizen AW (2013) Individual traits influence vigilance in wild female eastern grey kangaroos. Aust J Zool 61:332–341. https://doi.org/10.1071/ZO13025

Edwards HA, Burke T, Dugdale HL (2017) Repeatable and heritable behavioural variation in a wild cooperative breeder. Behav Ecol 28:668–676. https://doi.org/10.1093/beheco/arx013

Favreau F-R, Goldizen AW, Fritz H, et al (2014) Within-population differences in personality and plasticity in the trade-off between vigilance and foraging in kangaroos. Anim Behav 92:175–184. https://doi.org/10.1016/j.anbehav.2014.04.003

Found R, St Clair CC (2016) Behavioural syndromes predict loss of migration in wild elk. Anim Behav 115:35–46. https://doi.org/10.1016/j.anbehav.2016.02.007

Frère CH, Krützen M, Mann J, et al (2010) Social and genetic interactions drive fitness variation in a free-living dolphin population. PNAS 107:19949–19954. https://doi.org/10.1073/pnas.1007997107

Gélin U, Coulson G, Festa-Bianchet M (2015) Heterogeneity in reproductive success explained by individual differences in bite rate and mass change. Behav Ecol 27:1–7. https://doi.org/10.1093/beheco/arv209

Gélin U, Wilson ME, Coulson GM, Festa-Bianchet M (2013) Offspring sex, current and previous reproduction affect feeding behaviour in wild eastern grey kangaroos. Anim Behav 86:885–891. https://doi.org/10.1016/j.anbehav.2013.08.016

Gélin U, Wilson ME, Cripps J, et al (2016) Individual heterogeneity and offspring sex affect the growth–reproduction trade-off in a mammal with indeterminate growth. Oecologia 180:1127–1135. https://doi.org/10.1007/s00442-015-3531-z

Grant TR (1973) Dominance and association among members of a captive and a free-ranging group of grey kangaroos (*Macropus giganteus*). Anim Behav 21:449–456. https://doi.org/10.1016/S0003-3472(73)80004-1

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. J Stat Softw 33:1–22

Heathcote CF (1987) Grouping of eastern grey kangaroos in open habitat. Aust Wildl Res 14:343–8

Hertel AG, Hertel AG, Niemelä PT, et al (2020) A guide for studying among-individual behavioral variation from movement data in the wild. Mov. Ecol. 8:30

Jarman PJ (1987a) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Jarman PJ (1987b) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Jarman PJ, Coulson G (1989) Dynamics and adaptiveness of grouping in macropods. In: Grigg G, Jarman P, Hume I (eds) Kangaroos, wallabies and rat-kangaroos. Surrey Beatty & Sons Pty Ltd, New South Wales, pp 527–547

Johnson KV-A, Aplin LM, Cole EF, et al (2017) Male great tits assort by personality during the breeding season. Anim Behav 128:21–32. https://doi.org/10.1016/j.anbehav.2017.04.001

Kaufmann JH (1975) Field observations of the social behaviour of the eastern grey kangaroo, *Macropus giganteus*. Anim Behav 23:214–221. https://doi.org/10.1016/0003-3472(75)90066-4

Kerth G, Konig B (1999) Fission, fusion and nonrandom associations in female bechstein's bats (*Myotis bechsteinii*). Behaviour 136:1187–1202. https://doi.org/10.1163/156853999501711

King WJ (2015) Kinship and sociability in eastern grey kangaroos. The University of Queensland

King WJ, Festa-Bianchet M, Coulson G, Goldizen AW (2017) Long-term consequences of mother-offspring associations in eastern grey kangaroos. Behav Ecol Sociobiol 71:77. https://doi.org/10.1007/s00265-017-2297-1

King WJ, Forsyth DM, Coulson G, Festa-Bianchet M (2015a) Adoption in eastern grey kangaroos: a consequence of misdirected care? PLoS One 10:e0125182. https://doi.org/10.1371/journal.pone.0125182

King WJ, Garant D, Festa-Bianchet M (2015b) Mother-offspring distances reflect sex differences in fine-scale genetic structure of eastern grey kangaroos. Ecol Evol 5:2084–2094. https://doi.org/10.1002/ece3.1498

King WJ, Goldizen AW (2016) Few sex effects in the ontogeny of mother-offspring relationships in eastern grey kangaroos. Anim Behav 113:59–67. https://doi.org/10.1016/j.anbehav.2015.12.020

King WJ, Wilson ME, Allen T, et al (2011) A capture technique for free-ranging eastern grey kangaroos (*Macropus giganteus*) habituated to humans. Aust Mammal 33:47–51. https://doi.org/10.1071/AM10029

Kruuk LEB, Hill WG (2008) Introduction. Evolutionary dynamics of wild populations: The use of long-term pedigree data. Proc R Soc B Biol Sci 275:593–596. https://doi.org/10.1098/rspb.2007.1689

Maguire G, Ramp D, Coulson G (2006) Foraging behaviour and dispersion of eastern grey kangaroos (*Macropus giganteus*) in an ideal free framework. J Zool 268:261–269. https://doi.org/10.1111/j.1469-7998.2005.00026.x

Mathot KJ, Wright J, Kempenaers B, Dingemanse NJ (2012) Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. Oikos 121:1009–1020. https://doi.org/10.1111/j.1600-0706.2012.20339.x

Menz CS, Goldizen AW, Blomberg SP, et al (2017) Understanding repeatability and plasticity in multiple dimensions of the sociability of wild female kangaroos. https://doi.org/10.1016/j.anbehav.2017.01.012

Miller EJ, Eldridge MDB, Cooper DW, Herbert CA (2010) Dominance, body size and internal relatedness influence male reproductive success in eastern grey kangaroos

(*Macropus giganteus*). Reprod Fertil Dev 22:539–549. https://doi.org/10.1071/RD09061

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:no-no. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Niemelä PT, Dingemanse NJ (2018) On the usage of single measurements in behavioural ecology research on individual differences. Anim Behav 145:99–105. https://doi.org/10.1016/j.anbehav.2018.09.012

Olsen J, Judge D, Fuentes E, et al (2010) Diets of wedge-tailed eagles (*Aquila audax*) and little eagles (*Hieraaetus morphnoides*) breeding near Canberra, Australia. J Raptor Res 44:50–61

Owen-Smith N (2008) Effects of temporal variability in resources on foraging behaviour. In: Prins HHT, Langevelde F Van (eds) Resource Ecology: spatial and temporal dynamics of foraging. Springer Netherlands, Dordrecht, pp 159–181

Patrick SC, Weimerskirch H (2015) Senescence rates and late adulthood reproductive success are strongly influenced by personality in a long-lived seabird. Proc R Soc B Biol Sci 282:20141649–20141649. https://doi.org/10.1098/rspb.2014.1649

Patriquin KJ, Leonard ML, Broders HG, Garroway CJ (2010) Do social networks of female northern long-eared bats vary with reproductive period and age? Behav Ecol Sociobiol 64:899–913. https://doi.org/10.1007/s00265-010-0905-4

Petelle MB, Martin JGA, Blumstein DT (2015) Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (*Marmota flaviventris*). J Evol Biol 28:1840–1848. https://doi.org/10.1111/jeb.12700

Pollak E (2001) Quantitative Trait. In: Encyclopedia of Genetics. pp 1597–1598

Poole WE (1982) *Macropus giganteus*. Mamm Species 1. https://doi.org/10.2307/3504005

Poole WE (1973) A study of breeding in grey kangaroos, *Macropus giganteus* Shaw and *M. fuliginosus* (Desmarest), in Central New South Wales. Aust J Zool 21:183–212

Quesnel L, King WJ, Coulson G, Festa-bianchet M (2018) Tall young females get ahead: size-specific fecundity in wild kangaroos suggests a steep trade-off with growth. Oecologia 186:59–71. https://doi.org/10.1007/s00442-017-4003-4

Réale D, Dingemanse NJ (2010) Personality and individual social specialisation. In: Szekely T, Moore AJ, Komdeur J (eds) Social Behaviour: genes, ecology and evolution, 1st edn. Cambridge University Press, Cambridge, pp 417–441

Réale D, Festa-Bianchet M (2003) Predator-induced natural selection on temperament in bighorn ewes. Anim Behav 65:463–470. https://doi.org/10.1006/anbe.2003.2100

Réale D, Martin J, Coltman DW, et al (2009) Male personality, life-history strategies and reproductive success in a promiscuous mammal. J Evol Biol 22:1599–1607. https://doi.org/10.1111/j.1420-9101.2009.01781.x

Réale D, Reader SM, Sol D, et al (2007) Integrating animal temperament within ecology and evolution. Biol Rev 82:291–318. https://doi.org/10.1111/j.1469-185X.2007.00010.x

Rioux-Paquette E, Garant D, Martin AM, et al (2015) Paternity in eastern grey kangaroos: moderate skew despite strong sexual dimorphism. Behav Ecol 26:1147–1155. https://doi.org/10.1093/beheco/arv052

Robertshaw JD, Harden RH (1986) The ecology of the dingo in North-Eastern New South Wales IV.* Prey selection by dingoes, and its effect on the major prey species, the swamp wallaby, Wallabia bicolor (Desmarest). Aust Wildl Res 13:141–63

Roche DG, Careau V, Binning SA (2016) Demystifying animal "personality" (or not): why individual variation matters to experimental biologists. J Exp Biol 219,:3832–3843. https://doi.org/10.1242/jeb.146712

Schuett W, Dall SRX (2009) Sex differences, social context and personality in zebra finches, *Taeniopygia guttata*. Anim Behav 77:1041–1050. https://doi.org/10.1016/j.anbehav.2008.12.024

Seyfarth RM, Silk JB, Cheney DL (2013) Social bonds in female baboons: the interaction between personality, kinship and rank. Anim Behav 87:23–29. https://doi.org/10.1016/j.anbehav.2013.10.008

Seyfarth RM, Silk JB, Cheney DL (2012) Variation in personality and fitness in wild female baboons. Proc Natl Acad Sci 109:16980–16985. https://doi.org/10.1073/pnas.1210780109

Shelton DS, Martins EP (2017) Behavioral variation, adaptation, and evolution. In: APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior. American Psychological Association, Washington, pp 219–236

Sih A, Bell AM, Johnson JC, Ziemba RE (2004) Behavioral Syndromes: An Integrative Overview. Q Rev Biol 79:241–277. https://doi.org/10.1086/422893

Sih A, Cote J, Evans M, et al (2012) Ecological implications of behavioural syndromes. Ecol Lett 15:278–289. https://doi.org/10.1111/j.1461-0248.2011.01731.x

Silk MJ, Croft DP, Tregenza T, Bearhop S (2014) The importance of fission-fusion social group dynamics in birds. Int J Avian Sci 156:701–715. https://doi.org/10.1111/ibi.12191

Smith BR, Blumstein DT (2008) Fitness consequences of personality: a meta-analysis. Behav Ecol 19:448–455. https://doi.org/10.1093/beheco/arm144

Smith JE, Memenis SK, Holekamp KE (2007) Rank-related partner choice in the fission–fusion society of the spotted hyena (*Crocuta crocuta*). Behav Ecol Sociobiol 61:753–765. https://doi.org/10.1007/s00265-006-0305-y

Špinka M (2006) How important is natural behaviour in animal farming systems? Appl Anim Behav Sci 100:117–128. https://doi.org/10.1016/j.applanim.2006.04.006

St-Hilaire É, Réale D, Garant D (2017) Determinants, selection and heritability of docility in wild eastern chipmunks (*Tamias striatus*). Behav Ecol Sociobiol 71:88. https://doi.org/10.1007/s00265-017-2320-6

Strickland K, Frère CH (2019) Individual variation in the social plasticity of water dragons. Am Nat 194:194–206. https://doi.org/10.1086/704089

Stuart-Dick RI (1987) Parental investment and rearing schedules in the eastern grey kangaroo. University of New England

Taylor RW, Boon AK, Dantzer B, et al (2012) Low heritabilities, but genetic and maternal correlations between red squirrel behaviours. J Evol Biol 25:614–624. https://doi.org/10.1111/j.1420-9101.2012.02456.x

Toni P, Forsyth DM, Festa-Bianchet M (2020) Forage availability and maternal characteristics affect costs of reproduction in a large marsupial. Oecologia 193:97–107. https://doi.org/10.1007/s00442-020-04653-5

Urszán TJ, Garamszegi LZ, Nagy G, et al (2018) Experience during development triggers between-individual variation in behavioural plasticity. J Anim Ecol 87:1264–1273. https://doi.org/10.1111/1365-2656.12847

van Oers K, Drent PJ, Dingemanse NJ, Kempenaers B (2008) Personality is associated with extrapair paternity in great tits, *Parus major*. Anim Behav 76:555–563. https://doi.org/10.1016/j.anbehav.2008.03.011

van Schaik CP (2013) The costs and benefits of flexibility as an expression of behavioural plasticity: A primate perspective. Philos. Trans. R. Soc. B Biol. Sci. 368

Webster MM, Ward AJW (2011) Personality and social context. Biol. Rev. 86:759–773

Weckerly FW (1998) Sexual-size dimorphism: influence of mass and mating systems in the most dimorphic mammals. J Mammal 79:33–52

Weston MA, Mcleod EM, Blumstein DT, Guay PJ (2012) A review of flight-initiation distances and their application to managing disturbance to Australian birds. Emu - Austral Ornithol. 112:269–286

White SJ, Wilson AJ (2019) Evolutionary genetics of personality in the Trinidadian guppy I: maternal and additive genetic effects across ontogeny. Heredity (Edinb) 122:1–14. https://doi.org/10.1038/s41437-018-0082-1

Wilson AJ (2018) How should we interpret estimates of individual repeatability? Evol Lett 2:4–8. https://doi.org/10.1002/evl3.40

Wolf M, Weissing FJ (2012) Animal personalities: consequences for ecology and evolution. Trends Ecol Evol 27:452–461. https://doi.org/10.1016/j.tree.2012.05.001

Zablocki-Thomas PB, Herrel A, Karanewsky CJ, et al (2019) Heritability and genetic correlations of personality, life history and morphology in the grey mouse lemur (*Microcebus murinus*). R Soc Open Sci 6:1–11. https://doi.org/10.1098/rsos.190632

Chapter II.

Group differences in behaviour are more important than individual 'personality' in eastern grey kangaroos

Authorship

Weliton Menário Costa[1], Timothée Bonnet[1], Marco Festa-Bianchet[1,2], Loeske E.B. Kruuk[1]

1 Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT, Australia

[2] Département de biologie, Université de Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke, Québec, J1K 2R1, Canada

To be submitted to Animal Behaviour

Abstract

Animal personality is defined by behaviour, so understanding personality requires an understanding of the causes of variation in behaviour. For social, group-living species, it is not yet clear the extent to which behavioural responses are driven by individuals' personalities *vs* by the social environment of the group. Here, we analysed causes of behavioural variation in a wild population of eastern grey kangaroos (*Macropus giganteus*). Whilst previous work has established consistent between-individual repeatability of behaviour, or 'personality'; its magnitude has not been compared to the magnitude of the effect of the social environment – in this species nor, to our knowledge, in any other wild animal. We analyse here a suite of behavioural responses to different experimental stimuli for nearly 300 marked individuals over a three-year period. We found consistent individual repeatability for all traits. However, for all measures of behaviour of individuals in groups, among-group variance explained a larger proportion of the total variation in behaviour than individual repeatability. Using multivariate models, we show covariance between the repeatable components of different responses in the experiments. However, there was no evidence of associations between responses to experimental stimuli and individuals' social dominance. The study also illustrates the use of Bayesian multivariate generalised mixed models to estimate variance and covariances components for non-Gaussian behavioural response traits and the estimation of repeatability on both latent and data scales. In conclusion, our results highlight the importance of incorporating aspects of the social environment into investigations of behavioural repeatability and demonstrate how the use of repeatability data

from classic personality experiments, such as flight initiation tests, might exaggerate the importance of animal personality in social systems.

Introduction

There has been a rapid expansion of the study of 'animal personality' in recent years, with many studies establishing its existence and demonstrating important consequences. 'Personality' can be defined as individuals showing consistent, repeatable differences in behaviour across time and contexts (Stamps and Groothuis 2010). Personality may be heritable and also a target for selection (Dingemanse and Réale 2005; Dingemanse et al. 2012; Petelle et al. 2015). It can, therefore, have important ecological and evolutionary consequences. As such, the study of animal personality has recently become a highly active subfield of behavioural ecology (Beekman and Jordan 2017; Bell 2017; Dingemanse 2017; Sih 2017; Briffa 2017).

Personality is defined by behaviour, so understanding personality requires an understanding of the causes of variation in behaviour. Behavioural traits can be treated like any phenotypic trait that can be measured on individuals (Pollak 2001). Quantitative analyses can be used to dissect their different causes of variance – considering, for example, differences between sex or age classes, individual repeatability, plasticity in response to environmental variation, individual variation in plasticity, and also multivariate associations between different traits (Bell et al. 2009; Dingemanse et al. 2010b; Michelangeli et al. 2016; White et al. 2020). However, these various statistical terms have been given particular names when referring to behaviour. Specifically, individual repeatability has become 'personality' (Carter et al. 2013), plasticity may be referred to as 'behavioural reaction norms' (Dingemanse et al. 2010a), individual differences in plasticity as 'differences in the degree of specialisation' (Réale and Dingemanse 2010), and multivariate associations have become 'behavioural syndromes' (Sih et al. 2004). It is unclear whether the use of these (arguably more subjective) terms advances our understanding of the causes and consequences of variation between individuals in traits, but they undoubtedly divide opinion, which in turn has fuelled interest in the area (Beekman and Jordan 2017; Bell 2017; Dingemanse 2017; Sih 2017; Briffa 2017).

For social, group-living species, it is also not clear whether behavioural responses are likely to be more driven by individual personality or by the 'social environment' of the group. Personality could be tightly linked with 'group behaviour' (Massen and Koski 2014; Jolles et al. 2018), and fission-fusion groups provide excellent systems with which to disentangle that the relative role of intrinsic personality vs social environment, as individuals range across different groups throughout a short time (Aureli et al. 2008). Socially responsive individuals will adjust their behaviour according to the conduct of others (Dall et al. 2004), and many aspects of group decisions require a consensus, so the group does not split (Conradt and Roper 2005). Studies of animal behaviour typically consider observations of single focal individuals, but this may constitute an unrealistic scenario when considering social animals if it misses important variation driven by the social environment. In contrast, recording responses of a whole group should provide useful information on whether behaviour varies more strongly according to individuals' personalities or according to group differences. We expect personality to contribute to how an individual perceives a stimulus and gathers personal information about the environment (Bunford et al. 2019). Then, each individual in a group needs to weigh personal versus social information, available from the behaviour of others, in their response to a particular situation (East and Hofer 2010). Previous studies have demonstrated how social interactions can shape personality, behavioural plasticity, and individual differences in plasticity (Dall et al. 2004; Bergmüller et al. 2010; Dingemanse and Wolf 2013; Wolf and McNamara 2013; Montiglio et al. 2013; Planas-Sitjà et al. 2018), although more empirical studies are required to test how these processes play out in the wild. Most importantly, to our knowledge, no previous study has compared quantitative estimates of the importance of individual repeatability of behaviour relative to variation due to group effects.

Recent developments in statistical analyses have also afforded valuable advances for the study of behavioural variation. In particular, behaviour is often scored using measures that do not follow a standard normal distribution. These can be more challenging to analyse

appropriately, especially when estimating parameters such as repeatability or covariation between traits. Fortunately, Markov Chain Monte Carlo (MCMC) approaches implemented in a Bayesian framework provide a tractable means of fitting multivariate generalised linear mixed models and assess uncertainty on these parameters (Hadfield 2010; Nakagawa and Schielzeth 2010). Behavioural ecologists are now increasingly adopting MCMC methods (e.g., (Dingemanse et al. 2012; Aplin et al. 2015; St-Hilaire et al. 2017; Urszán et al. 2018). We demonstrate here the advantages of this method for addressing a variety of questions, including calculation of data-scale repeatabilities from generalised linear models of non-normal data, in which parameter estimates are provided on latent scales (Nakagawa and Schielzeth 2010); for such models, estimation of repeatability on the data scale requires further back-transformation. Furthermore, considering analytical issues, the summary of behavioural traits into principal components is commonly preferred over analysing the observed data (Budaev 2010), and we compare here analyses of raw trait data versus principal components to see how results or interpretations can be affected by this approach.

In this study, we analysed causes of behavioural variation in a wild population of eastern grey kangaroos (*Macropus giganteus,* Shaw 1790) in Victoria, Australia, which has been the subject of a long-term study (Toni et al. 2020; Montana et al. 2020). Eastern grey kangaroos are large-bodied grazing marsupials in the Macropod family (Macropodidae; Gray 1821). Their social structure typically consists of non-random fission-fusion dynamics (Carter et al. 2009a), and they spend most of their active time feeding in fission-fusion groups (Carter et al. 2009b). Previous studies of eastern grey kangaroos have revealed consistent individual differences in vigilance duration (Edwards et al. 2013), bite rate (Gélin et al. 2015), 'flight initiation distance' (FID; (Edwards et al. 2013; Best et al. 2015)) and mean foraging group size (Best et al. 2015; Menz et al. 2017). There is also evidence of plastic responses in kangaroo behaviour across temporal and social variables. For example, FID decreases with test number, indicating habituation (Edwards et al. 2013; Austin and Ramp 2019a) and with group size (Best et al. 2015). Analyses of vigilance

have also found effects of group size: larger groups showed more 'social vigilance' (i.e., monitoring of group members) (Carter et al. 2009b; Favreau et al. 2010) but less anti-predator vigilance (Favreau et al. 2010). There is also evidence for correlations between FID and other individual traits, such as vigilance duration (positive covariation; (Edwards et al. 2013)) and the number of preferred associates (negative; (Best et al. 2015). However, these previous studies have considered adult females only, or not distinguished between adult and juvenile females, relying in most cases on photo-identification of non-marked individuals in a study population in Sundown National Park, Queensland, Australia (Edwards et al. 2013; Best et al. 2015; Menz et al. 2017; Favreau et al. 2018). Further, whilst this previous work has therefore established consistent between-individual variation (repeatability or personality) in behaviour of eastern grey kangaroos, the magnitude of this repeatability has not been compared to that of the effect of the social environment – in this species nor, to our knowledge, in any other wild animal. Also, whilst individuals show plasticity in the form of habituation, the extent of individual variation in plasticity has not yet been estimated: in particular, do some individuals habituate more rapidly than others?

We report here analyses of a suite of behavioural traits of nearly 300 individuals in response to a range of artificial stimuli presented over a three-year period. Specifically, we assessed individuals' responses (i) to an approaching human (specifically, flight initiation distance FID, which is a classic test of personality in wild species, e.g. (Blumstein et al. 2003, 2015; Carter et al. 2010; Weston et al. 2012; Guay et al. 2013; Garamszegi and Møller 2017; Allan et al. 2020)), (ii) to a foreign object (Patrick and Weimerskirch 2014; Clary et al. 2014) and (iii) to novel sounds (Found and St Clair 2016). We assessed the response of only a target individual in the first test and all group members in the second and third tests. For experiments where we measured several response variables, we also summarised the traits via principal component analyses, and then analysed the first principal component in the same way as the raw trait data. We also included analyses of winners of pairwise dominance interactions based on long-term

data across nine years. We provide a comprehensive analysis of different components of behavioural variance: individual repeatability; among-group variance; sex, age and location effects; plasticity to changing environment and time (including possible habituation); and individual variance in plasticity. For traits measured on all individuals in a group, we estimated the proportion of variance due to repeatable differences between individuals (individual repeatability or 'personality') and an equivalent component for among-group differences (i.e., the proportion of total variance attributed to immediate group member effects or 'among-group proportion'), and compared the magnitude of the two. Finally, we tested for multivariate associations of the different behavioural traits across individuals, in particular to test for associations between individuals' responses to experimental stimuli vs their social dominance within the population.

Methods

Study population

Our study population is a wild population of eastern grey kangaroos at Wilsons Promontory National Park, VIC (38º57'S, 146º17'E), in which individuals have been artificially marked and monitored since 2008 (King et al. 2011). The grassy study site is mostly open, covering approximately 110 ha (Figure S1.1). The western side of the study area is bound by a road, with a visitor walking track (Figure S1.1); therefore, individuals in the west of the study area are more habituated to humans and are thus more approachable. The population at any one time comprises approximately 120-140 adult females (where 'adult' refers to individuals aged three years and above, see below), 40-80 adult males, and 30-80 juveniles every year. All individuals are ear-tagged with a unique colour combination, and females weighing over 18 kg and males weighing over 38 kg receive a unique plastic collar. Individuals are recaptured once a year (Mackay et al. 2018), at which time body mass and size are measured, female lactation status is assessed, and markings are renewed if necessary. There is no evidence of predation on adults, but wedge-tailed eagles (*Aquila audax* Latham 1801) and red foxes (*Vulpes vulpes* Linnaeus 1758) are potential predators on young. We use behavioural data from experimental tests collected from March 2017 to April 2019 and longer-term data on dominance interactions collected from March 2012 to April 2019.

Behavioural experiments

We collected behavioural data from experiments designed to investigate variation between individuals in response to various stimuli. We measured individual responses to (1) an approaching human, (2) a foreign object (i.e., a remote-controlled car), and (3) novel sounds (i.e., instrumental music playbacks) and recorded several aspects of behaviour for each test. Full details of the three experimental designs and all 16 behavioural responses recorded are given in

Supplementary Material 1, Table S1.1, but we focused our analyses on the following traits (trait names are given in italics):

- in response to an approaching human: Flight Initiation Distance (*FID*), defined as the distance between the human and the focal individual at the point at which it displaced in response to approaching. We recorded the response of one target individual per group in these tests.
- in response to either an approaching foreign object or novel sounds, whether or not an individual *Hopped Away* and the percentage of time spent feeding (*Duration Feeding)* during the trial, from 3-minute videos of each test. We recorded the response of all group members in the foreign object test in the novel sounds test.

Summary statistics and samples sizes for these five traits (*FID*, and *Hopped Away* and *Duration Feeding* for both object and sounds tests) are given in Table 1.

We carried out experiments within two hours of sunrise or sunset when kangaroos were active and easily identifiable. All trials were carried out by WMC, with the exception of 9% of the approaching human trials, which were carried out by one other *observer* (CD). We noted temporal and environmental variables at the time of testing, specifically: *date*, *temperature*, *observer*, *easting* location (distance along the west-east axis, in m), *initial distance to stimulus*, *group size*, and *group associates* (see details below). Ambient *temperature* (°C) was measured with a Kestrel® 2000 Weather-Environmental Meter, and *easting* was recorded with a Garmin® GPSMAP 64 device.

We conducted behavioural tests on kangaroos aged one year or older. *Sex* and *age* of each kangaroo were known from the long-term monitoring of the population. However, only individuals who were first caught as two-year-olds or less could be aged precisely. We, therefore, fitted *age* as a two-level category (1-year-old, older) in our models; we chose to define the age categories as 1-year-olds versus older because, for these experiments, preliminary analyses indicated that 2-year-olds behaved similarly to adults.

For each individual, we also noted the size of the group it was in and its relationship to other individuals in the group. To determine the membership of a group, we followed a '10-m chain rule' (Jarman 1987; King et al. 2015) whereby individuals are defined as being within the same group if they are within 10 m of the most peripheral individual (Supplementary Material 1). *Group size* was then defined as the number of individuals aged 1 or older within a group. We classified *group associates* of different sorts, considering more durable bonds within a social group compared to a general fission-fusion association. Levels of group associates depended on whether key individuals were also in the group: the mothers of young kangaroos, the offspring of adult females, and for adult males, a female they were courting. This resulted in a six-level factor with the following possible levels for each individual: 1. a young (aged <2) with its mother; 2. adult female with small pouch young; 3. adult female with large pouch young; 4. Adult female with 1-year-old offspring; 5. Adult male courting a female; 6. No associates. A small pouch young has a maximum age of 5 months, visible as a small or medium bulge in the pouch. A large pouch young is between 6 and 11 months of age, evident as it keeps its head out of the pouch and is completely furred. There were rare cases of females with both a large pouch young and a yearling, so to reduce a parameter in the models, we considered those cases within class 3 (with a large pouch young). Male courtship is apparent by the sexual behavioural patterns displayed (Coulson 1997).

We also tested for effects of habituation in our behavioural experiments. For each trial, we recorded *test number*, i.e., whether it was the first, second, or *n*th time an individual was being tested for that trial. To take account of the considerable variation between trials that resulted from repeats both within and between years, we also considered the *number of days since the previous test*. As this could not be defined for the first test, the first value was set to an arbitrary value larger than the largest possible number of days since the previous test (500 days; use of a different value did not affect the results).

Principal component analysis of foreign object and novel sound test traits

Our analyses focused on the five behavioural response traits described above. However, we also collected data on a range of 16 other traits (see Table S1.1). Therefore, we summarised the information across all behavioural traits from the foreign-object and novel-sound tests using a Principal Component Analysis (PCA; Table S1.2) to compare the results using principal components with those using raw traits. For each of the foreign-object and novel-sounds tests, we included all traits in a PCA (see Supplementary Material 1) and extracted the first Principal Component (PC1) for each. These two PC1s were then analysed in the same way as the five observed traits given above. We focused on analysis of just PC1 instead of using all PCs with eigenvalues greater than 1 (as sometimes adopted in PCA studies) because this part of the analysis was just used as a comparison example to see what happens to the results when we use a PC summary of traits versus the raw traits themselves. The main focus of our analyses was on the raw traits.

Dominance measures

Social dominance was assessed from observations of the outcomes of agonistic interactions between pairs of individuals (dyads), which were collected *ad libitum* from 2012 to 2019. For each dyad, one individual was randomly assigned to be the 'focal' and the other the 'opponent', following (Wilson et al. 2011), and the response variable analysed was whether the focal individual won or lost. Here, we distinguished age classes of 1-year-olds, 2-year-olds, and older. We separated 2-year-olds from older here because, at this age, their agonistic behavioural patterns are mostly different from older individuals, with 2-year olds still engaging more in playfights than fights or display based on my field observations. We also included an *'age difference'* category for the dyad and a *sex combination* code. The *age difference* category between focal and opponent was a nine-level factor denoting all possible age combinations of focal-opponent pairs (older-older, older-2yo, 2yo-older, older-1yo, 1yo-older, 2yo-2yo, 2yo-1yo, 1yo-2yo, and 1yo-

1yo). The *sex-combination* code was a four-level factor corresponding to the four possible dyad sex combinations of focal-opponent pairs (1 = male-male, 2 = female-female, 3 = male-female, and 4 = female-male). 75% of the dominance observations were among adult males, with the next most frequent (though only 6%) category being among 2-year-old males. Usually, males spotted an opponent, moved towards it and started a dominance assertion by fighting or displaying, or they approached another male that was courting a female and caused it to move away if they were clearly larger and more dominant. Encounters among subadults (i.e. ages 1 and 2) were mostly playfights, in which we designated as the loser the first individual to move away. Agonistic interactions among females were rarely observed in the field (8.7% of all observations, across all age classes). Female-female or older female-subadult interactions happened mostly when an individual approached them at a close distance. In such cases, the dominant female often batted, growled, or chased the arriving individual. Occasionally, an older female chased another individual in the group out, usually a subadult.

The dominance data involved a total of 1461 interactions involving 218 individuals recorded between 2012 and 2019 (Table 1). When carrying out the multivariate analyses to test for associations between the different behavioural traits, we restricted the interaction data set to 2017-2019 in order to match the sampling period of the experimental tests. This involved 888 pairwise interactions between a total of 176 individuals, recorded on 156 dates.

Statistical analysis

Univariate analyses: causes of variation in behavioural traits

We used a Bayesian MCMC framework (Hadfield 2010) to dissect different sources of variance between individuals in the behavioural traits scored: individual repeatability, among-group variance, sex and age classes, plasticity in response to changing environment and time, and individual variation in plasticity. We used five traits from the experimental tests (FID, and Hopped Away and Duration Feeding for both object and sounds tests) and the sixth trait, which

was a measure of dominance from the dominance observations (full list in Supplementary Material 1).

We quantified individual repeatability and among-group variance, sex and age effects, and environmental and temporal effects. We included *individual identity* (to quantify individual repeatability) and *group identity* (to quantify among-group variance) as random effects in all models, except the analyses of FID and Dominance, in which *group identity* was not included because we only measured the response of one or two target individuals, respectively. We assigned a unique *group identity* code for each tested group, so groups with the same set of individuals on a different day received a different code. We initially also included a random effect of *date of testing*, but removed it from the models as there was no support for variance in this effect. For each random effect, the models returned a posterior distribution of variance estimates, from which we could estimate a posterior mean and 95% higher posterior density credible interval (CI). We estimated repeatability of individual effects by dividing the estimates of among-individual variance by the total variance (we refer to this as 'individual repeatability'), and repeatability of group effects by dividing the estimates of among-group variance by the total variance (we refer to this as 'among-group proportion'). Calculations were carried out on each sample of the posterior distribution to generate a posterior distribution of repeatability estimates, from which we could estimate means and 95% CIs of these derived statistics. The binary traits were analysed with generalised mixed-effect models, a probit link function and residual variance set to 1?. Parameter estimates from these models are on a latent scale (Villemereuil et al. 2016); we, therefore, used the QGicc function in the R package QGglmm (Villemereuil et al. 2016) to back-transform the latent-scale variance estimates and calculate repeatability on the original data scale.

As fixed effects, we included individual intrinsic characteristics (*sex* and *age* (1-year-old or older)), variables of the physical and social environment (*temperature*, *easting*, *group size*, and *group associates*), and temporal variables (*test number*, *number of days from the previous test* to assess

habituation, and *year* (2017, 2018 or 2019)). These were used to test for *sex* and *age* effects and for plasticity to changing environment or time. We did not include *northing* (location along the south-north axis) because not much variation in habitat use by kangaroos in our study area could be captured by this variable in comparison with *easting* (Figure S1.1). Finally, we also included a fixed effect of *observer* for FID (two levels), and *initial distance to the stimulus* for Object and Sound traits, to account for any technical differences.

We initially started with a model with all fixed effects (Model 1), and then removed some fixed effects to assess the extent to which their inclusion affected the estimates of variance components. We first removed fixed effects that might be confounded with *group identity*, i.e., environmental variables, and so fitted a second model (Model 2) in which the only fixed effects were *sex*, *age*, *test number* and *number of days since the previous test*. Next, we removed additional fixed effects that might be confounded with *individual identity*, namely *sex* and *age* classes, and fitted Model 3 with only *test number* and *number of days since the previous test*. In summary, Model 1 contained all fixed effects; Model 2 excluded group-level effects but included individual-level and habituation effects, and Model 3 contained only habituation effects. The variance explained by the potentially confounding fixed effects was expected to be added to individual or group variance, and hence to affect the repeatability estimates (Wilson 2008; de Villemereuil et al. 2018).

Estimation of repeatability for the trait Hopped-Away-Sound was not feasible in Models 1 and 2. These models had more parameters than Model 3, and were prone to a statistical problem called 'quasi-complete separation' in logistic/probit regression due to some fixed effect levels being only 0 or 1 (Sauter and Held 2016). This resulted in unrealistically large variance estimates in these models (see Table S2.5), and back-conversion to the data scale (see below) was not possible. We therefore only estimated repeatabilities from Model 3 for this trait.

We found consistent effects of *test number* (see results, Table 2), indicating plasticity in all traits in this form of habituation. We, therefore, tested for individual variation in the plasticity of

habituation by fitting a random interaction between *individual identity* and *test number* (Model 4). All other fixed and random effects were the same as in Model 1. Model 4 gave a posterior distribution for the estimate of the variance in individuals' slopes in response to test number, from which we could calculate a mean and 95% CI. Slope variance estimates in the 'Hopped-Away-Sound' were disproportionally high compared to variance in other traits, presumably due to the same statistical problem of quasi-complete separation in logistic/probit regression described above (see Table S2.5; we, therefore, do not consider them further here).

In summary, we fitted four univariate models for each of the seven behavioural response traits (FID, and Hopped-Away, Duration-Feeding and PC1 for the two response trials):

- Model 1 - random intercept model with all fixed effects;
- Model 2 - random intercept model with *age*, *sex*, *test number* and *number of days from the previous test* (i.e. without group-level covariates);
- Model 3 - random intercept model with only *test number* and *number of days from the previous test* (i.e. without individual- or group-level covariates);
- Model 4 - random slope model for *test number* with all fixed effects.

We specified prior distributions for all random effects, as required for MCMCglmm (codes are given in Supplementary Material 1). We used default inverse-Wishart prior distributions (Hadfield 2018), and the residual variance was fixed at 1 for binary traits. The sample size of iterations stored was 1000 for all models and specific values of iterations, burning, and thinning intervals for each model are included in Supplementary Material 2. All models were checked for possible autocorrelations in the posterior distribution of the random variables and run to ensure autocorrelation values were below 0.1 (Hadfield 2018) (also, see Supplementary Material 1).

The Dominance trait was analysed slightly differently as it is an outcome of dyadic interactions. For each dyad, the response variable was whether the focal individual won or lost. We included sex code and age difference category as fixed effects (see above) and identity of the

focal individual, the identity of the *opponent*, and *year* (here, a 9-level factor because of the long-term data) as random effects (following Wilson et al. deer paper). The structure of the MCMCglmm priors was as above.

Multivariate analyses: covariation between traits

We estimated individual-level and (where feasible) group-level covariances between the eight behavioural traits by fitting bivariate random intercept models in MCMCglmm (Hadfield 2010) for each possible pairwise combination of traits. We initially attempted to encompass all traits into a single multivariate model, but this generated an excessive number of parameters, leading to problems with convergence and reduced reliability and requiring us to consider a series of bivariate models (one for each pair of traits) instead. We maintained the fixed and random effects used in Model 1 for the bivariate models, estimating covariances across individuals and group intercepts, when applicable. We extracted the covariance estimates and their respective 95% credible intervals from the bivariate models. A covariance was deemed significant when its credible intervals did not overlap zero (Hadfield 2018). Individual-level covariances were estimated for all traits – i.e. to assess the covariance between an individual's value for one trait with that for another. We also estimated covariance across groups between different traits in the same way – i.e. were groups with a high value for one trait also have high values for another?; however, we only estimated this group covariance for traits within either the Object or Sound tests because group identity was not fitted for the FID and dominance interaction data. Estimation of covariances between Hopped-Away-Sound and both Duration-Feeding-Object and Duration-Feeding-Sound was not feasible, following the same non-convergence problems described above.

We specified priors and checked for autocorrelation as described above. However, when specifying the priors, we used inverse Wishart priors with identity variance-covariance matrices and varying degree of belief parameter (nu) equal to the dimension of the random effect variance-covariance matrix (complete codes in Supplementary Material 1).

Ethical note

Captures were undertaken with ethics approval from the Université de Sherbrooke (no.s MFB2008-03, MFB2012-02, MFB2016-01), the University of Melbourne (no.s 0911512.1, 1312902.1) and the Australian National University (no. A2018/02) and research permits from the Victorian Department of Environment, Land, Water & Planning (no.s 10005558, 10007062, 10008630). Behavioural experiments and observations were conducted with animal ethics approval from the Australian National University (no.s A2017/17, A2018/02).

Table 1. Summary statistics and sample sizes for the behavioural traits analysed.

Test	Trait	Number of Observations	Number of Individuals	Mean Number of Tests or Interactions per Individual	Number of Groups	Mean Group Size	Trait Mean
Approaching Human	FID (m)	1642	292	5.63 (±2.91)	-	1.80 (±0.93)	7.03 (±4.32)
Foreign Object	Hopped-Away-Object (0/1)	930	240	3.88 (±2.03)	380	2.62 (±1.41)	0.51 (±0.50)
	Duration-Feeding-Object (%)						45.81 (±25.79)
	PC1-Object						0.000 (±1.75)
Novel Sounds	Hopped-Away-Sound (0/1)	657	215	3.06 (±1.57)	260	2.70 (±1.52)	0.30 (±0.46)
	Duration-Feeding-Sound (%)						55.14 (±25.57)
	PC1-Sound						0.00 (±1.93)
Dominance Interactions	Dominance (0/1)	1461	218	6.70 (±10.61)	-	-	0.52 (±0.50)

The respective standard deviations (in parentheses) follow the mean values. See Supplementary Materials 1 (Tables S1.1-3) for further details of traits and derivation of the PC1s. All measures were made in 2017-2019, with the exception of Dominance Interactions, which were assessed from 2012-2019.

Results

We first present results for the univariate analyses of each behavioural trait and the dominance measures, then for univariate analyses of the principal component scores, and finally for the multivariate analyses of covariation among traits. The main results are summarised in Table 2, but see Supplementary Material 2 (Tables S2.1-S2.8) for complete details and all parameter estimates of each model. We focus mainly on the results of Model 1 (with all fixed effects), but all four models are presented in the Suppl. Material, and we discuss the differences between them below.

Sex and age effects

Most behavioural traits did not vary with either *sex* or *age* class (Table 2a, Model 1); however, for FID*,* males were less approachable, by an average of 1.5 metres, than females (Table 2, Table S2.1). Duration-Feeding-Object was lower for older animals than 1-year-olds (Table 2). Also, the *age difference* was significant for Dominance, with focal individuals being more likely to win an agonistic encounter if they were older than the opponent or to lose if they were younger, but with no difference if focal and opponent had the same age class (Table 2, Table S2.8).

Environmental effects and plasticity in behavioural traits

We observed plasticity in behavioural traits in response to several different independent variables (Table 2, Model 1). There was a significant effect of *easting* location on FID (Table 2a). On average, individuals 1000 m to the east were less approachable by 1 m than individuals at 0 m. However, the location did not affect the Object and Sound traits (Table 2a). Moreover, there were no effects of *temperature* on any trait.

The social environment (both *group size* and g*roup associates*) also affected several responses. *Group size* affected three traits (FID, Hopped-Away-Object, and Duration-Feeding-

Object): on average, individuals in larger groups had lower FIDs and reduced responses to a foreign object (Table 2a). *Group associates* affected two traits (FID and Hopped-Away-Sound): males that were courting a female had lower FIDs, whereas females with a large pouch young had higher FIDs (i.e. were less tolerant of an approaching human) (Table 2a). However, when exposed to novel sounds, females with a yearling were less likely to hop away (Table 2a).

Traits varied with all the temporal variables measured (*test number*, *number days from the previous test,* and *year*). Most consistently, *test number* affected all traits: on average, individuals showed reduced responses with repeated tests for all seven traits, clearly indicating habituation (Table 2a). With increasing test numbers, individuals became more approachable (lower FID), spent more time feeding (higher Duration-Feeding), and were less likely to hop away in both the object and sounds tests. In line with these results, *number of days since the previous test* had the opposite effect to *test number*: with an increased *number of days since the previous test,* FID increased and Duration-Feeding-Object decreased (Table 2a), indicating some loss of habituation. Year had an effect on many traits, with individuals being less responsive in 2018 than in 2017 for all but one trait; FIDs were smaller in both 2018 and 2019 than in 2017.

Individual repeatability of behavioural traits

We found consistent individual variation in FID, Duration-Feeding-Object, and Dominance, with repeatability estimates of 49%, 8%, and 27%, respectively, all clearly above zero (Table 2, Model 1; Figure 1). Nevertheless, the repeatability estimate for Duration-Feeding-Object was <10%. The other traits (Hopped-Away-Object, Duration-Feeding-Sound, Hopped-Away-Sound) had individual repeatability estimates close to zero (Table 2). Estimates of individual repeatability did not increase when removing individual-level covariates in the majority of the traits, although there was a slight increase in FID, but the 95% CIs were overlapping, so no statistical support for a change in FID and Duration-Feeding-Object (Table 2, Model 3).

For Dominance, the random effect of Opponent explained a similar amount of variance as did Focal, as expected: individuals were as consistently likely to lose as they were to win (Table S2.8).

Among-group variance of behavioural traits

The among-group proportion of variance ranged from 19% (Duration-Feeding-Object) to 41% (Hopped-Away-Object), with posterior distributions that were clearly distinct from zero in all Object and Sound traits (Table 2b, Model 1). For these four behavioural traits, estimates of among-group proportion were consistently higher than individual repeatability estimates (Table 2; Figure 1). Dropping *group size* and *associates* resulted in slightly, but not substantially, higher estimates of among-group proportion (Table 2; Models 2 vs 1), indicating that group size and associates accounted for some but not a large amount of the differences between groups.

Individual variance in habituation

For all traits, there was evidence for among-individual variance in plasticity across *test number*, with the posterior distribution for the estimate of slope variance being clearly distinct from zero (Table 2, Model 4), thus reflecting individual variation in habituation. There was a negative covariance between individual intercepts and slopes for test number in FID and Duration-Feeding-Object (Table 2b), which may indicate more individual behavioural variation in earlier tests than in later ones.

Analysis of principal component values

We used principal component analysis to summarise the variation in the full set of 14 different behavioural responses (Table S1.1) observed, 7 for each of the Object and Sound tests (Supplementary Material 1, Table S1.2). The first principal component explained 53% of the variation in the Object traits (PC1-Object), and 53% of the variation in the Sound tests (PC1-

Sound). Duration of feeding contributed more information than any other trait to PC1, in both the Object and Sound test traits, with 'Latency from First Vigilance to Feeding' being the next most important measure in PC-Object and 'Hopped Away' in PC-Sound (Supplementary Material 1, Table S1.2).

The analysis of the principal component scores gave results that were qualitatively similar to the observed trait values (Table 2, Model 1). *Age* class, *group size, test number*, and *year* affected PC1-Object (Table 2). *Test number* and *year* affected PC1-Sound (Table 2a). We found support for individual repeatability in PC1-Object (18%), but not for PC1-Sound (0%; Table 2b, Model 1, Figure 1). There was evidence for a substantial among-group proportion of variance in both PC1-Object (16%) and PC1-Sound (59%) (Table 2b, Model 1). However, these variance proportions were not always consistent with the proportions found in the raw traits (Table 2b, Table S2.4 & S2.7). The among-group proportion was higher than individual repeatability in PC1-Sound (59% and 'no support', respectively), but the results changed for PC1-Object: among-group proportion was slightly lower than individual repeatability (16% and 18%, respectively).

Finally, we found support for individual differences in slopes of response to habituation and support for negative covariance between individual intercept and slope variance for *test number* in both PC1-traits (Table 2b, Model 4).

	Variables	Human	Object			Sounds			Dominance
		FID	Hopped Away-Object	Duration of Feeding-Object	PC 1-Object	Hopped Away-Sound	Duration of Feeding-Sound	PC 1-Sound	Won Encounter
Fixed Effects (Model 1)	Age (Older)	ns	ns	-	-	ns	ns	ns	.
	Sex (Male)	+	ns	ns	ns	ns	ns	ns	.
	Easting	+	ns	ns	ns	ns	ns	ns	.
	Group Size	-	-	+	+	ns	ns	ns	.
	Group Associates (relative to no associate):								
	- Male Courting a Female	-	ns	ns	ns	ns	ns	ns	.
	- Mother with her Large Pouch Young	+	ns	ns	ns	ns	ns	ns	.
	- Mother with her 1yo	ns	ns	ns	ns	-	ns	ns	.
	Test Number	-	-	+	+	-	+	+	.
	Number of Days from Previous Test	+	ns	-	ns	ns	ns	ns	.
	Year, relative to 2017: - 2018	-	+	ns	-	+	-	-	.
	- 2019	-	ns	ns	ns	ns	ns	ns	.
	Observer (WMC, relative to CD)	-	.	.	.	.	.	.	.
	Initial distance to the stimulus	.	ns	ns	ns	-	ns	ns	.
	Age Difference (relative to 1YO-1YO)								
	- Focal was older than opponent (2YO-1YO,	.	.	.	.	.	.	.	+
	Older-1YO, Older-2YO)	.	.	.	.	.	.	.	-
	- Focal was younger (1YO-Older, 2YO-Older)								

(b) Random effects (all models) variance and repeatability estimates	*Individual repeatability (and 95% CI)* Model 1 (all fixed effects)	0.487 (0.436, 0.560)*	Back-transformed 0.053 (0.000, 0.103)	0.079 (0.019, 0.153)*	0.183 (0.097, 0.239)*	|	0.000 (0.000, 0.090)	0.001 (0.000, 0.051)	Back-transformed 0.267 (0.214, 0.353)*
	Model 2 (without group-level fixed effects)	0.544 (0.471, 0.581)*	0.045 (0.000, 0.091)	0.083 (0.013, 0.144)*	0.153 (0.085, 0.225)*	|	0.000 (0.000, 0.087)	0.001 (0.000, 0.048)	.
	Model 3 (without individual or group-level fixed effects)	0.523 (0.473, 0.584)*	0.038 (0.000, 0.089)	0.076 (0.025, 0.152)*	0.172 (0.090, 0.228)*	0.002 (0.000, 0.061)	0.000 (0.000, 0.105)	0.000 (0.000, 0.058)	.
	Among-group proportion (and 95% CI) Model 1	.	0.408 (0.334, 0.496)*	0.185 (0.126, 0.286)*	0.155 (0.069, 0.218)*	|	0.361 (0.260, 0.445)*	0.588 (0.520, 0.666)*	.
	Model 2	.	0.426 (0.336, 0.489)*	0.225 (0.142, 0.292)*	0.172 (0.099, 0.238)*	|	0.372 (0.271, 0.454)*	0.618 (0.546 , 0.681)*	.
	Model 3	.	0.424 (0.339, 0.496)*	0.220 (0.159, 0.305)*	0.165 (0.100, 0.250)*	0.605 (0.517, 0.695)*	0.368 (0.263, 0.434)*	0.613 (0.539, 0.681)*	.
	Slope variance for Test Number – Model 4 (and 95% CI)	0.167 (0.098, 0.238)*	1.132 (0.074, 2.859)*	6.927 (1.923, 12.920)*	0.129 (0.077, 0.191)*	|	1.470 (0.171, 3.734)*	0.126 (0.010, 0.271)*	.
	Individual Intercept x Individual Test Number Slope covariance – Model 4	-1.014 (-1.509, -0.533)*	-4.419 (-11.98, 0.140)	-38.63 (-66.06, -14.15)*	-0.397 (-0.606, -0.206)*	|	-1.013 (-7.824, 4.127)	-0.126 (-0.271, -0.009)*	.

Table 2. Summary of the models of behavioural traits.

We show here explanatory variables/levels with significant results only – for complete outputs, see Supplementary Material 2 (Tables S2.1-S2.8). ***(a)*** *Summary of the fixed effects with at least a significant result across the traits, for Model 1. + and - symbols reflect the direction of the significant effect (pMCMC<0.05). 'ns' indicates no statistical support for a fixed effect, and '.' indicates that that fixed effect was not fitted in the model. Parameter estimates are given in full detail in Tables S2.1-S2.8.* ***(b)*** *Estimates of individual repeatability and among-group*

*proportion and slope variance from Models 1-4. '*****' indicates that credible intervals did not overlap zero, so there was statistical support. Model 1 - random intercept model with all fixed effects, Model 2 - random intercept model with age, sex, test number and number of days from the previous test, Model 3 - random intercept model with only test number and number of days from the previous test, and Model 4 - random slope model for test number with all fixed effects (see Methods for further details). Values shown in slope variance/covariance are the posterior distribution means, followed by the 95% credible intervals in parentheses. Hopped-Away-Sound random effect estimates were not feasible for Models 1, 2, and 4 ('|', see Methods). The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017), Observer (CD), Age Difference (1-year-old/1-year-old). Results of raw traits are not shaded and of principal components are shaded in grey.*

Trait Covariance across Individuals								
		Human	Object			Sounds		
	Variables	FID	Hopped-Away-Object	Duration-Feeding-Object	PC 1-Object	Hopped-Away-Sound	Duration-Feeding-Sound	PC 1-Sound
Object	Hopped-Away-Object	3.193 (1.682, 4.845)*						
	Duration-Feeding-Object	-15.002 (-22.482, -7.496)*	-1.251 (-5.745, 2.706)					
	PC 1-Object	-1.335 (-1.906, -0.744)*	.	.				
Sounds	Hopped-Away-Sound	32.700 (3.053, 71.430)*	1.000 (-36.918, 42.959)	|	-0.039 (-1.184, 1.679)			
	Duration-Feeding-Sound	-9.605 (-17.525, -2.896)*	-0.324 (-4.095, 2.895)	22.140 (-1.320, 50.110)	0.983 (-0.327, 2.974)	|		
	PC 1-Sound	-0.756 (-0.245, -1.229)*	-0.011 (-0.802, 0.927)	-0.982 (-2.513, 0.358)	0.106 (0.239, -0.017)	.	.	
Dominance		4.766 (-0.076, 9.182)	-1.580 (-3.882, 0.509)	25.740 (3.355, 48.950)*	0.845 (-0.774, 2.500)	-0.135 (-3.223, 3.490)	-9.650 (-31.251, 16.120)	0.827 (-0.873, 2.614)

Table 3. Covariance between individual-level behavioural traits from bivariate mixed models.

*Values shown are the means of the posterior distribution of covariance estimates, followed by the 95% credible intervals in parentheses. '.' indicates that the covariance was not estimated between the corresponding traits, '*****' indicates that credible intervals did not overlap zero, indicating statistical support. Covariances between Hopped-Away-Sound and Duration-Feeding-Object and Duration-Feeding-Sound were not feasible ('|', see Methods). Sample sizes: 1642 observations on 292 individuals (Approaching Human test), 930 observations on 240 individuals (Object test), 657 observations on 215 individuals (Sound test), and 888 pairwise observations on 218 individuals (Dominance observations).*

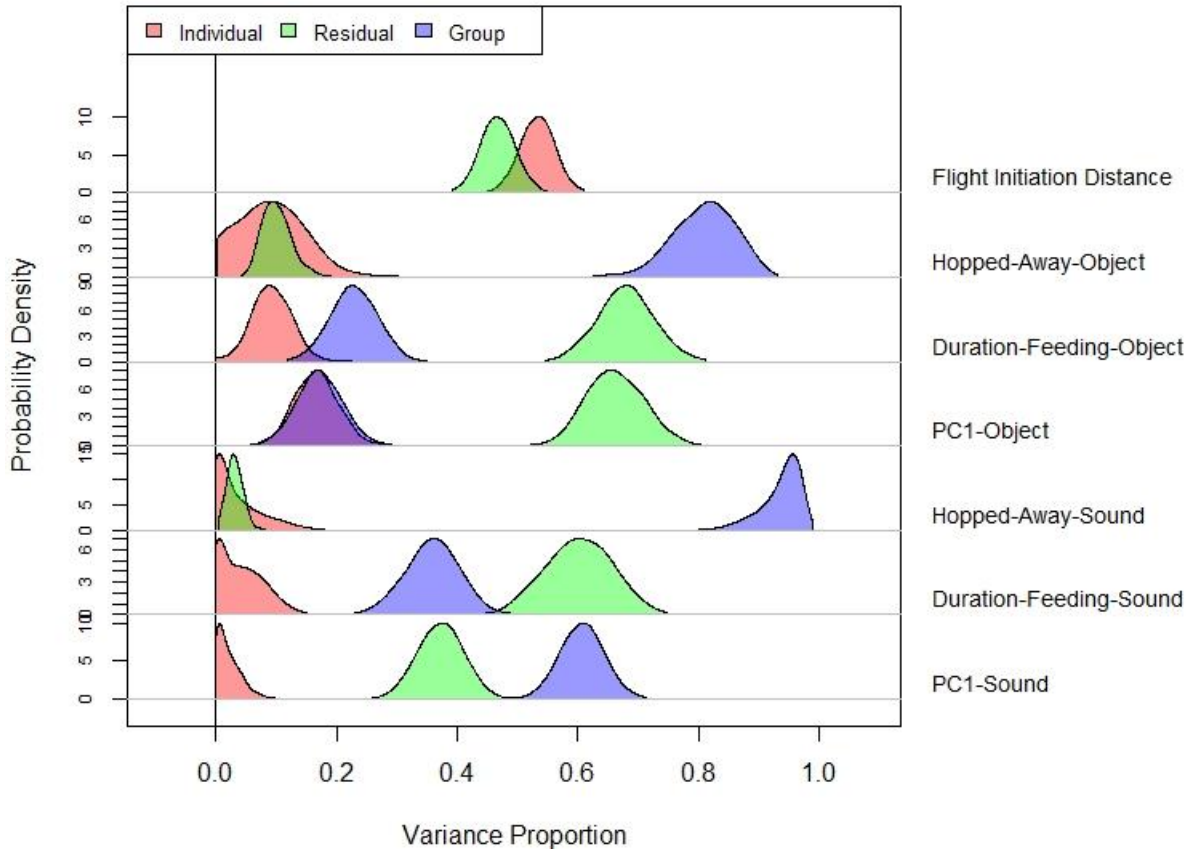


Figure 1. Variance decomposition for seven behavioural traits, showing the relative importance of among-individual versus among-group differences.

For each trait, the posterior distribution of each variance component contained in the model is shown as a proportion. For the trait FID, we did not model group-level variance because there was only one target individual per group (see Methods).

Covariance among traits

We found evidence for covariances within individuals between several of the different traits. In particular, there was support for individual-level covariances between FID and most other behavioural traits measured (Table 3a): individuals with a high FID value were more likely to hop away in both the foreign object and novel sound test and correspondingly to spend less time feeding. There were also covariances between FID and the PC1 values for both traits; however, not between FID and Dominance.

However, despite the covariances of FID with other traits, we found little indication of associations between the traits measured in response to an experimental trial and our measure of social dominance. Duration-Feeding-Object was the only trait that covaried with Dominance, with more dominant individuals spending a longer time feeding. Still, there was no evidence of covariance with any of the other traits.

At the group level, we could only estimate covariance within the foreign Object test, but we found a negative covariance among Duration-Feeding-Object and Hopped-Away-Object: as expected, groups with higher average feeding duration had a lower frequency of members that hopped away in response to the foreign object trial (-12.130 (CI -20.041,-3.011)).

Discussion

This chapter dissects the sources of variation, including individual repeatability and among-group variance, sex, age class, habituation and aspects of sociality and of the environment, in a suite of behavioural traits in our study population of eastern grey kangaroos. Our analyses indicated consistent individual and group differences in behaviour. Sex and age class did not have a large effect on the majority of the experimental traits, but age differences contributed to an individual's propensity to win a dominance encounter. Group size, group associates, and easting location contributed to some responses, but test number was particularly significant across all traits, indicating habituation to tests. We also demonstrated that individuals varied in their rates of habituation. Finally, the multivariate analyses indicated covariance between several components of the traits, although not between the experimental traits and social dominance.

Personality versus social plasticity

There were consistent differences between individuals and also between groups in the responses to our experimental trials. Despite evidence for repeatable individual variance (personality), group variances were consistently larger than individual variances for all raw traits where we estimated both (Figure 1). This suggests that the immediate group environment is more important than personality for behavioural variation in a natural setting where individuals rely on both social and personal information. This is a good reminder that, in social species, interacting individuals may exchange information, affecting their odds of responding in a certain way (Fögen 2014). In the Object and Sound experiments, the stimuli were introduced to all group members and responses were recorded for a period of time after the stimuli, so that animals had time to assess each other's responses while reacting (East and Hofer 2010). Individual FID is a trait that is relatively easily measured in the field (Weston et al. 2012), and is frequently repeatable across individuals (Møller 2008; Carrete and Tella 2010; Guay et al. 2013).

However, because it only considers the first few seconds of a single individual's reaction per group, it may not capture the full complexity of behavioural responses in a social system.

Among-group variance can be considered to indicate the effect of social plasticity on behaviour in the same way as individual repeatability indicates the importance of personality. Our results therefore indicate that social plasticity is more important for an individual's behavioural phenotype than personality. Many studies, for example of social insects, fishes, birds, and mammals (see review, (Rieucau and Giraldeau 2011)), have shown that individuals may rely on 'social information' (i.e. derived from conspecifics) as a 'backup' plan, mostly in situations where personal information is unreliable, out-of-date, or too costly to gather. However, generally, kangaroos might rely more on social rather than personal information under the social context. This result is particularly interesting given that kangaroos are thought to have a fluid social structure with few strong affiliations (Best et al. 2014).

Group differences were important for behavioural variation, and specific social factors also contributed to change in the responses, for example, group size and the presence of specific associates. An increase in group size reduced FID and the probability of hopping away (in the object trial), and increased duration of feeding (also in the object trial); all suggest animals showing increased 'confidence' in larger groups. These results are in lines with findings in other eastern grey kangaroo populations (Carter et al. 2009b; Favreau et al. 2010; Best et al. 2015). They also support the 'dilution effect' theory (Foster and Treherne 1981), which suggests that individuals should gain protection from predators by grouping with conspecifics (Foster 2017). The larger the group size, the smaller the chance of being preyed upon, so there is safety in numbers (Foster 2017). Because individuals may be safer among more conspecifics, they will likely spend less time in anti-predator activities and more time acquiring resources, such as food (Barbosa 2002; Beauchamp 2008).

Eastern grey kangaroos inhabit mostly open grassy areas (Heathcote 1987) and their grouping behaviour may have evolved as an anti-predator defence (Jarman and Coulson 1989;

Banks 2001). Individuals display both social vigilance, in which they monitor the behaviour of group members, and anti-predator vigilance, in which they monitor for potential threats from the environment (Favreau et al. 2015). (Carter et al. 2009b). An increase in group size increases levels of social vigilance (Carter et al. 2009b; Favreau et al. 2010), but reduces anti-predator vigilance (Favreau et al. 2010). When anti-predator vigilance is high, which is normally associated with a sense of potential risk, group members synchronise feeding and vigilance bouts (Favreau et al. 2010). Therefore, eastern grey kangaroos use social information under several contexts and can adjust their behaviour cohesively with other group members.

Specific *group associates* such as those in male courtship and certain mother-offspring associations also affected some responses. Although males were less approachable than females in general, they had shorter FIDs when they were courting a female, likely because they were reluctant to move away and break the bond with the female. Mothers had higher FID when they had a large pouch young, indicating higher wariness, which suggests that kangaroo mothers increase their escape response when their pouch young are bigger. Young kangaroos are particularly susceptible to fox predation around pouch emergence (Banks et al. 2000). Predation risk has been reported to affect the strength of antipredator responses in an African ungulate community (Creel et al. 2014), and predator risk has also been correlated with offspring growth via maternal effects in great tits, which shows that mothers might adjust their influence on offspring in response to a change in the level of predation risk through different mechanisms, e.g., hormonally or behaviourally (Coslovsky and Richner 2011). Mothers were less likely to hop away in response to my novel tests sound if they had a young-at-foot (the category of 'group associate' described above as a mother with its 1-year-old offspring), which suggests that they may avoid separation from the young at this stage as it can result in loss of young (King et al. 2015).

In summary, we have found that individuals were more variable than consistent in behaviour under the social context due to social influences. However, there is still support for

the existence of 'personalities' in our kangaroo study population. In my next chapter, I focus on explaining the causes of among-group behavioural differences and test whether personality is important to group-level outcomes and social organisation (Jolles et al. 2017, 2020). For example, I test whether there a disproportionate influence of certain personalities in the group for group-level responses and whether individuals are grouped based on their personalities.

Sex and age effects

Differences between the sexes were more prominent in FID than for the other traits, whereas age class affected feeding duration (in the Object trial) and dominance interactions. The sex effect observed in the FID trait was possibly associated with a difference in long-term habituation between males and females. As most males migrate away for most of the year, staying in areas with less experience of visitors (Montana et al. 2020), they are less habituated over the long term than females. The sexes of the interacting individuals did not affect the likelihood of winning a dominance encounter, but note that interactions between males comprised 91% of all recorded observations, so male kangaroos are clearly more prone to engage in dominance disputes than females. A recent comprehensive meta-analysis on personality differences between males and females using 2167 effects across 226 species (from invertebrates to mammals) has found little support for widespread sex differences in either personality means or variability (Harrison, Noble, Jennions, manuscript in preparation).

The effects of age on behavioural traits, and dominance, in particular, may be driven in part by body size differences. Age will be correlated with size as an older individual will usually be bigger (kangaroos show indeterminate growth), and we have not corrected for size in this study to isolate the effect of age. However, age itself (after accounting for a size effect) is positively associated with success in dominance interactions in several other species, such as the European vulture guild (Moreno-Opo et al. 2020), free-ranging domestic dogs (Cafazzo et al. 2010), and brown trout (Jacob et al. 2007), and is even more important than size in female beef

cattle (Šárová et al. 2013). Here, older individuals spent less time feeding than 1-year-olds in response to the Object test, which indicates that 1-year-olds were not as vigilant in response to the foreign object or focused more time on food intake required for growth (Toni et al. 2020). We are not aware of other studies testing for age differences in response to novel experimental stimuli in other wild animal taxa. However, in humans, a study tested the personality of individuals from 14 months of age to adulthood (three decades, n: >100 individuals) (Tang et al. 2020). They found larger among-individual differences in behaviour than intra-individual differences, suggesting insufficient age effects: highly inhibited infants were consistently more introverted and less socially functional 25 years later than infants displaying less inhibited behaviours.

Plasticity

As discussed above, eastern grey kangaroos may change their behaviour across social groups (captured by among-group proportions) and according to changes in group size or the presence of particular associates. Several traits also varied between the different years. We have observed strong year effects on juvenile survival (King et al. 2017), which could influence annual social dynamics and, hence, affect behaviour (Chase et al. 2002). Nevertheless, the main effects of plasticity across all traits were associated with habituation, as discussed below.

Habituation

We found evidence for habituation captured through the effects of location (*easting*) and *test number*. The effect of location on FID shows a difference in long-term habituation in individuals who lived in different locations along the west-east axis, in line with our expectations due to their experience of visitors (Supplementary Materials 1, Figure S1.1). The effect of *test number* indicates habituation to the stimuli across time, an effect that was significant across all experimental traits. Individuals showed increased habituation by reducing their FID and

probability of hopping away and by increasing their duration of feeding (hence, reducing duration of vigilance, see Supplementary Material 1, Table S1.3). Similar results have been found, for example, for FID in feral horses (*Equus caballus* L., 1758) (Cabrera et al. 2017). Habituation is a universal non-associative form of learning by which individuals reduce responses to frequent stimuli existing in the surrounding environment to focus normally on stronger stimuli (Rankin et al. 2009; Thompson 2010). The effect of *number of days from the previous test* was significant and opposite to the effect of *test number* in FID and Duration-Feeding-Object, which indicates a recovery in response following a prolonged period of lack of stimulation, which is itself further evidence that habituation occurred (Thompson 2010). Technically speaking, it is generally good practice to consider how close together in time observations are taken, in addition to the individual temporal plasticity, to avoid the effects of temporal autocorrelation on repeatability and plasticity estimates (Mitchell et al. 2020).

Individual differences in habituation

We extended our analyses to test for individual variation in habituation as this could be of evolutionary importance (Nussey et al. 2007; Dingemanse and Wolf 2013; Wheat and Wilmers 2016). We found evidence for individual variation in habituation, which means that some individuals habituated more rapidly than others, which could be a substrate for selection to act on (Scheiner 2002; Dingemanse et al. 2012; Dingemanse and Wolf 2013). We found a negative covariance between individual intercept and slope for test number across all traits, which implies that individuals with larger intercepts had relatively more negative slopes, meaning that more responsive individuals had lower rates of habituation. Therefore, the support for covariance between intercept and slope indicates a link between personality and plasticity (Dingemanse et al. 2010a). The negative covariance possibly also indicates more variation earlier on than at the last tests, so that the individual reaction norms would have a fanning in shape (Pillinger 2013). A study revealed habituation to novel environments in great tits (*Parus major*), with a mean increase

in exploration in repeated tests and support for individual differences in habituation slopes (Dingemanse et al. 2012). They observed a positive covariance between the individual intercept and slope, with faster explorers exploring even faster in the repeated tests (Dingemanse et al. 2012).

Observed traits versus principal components

We compared results from analyses of raw trait values vs of the first principal component from the full set of traits measured in each trial. In general, we found that variance proportions were inconsistent, and interpretation was more difficult for principal components. The inconsistency in principal component variance estimates compared to raw traits was of particular concern because much animal personality research summarises data into principal components (Budaev 2010), and these variance estimates are used to answer critical questions (Chong et al. 2018). Therefore, the use of observed traits was more straightforward, easier to interpret, easily reproducible, and avoided the pitfall of performing 'statistics on statistics'. Furthermore, it helped avoid the use of conglomerate terms, such as 'boldness' and 'exploration', which are likely to cause problems with study comparisons and reproducibility (Beekman and Jordan 2017).

Covariation between multiple traits

We investigated individual-level covariances and found support for covariance between several traits. FID and Hopped-Away traits are types of anti-predator responses (Austin and Ramp 2019b, a). Positive individual-level covariance between these traits indicates that some individuals displayed stronger anti-predator responses than others, regardless of the stimulus. In addition, individuals that display stronger anti-predator responses spend less time foraging in the face of disturbance and vice versa (Favreau et al. 2010; Dannock et al. 2013). Previous studies on eastern grey kangaroos had found an individual-level correlation between FID and vigilance

duration (positive) (Edwards et al. 2013). However, these studies considered adult females only; hence, here, we extended these analyses to all sexes and age groups. Dominance covaried with Duration-Feeding-Object across individuals (positive), which indicates that more dominant individuals were less disturbed by the object and focused more time on resource acquisition. However, Dominance did not covary with any other traits. These results could either have been affected by some challenges during data collection or particularities of kangaroo biology. For example, almost all observations of dominance interactions were among adult males, but we were not able to conduct multiple experimental behavioural measures on all adult males. Also, adult males usually spend most of the time in areas with less tourist (including all the most dominant males in our study), so they were less-long term habituated, which would have an impact on the behavioural estimates and hence, on the covariances among dominance and other behavioural traits. Other studies have found a correlation between dominance and personality, with higher-rank individuals being less reactive under risky situations, i.e., bolder (e.g., male rainbow fish, *Melanotaenia duboulayi* (Colléter and Brown 2011), male Australian field crickets, *Teleogryllus oceanicus* (Rudin et al. 2017), and female zebra finches, *Taeniopygia guttata* (David et al. 2011)).

Finally, we did not find support for individual-level covariance between the Duration of Feeding and Hopped Away traits, but there was support for covariance across groups in these traits within the Object test (Table 3b). We believe that there was no covariance across individuals due to the low individual repeatabilities in comparison with among-group variance. This result emphasises the importance of the social environment in the way individuals respond, indicating that in groups with higher average Feeding Duration, a lower frequency of members Hopped Away.

Conclusion

Our study extends our knowledge of behavioural variation and covariation in a wild social macropod, using large sample sizes, powerful statistical methods, and reliable observations on marked individuals. We present, to our knowledge for the first time, comparisons of the magnitude of individual repeatability versus among-group proportion of variance in behavioural traits in a wild animal population, highlighting the importance of incorporating various aspects of the social environment into experiments investigating behavioural repeatability. Therefore, the use of repeatability data from classic personality experiments, such as FID tests, may exaggerate the importance of animal personality in natural social systems. Nevertheless, the field will benefit from further studies investigating how individual personality might contribute to group differences in behaviour and social organisation – which is addressed in the next chapter.

Acknowledgements

This project was supported by The Holsworth Wildlife Research Endowment & The Ecological Society of Australia between 2017 and 2020. We thank Carolina Dólan (CD) for assistance with fieldwork, especially data collection on Flight Initiation Distance, and all other students and field assistants for conducting dominance observations. We thank Prof Robert Magrath for the resourceful help during the preparations of the playback experiment.

References

Allan ATL, Bailey AL, Hill RA (2020) Habituation is not neutral or equal: Individual differences in tolerance suggest an overlooked personality trait. Sci Adv 6:870–878. https://doi.org/10.1126/sciadv.aaz0870

Aplin LM, Firth JA, Farine DR, et al (2015) Consistent individual differences in the social phenotypes of wild great tits, Parus major. Anim Behav 108:117–127. https://doi.org/10.1016/j.anbehav.2015.07.016

Austin CM, Ramp D (2019a) Flight responses of eastern gray kangaroos to benign or harmful human behavior. Ecol Evol 9:13824–13834. https://doi.org/10.1002/ece3.5818

Austin CM, Ramp D (2019b) Behavioural plasticity by eastern grey kangaroos in response to human behaviour. Animals 9:244–260. https://doi.org/10.3390/ani9050244

Banks PB (2001) Predation-sensitive grouping and habitat use by eastern grey kangaroos: a field experiment. Anim Behav 61:1013–1021. https://doi.org/http://dx.doi.org/10.1006/anbe.2001.1686

Banks PB, Newsome AE, Dickman CR (2000) Predation by red foxes limits recruitment in populations of eastern grey kangaroos. Austral Ecol 25:283–291. https://doi.org/10.1046/j.1442-9993.2000.01039.x

Barbosa A (2002) Does vigilance always covary negatively with group size? Effects of foraging strategy. Acta Ethol 5:51–55. https://doi.org/10.1007/s10211-002-0059-0

Beauchamp G (2008) What is the magnitude of the group-size effect on vigilance? Behav. Ecol. 19:1361–1368

Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? Behav Ecol 28:617–623. https://doi.org/10.1093/beheco/arx022

Bell AM (2017) There is no special sauce: a comment on Beekman and Jordan. Behav Ecol. https://doi.org/10.1093/beheco/arx031

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Bergmüller R, Taborsky M, Bergmü Ller R, et al (2010) Animal personality due to social niche specialisation. Trends Ecol Evol 25:504–511. https://doi.org/10.1016/j.tree.2010.06.012

Best EC, Blomberg SP, Goldizen AW (2015) Shy female kangaroos seek safety in numbers and have fewer preferred friendships. Behav Ecol 26:639–646. https://doi.org/10.1093/beheco/arv003

Best EC, Dwyer RG, Seddon JM, Goldizen AW (2014) Associations are more strongly correlated with space use than kinship in female eastern grey kangaroos. Anim Behav 89:1–10. https://doi.org/10.1016/j.anbehav.2013.12.011

Blumstein DT, Anthony LL, Harcourt R, Ross G (2003) Testing a key assumption of wildlife buffer zones: is flight initiation distance a species-specific trait? Biol Conserv 110:97–100. https://doi.org/10.1016/S0006-3207(02)00180-5

Blumstein DT, Flores G, Munoz NE (2015) Does locomotor ability influence flight initiation distance in yellow-bellied marmots? Ethology 121:434–441. https://doi.org/10.1111/eth.12351

Briffa M (2017) Abandoning animal personality would cause obfuscation: A comment on Beekman and Jordan. Behav Ecol 28:625–626. https://doi.org/10.1093/beheco/arx025

Budaev S V. (2010) Using principal components and factor analysis in animal behaviour research: caveats and guidelines. Ethology 116:472–480. https://doi.org/10.1111/j.1439-0310.2010.01758.x

Bunford N, Csibra B, Gácsi M (2019) Individual differences in response to ambiguous stimuli in a modified go/no-go paradigm are associated with personality in family dogs. Sci Rep 9:. https://doi.org/10.1038/s41598-019-47510-z

Cabrera D, Andres D, McLoughlin PD, et al (2017) Island tameness and the repeatability of flight initiation distance in a large herbivore. Can J Zool 95:771–778. https://doi.org/10.1139/cjz-2016-0305

Cafazzo S, Valsecchi P, Bonanni R, Natoli E (2010) Dominance in relation to age, sex, and competitive contexts in a group of free-ranging domestic dogs. Behav Ecol 21:443–455. https://doi.org/10.1093/beheco/arq001

Carrete M, Tella JL (2010) Individual consistency in flight initiation distances in burrowing owls: A new hypothesis on disturbanceinduced habitat sele. Biol Lett 6:167–170. https://doi.org/10.1098/rsbl.2009.0739

Carter AJ, Feeney WE, Marshall HH, et al (2013) Animal personality: What are behavioural ecologists measuring? Biol Rev 88:465–475. https://doi.org/10.1111/brv.12007

Carter AJ, Goldizen AW, Tromp SA (2010) Agamas exhibit behavioral syndromes: Bolder males bask and feed more but may suffer higher predation. Behav Ecol 21:655–661. https://doi.org/10.1093/beheco/arq036

Carter AJ, Macdonald SL, Thomson VA, Goldizen AW (2009a) Structured association patterns and their energetic benefits in female eastern grey kangaroos, *Macropus giganteus*. Anim Behav 77:839–846. https://doi.org/10.1016/j.anbehav.2008.12.007

Carter AJ, Pays O, Goldizen AW (2009b) Individual variation in the relationship between vigilance and group size in eastern grey kangaroos. Behav Ecol Sociobiol 64:237–245. https://doi.org/10.1007/s00265-009-0840-4

Chase ID, Tovey C, Spangler-Martin D, Manfredonia M (2002) Individual differences versus social dynamics in the formation of animal dominance hierarchies. Proc Natl Acad Sci U S A 99:5744–5749. https://doi.org/10.1073/pnas.082104199

Chong VK, Fung HF, Stinchcombe JR (2018) A note on measuring natural selection on principal component scores. Evol Lett. https://doi.org/10.1002/evl3.63

Clary D, Skyner LJ, Ryan CP, et al (2014) Shyness-boldness, but not exploration, predicts glucocorticoid stress response in richardson's ground squirrels (*Urocitellus richardsonii*). Ethology 120:1101–1109. https://doi.org/10.1111/eth.12283

Colléter M, Brown C (2011) Personality traits predict hierarchy rank in male rainbowfish social groups. Anim Behav 81:1231–1237. https://doi.org/10.1016/j.anbehav.2011.03.011

Conradt L, Roper TJ (2005) Consensus decision making in animals. Trends Ecol. Evol. 20:449–456

Coslovsky M, Richner H (2011) Predation risk affects offspring growth via maternal effects. Funct Ecol 25:878–888. https://doi.org/10.1111/j.1365-2435.2011.01834.x

Coulson G (1997) Repertoires of social behaviour in captive and free-ranging grey kangaroos, *Macropus giganteus* and *Macropus fuliginosus* (Marsupialia: Macropodidae). J Zool 242:119–130. https://doi.org/10.1111/j.1469-7998.1997.tb02933.x

Creel S, Schuette P, Christianson D (2014) Effects of predation risk on group size, vigilance, and foraging behavior in an African ungulate community. Behav Ecol 25:773–784. https://doi.org/10.1093/beheco/aru050

Dall SRX, Houston AI, McNamara JM (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. Ecol Lett 7:734–739. https://doi.org/10.1111/j.1461-0248.2004.00618.x

Dannock RJ, Blomberg SP, Goldizen AW (2013) Individual variation in vigilance in female eastern grey kangaroos. Aust J Zool 61:312–319. https://doi.org/10.1071/ZO12122

David M, Auclair Y, Cézilly F (2011) Personality predicts social dominance in female zebra finches, Taeniopygia guttata, in a feeding context. Anim Behav 81:219–224. https://doi.org/10.1016/j.anbehav.2010.10.008

de Villemereuil P, Morrissey MB, Nakagawa S, Schielzeth H (2018) Fixed-effect variance and the estimation of repeatabilities and heritabilities: issues and solutions. J Evol Biol 31:621–632. https://doi.org/10.1111/jeb.13232

Dingemanse NJ (2017) The role of personality research in contemporary behavioral ecology: A comment on Beekman and Jordan. Behav Ecol 28:624–625. https://doi.org/10.1093/beheco/arx027

Dingemanse NJ, Bouwman KM, van de Pol M, et al (2012) Variation in personality and behavioural plasticity across four populations of the great tit Parus major. J Anim Ecol 81:116–126. https://doi.org/10.1111/j.1365-2656.2011.01877.x

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010a) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Dingemanse NJ, Kazem AJNN, Ré Ale D, et al (2010b) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Dingemanse NJ, Réale D (2005) Natural selection and animal personality. Behaviour 142:1159–1184. https://doi.org/10.1163/156853905774539445

Dingemanse NJ, Wolf M (2013) Between-individual differences in behavioural plasticity within populations: Causes and consequences. Anim Behav 85:1031–1039. https://doi.org/10.1016/j.anbehav.2012.12.032

East ML, Hofer H (2010) Social environments, social tactics and their fitness consequences in complex mammalian societies. In: Social Behaviour: Genes, Ecology and Evolution. Cambridge University Press, pp 360–390

Edwards AM, Best EC, Blomberg SP, Goldizen AW (2013) Individual traits influence vigilance in wild female eastern grey kangaroos. Aust J Zool 61:332–341. https://doi.org/10.1071/ZO13025

Favreau F-R, Goldizen AW, Pays O (2010) Interactions among social monitoring, anti-predator vigilance and group size in eastern grey kangaroos. Proc R Soc B Biol Sci 2089–2095. https://doi.org/10.1098/rspb.2009.2337

Favreau F-R, Pays O, Fritz H, et al (2015) Predators, food and social context shape the types of vigilance exhibited by kangaroos. Anim Behav 99:109–121. https://doi.org/10.1016/j.anbehav.2014.11.001

Favreau FR, Goldizen AW, Fritz H, Pays O (2018) Food supply fluctuations constrain group sizes of kangaroos and in turn shape their vigilance and feeding strategies. Anim Behav 135:165–176. https://doi.org/10.1016/j.anbehav.2017.11.020

Fögen T (2014) Animal Communication. Oxford University Press

Foster W (2017) Safety in numbers. J Exp Biol 220:4551–4553. https://doi.org/10.1242/jeb.160721

Foster WA, Treherne JE (1981) Evidence for the dilution effect in the selfish herd from fish predation on a marine insect

Found R, St Clair CC (2016) Behavioural syndromes predict loss of migration in wild elk. Anim Behav 115:35–46. https://doi.org/10.1016/j.anbehav.2016.02.007

Garamszegi LZ, Møller AP (2017) Partitioning within-species variance in behaviour to within- and between-population components for understanding evolution. Ecol Lett 20:599–608. https://doi.org/10.1111/ele.12758

Gélin U, Coulson G, Festa-Bianchet M (2015) Heterogeneity in reproductive success explained by individual differences in bite rate and mass change. Behav Ecol 27:1–7. https://doi.org/10.1093/beheco/arv209

Guay P-J, McLeod EM, Cross R, et al (2013) Observer effects occur when estimating alert but not flight-initiation distances. Wildl Res 40:289. https://doi.org/10.1071/WR13013

Hadfield J (2018) MCMCglmm Course Notes

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. J Stat Softw 33:1–22

Heathcote CF (1987) Grouping of eastern grey kangaroos in open habitat. Aust Wildl Res 14:343–8

Jacob A, Nusslé S, Britschgi A, et al (2007) Male dominance linked to size and age, but not to "good genes" in brown trout (Salmo trutta). BMC Evol Biol 7:207. https://doi.org/10.1186/1471-2148-7-207

Jarman PJ (1987) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Jarman PJ, Coulson G (1989) Dynamics and adaptiveness of grouping in macropods. In: Grigg G, Jarman P, Hume I (eds) Kangaroos, wallabies and rat-kangaroos. Surrey Beatty & Sons Pty Ltd, New South Wales, pp 527–547

Jolles JW, Boogert NJ, Sridhar VH, et al (2017) Consistent individual differences drive collective behaviour and group functioning of schooling fish. Curr Biol 27:2862-2868.e7. https://doi.org/10.1016/j.cub.2017.08.004

Jolles JW, King AJ, Killen SS (2020) The role of individual heterogeneity in collective animal behaviour. Trends Ecol Evol 35:278–291. https://doi.org/10.1016/j.tree.2019.11.001

King WJ, Festa-Bianchet M, Coulson G, Goldizen AW (2017) Long-term consequences of mother-offspring associations in eastern grey kangaroos. Behav Ecol Sociobiol 71:77. https://doi.org/10.1007/s00265-017-2297-1

King WJ, Forsyth DM, Coulson G, Festa-Bianchet M (2015) Adoption in eastern grey kangaroos: a consequence of misdirected care? PLoS One 10:e0125182. https://doi.org/10.1371/journal.pone.0125182

King WJ, Wilson ME, Allen T, et al (2011) A capture technique for free-ranging eastern grey kangaroos (*Macropus giganteus*) habituated to humans. Aust Mammal 33:47–51. https://doi.org/10.1071/AM10029

Mackay AE, Forsyth DM, Coulson G, Festa-Bianchet M (2018) Maternal resource allocation adjusts to timing of parturition in an asynchronous breeder. Behav Ecol Sociobiol 72:7. https://doi.org/10.1007/s00265-017-2419-9

Menz CS, Goldizen AW, Blomberg SP, et al (2017) Understanding repeatability and plasticity in multiple dimensions of the sociability of wild female kangaroos. https://doi.org/10.1016/j.anbehav.2017.01.012

Michelangeli M, Chapple DG, Wong BBM (2016) Are behavioural syndromes sex specific? Personality in a widespread lizard species. Behav Ecol Sociobiol 70:1911–1919. https://doi.org/10.1007/s00265-016-2197-9

Mitchell DJ, Dujon AM, Beckmann C, Biro PA (2020) Temporal autocorrelation: A neglected factor in the study of behavioral repeatability and plasticity. Behav Ecol 31:222–231. https://doi.org/10.1093/beheco/arz180

Møller AP (2008) Flight distance and blood parasites in birds. Behav Ecol 19:1305–1313. https://doi.org/10.1093/beheco/arn074

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Montiglio PO, Ferrari C, Réale D (2013) Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. Philos. Trans. R. Soc. B Biol. Sci. 368

Moreno-Opo R, Trujillano A, Margalida A (2020) Larger size and older age confer competitive advantage: dominance hierarchy within European vulture guild. Sci Rep 10:1–12. https://doi.org/10.1038/s41598-020-59387-4

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:no-no. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Nussey DH, Wilson AJ, Brommer JE (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. J Evol Biol 20:831–844. https://doi.org/10.1111/j.1420-9101.2007.01300.x

Patrick SC, Weimerskirch H (2014) Personality, foraging and fitness consequences in a long lived seabird. PLoS One 9:e87269. https://doi.org/10.1371/journal.pone.0087269

Petelle MB, Martin JGA, Blumstein DT (2015) Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (*Marmota flaviventris*). J Evol Biol 28:1840–1848. https://doi.org/10.1111/jeb.12700

Pillinger R (2013) Random slope models. In: Cent. Multilevel Model. | Univ. Bristol. http://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html#covar. Accessed 14 Nov 2020

Planas-Sitjà I, Nicolis SC, Sempo G, Deneubourg J-L (2018) The interplay between personalities and social interactions affects the cohesion of the group and the speed of aggregation. PLoS One 13:e0201053. https://doi.org/10.1371/journal.pone.0201053

Pollak E (2001) Quantitative Trait. In: Encyclopedia of Genetics. pp 1597–1598

Rankin CH, Abrams T, Barry RJ, et al (2009) Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92:135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Réale D, Dingemanse NJ (2010) Personality and individual social specialisation. In: Szekely T, Moore AJ, Komdeur J (eds) Social Behaviour: genes, ecology and evolution, 1st edn. Cambridge University Press, Cambridge, pp 417–441

Rieucau G, Giraldeau LA (2011) Exploring the costs and benefits of social information use: An appraisal of current experimental evidence. Philos. Trans. R. Soc. B Biol. Sci. 366:949–957

Rudin FS, Tomkins JL, Simmons LW (2017) Changes in dominance status erode personality and behavioral syndromes. Behav Ecol 28:270–279. https://doi.org/10.1093/beheco/arw151

Šárová R, Špinka M, Stěhulová I, et al (2013) Pay respect to the elders: Age, more than body mass, determines dominance in female beef cattle. Anim Behav 86:1315–1323. https://doi.org/10.1016/j.anbehav.2013.10.002

Sauter R, Held L (2016) Quasi-complete separation in random effects of binary response mixed models. J Stat Comput Simul 86:2781–2796. https://doi.org/10.1080/00949655.2015.1129539

Scheiner SM (2002) Selection experiments and the study of phenotypic plasticity. J Evol Biol 15:889–898. https://doi.org/10.1046/j.1420-9101.2002.00468.x

Sih A (2017) Insights for behavioral ecology from behavioral syndromes: a comment on Beekman and Jordan. Behav Ecol 28:627–628. https://doi.org/10.1093/beheco/arx032

Sih A, Bell AM, Johnson JC, Ziemba RE (2004) Behavioral Syndromes: An Integrative Overview. Q Rev Biol 79:241–277. https://doi.org/10.1086/422893

St-Hilaire É, Réale D, Garant D (2017) Determinants, selection and heritability of docility in wild eastern chipmunks (*Tamias striatus*). Behav Ecol Sociobiol 71:88. https://doi.org/10.1007/s00265-017-2320-6

Stamps J, Groothuis TGG (2010) The development of animal personality: Relevance, concepts and perspectives. Biol Rev 85:301–325. https://doi.org/10.1111/j.1469-185X.2009.00103.x

Tang A, Crawford H, Morales S, et al (2020) Infant behavioral inhibition predicts personality and social outcomes three decades later. Proc Natl Acad Sci U S A 117:9800–9807. https://doi.org/10.1073/pnas.1917376117

Thompson RF (2010) Habituation. In: Encyclopedia of Behavioral Neuroscience. Elsevier Inc., pp 1–5

Toni P, Forsyth DM, Festa-Bianchet M (2020) Forage availability and maternal characteristics affect costs of reproduction in a large marsupial. Oecologia 193:97–107. https://doi.org/10.1007/s00442-020-04653-5

Urszán TJ, Garamszegi LZ, Nagy G, et al (2018) Experience during development triggers between-individual variation in behavioural plasticity. J Anim Ecol 87:1264–1273. https://doi.org/10.1111/1365-2656.12847

Villemereuil P de, Schielzeth H, Nakagawa S, Morrissey M (2016) General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics 204:1281–1294. https://doi.org/10.1534/genetics.115.186536

Weston MA, Mcleod EM, Blumstein DT, Guay PJ (2012) A review of flight-initiation distances and their application to managing disturbance to Australian birds. Emu - Austral Ornithol. 112:269–286

Wheat RE, Wilmers CC (2016) Habituation reverses fear-based ecological effects in brown bears (*Ursus arctos*). Ecosphere 7:. https://doi.org/10.1002/ecs2.1408

White SJ, Pascall DJ, Wilson AJ (2020) Towards a comparative approach to the structure of animal personality variation. Behav Ecol 31:340–351. https://doi.org/10.1093/beheco/arz198

Wilson AJ (2008) Why h2 does not always equal VA/VP? J. Evol. Biol. 21:647–650

Wilson AJ, Morrisey MB, Adams MJ, et al (2011) Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, <i>Cervus elaphus<i>. J Evol Biol 24:772–783. https://doi.org/10.1111/j.1420-9101.2010.02212.x

Wolf M, McNamara JM (2013) Adaptive between-individual differences in social competence. Trends Ecol. Evol. 28:253–254

Allan ATL, Bailey AL, Hill RA (2020) Habituation is not neutral or equal: Individual differences in tolerance suggest an overlooked personality trait. Sci Adv 6:870–878. https://doi.org/10.1126/sciadv.aaz0870

Aplin LM, Firth JA, Farine DR, et al (2015) Consistent individual differences in the social phenotypes of wild great tits, Parus major. Anim Behav 108:117–127. https://doi.org/10.1016/j.anbehav.2015.07.016

Austin CM, Ramp D (2019a) Flight responses of eastern gray kangaroos to benign or harmful human behavior. Ecol Evol 9:13824–13834. https://doi.org/10.1002/ece3.5818

Austin CM, Ramp D (2019b) Behavioural plasticity by eastern grey kangaroos in response to human behaviour. Animals 9:244–260. https://doi.org/10.3390/ani9050244

Banks PB (2001) Predation-sensitive grouping and habitat use by eastern grey kangaroos: a field experiment. Anim Behav 61:1013–1021. https://doi.org/http://dx.doi.org/10.1006/anbe.2001.1686

Banks PB, Newsome AE, Dickman CR (2000) Predation by red foxes limits recruitment in populations of eastern grey kangaroos. Austral Ecol 25:283–291. https://doi.org/10.1046/j.1442-9993.2000.01039.x

Barbosa A (2002) Does vigilance always covary negatively with group size? Effects of foraging strategy. Acta Ethol 5:51–55. https://doi.org/10.1007/s10211-002-0059-0

Beauchamp G (2008) What is the magnitude of the group-size effect on vigilance? Behav. Ecol. 19:1361–1368

Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? Behav Ecol 28:617–623. https://doi.org/10.1093/beheco/arx022

Bell AM (2017) There is no special sauce: a comment on Beekman and Jordan. Behav Ecol. https://doi.org/10.1093/beheco/arx031

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Bergmüller R, Taborsky M, Bergmü Ller R, et al (2010) Animal personality due to social niche specialisation. Trends Ecol Evol 25:504–511. https://doi.org/10.1016/j.tree.2010.06.012

Best EC, Blomberg SP, Goldizen AW (2015) Shy female kangaroos seek safety in numbers and have fewer preferred friendships. Behav Ecol 26:639–646. https://doi.org/10.1093/beheco/arv003

Best EC, Dwyer RG, Seddon JM, Goldizen AW (2014) Associations are more strongly correlated with space use than kinship in female eastern grey kangaroos. Anim Behav 89:1–10. https://doi.org/10.1016/j.anbehav.2013.12.011

Blumstein DT, Anthony LL, Harcourt R, Ross G (2003) Testing a key assumption of wildlife buffer zones: is flight initiation distance a species-specific trait? Biol Conserv 110:97–100. https://doi.org/10.1016/S0006-3207(02)00180-5

Blumstein DT, Flores G, Munoz NE (2015) Does locomotor ability influence flight initiation distance in yellow-bellied marmots? Ethology 121:434–441. https://doi.org/10.1111/eth.12351

Briffa M (2017) Abandoning animal personality would cause obfuscation: A comment on Beekman and Jordan. Behav Ecol 28:625–626. https://doi.org/10.1093/beheco/arx025

Budaev S V. (2010) Using principal components and factor analysis in animal behaviour research: caveats and guidelines. Ethology 116:472–480. https://doi.org/10.1111/j.1439-0310.2010.01758.x

Bunford N, Csibra B, Gácsi M (2019) Individual differences in response to ambiguous stimuli in a modified go/no-go paradigm are associated with personality in family dogs. Sci Rep 9:. https://doi.org/10.1038/s41598-019-47510-z

Cabrera D, Andres D, McLoughlin PD, et al (2017) Island tameness and the repeatability of flight initiation distance in a large herbivore. Can J Zool 95:771–778. https://doi.org/10.1139/cjz-2016-0305

Cafazzo S, Valsecchi P, Bonanni R, Natoli E (2010) Dominance in relation to age, sex, and competitive contexts in a group of free-ranging domestic dogs. Behav Ecol 21:443–455. https://doi.org/10.1093/beheco/arq001

Carrete M, Tella JL (2010) Individual consistency in flight initiation distances in burrowing owls: A new hypothesis on disturbanceinduced habitat sele. Biol Lett 6:167–170. https://doi.org/10.1098/rsbl.2009.0739

Carter AJ, Feeney WE, Marshall HH, et al (2013) Animal personality: What are behavioural ecologists measuring? Biol Rev 88:465–475. https://doi.org/10.1111/brv.12007

Carter AJ, Goldizen AW, Tromp SA (2010) Agamas exhibit behavioral syndromes: Bolder males bask and feed more but may suffer higher predation. Behav Ecol 21:655–661. https://doi.org/10.1093/beheco/arq036

Carter AJ, Macdonald SL, Thomson VA, Goldizen AW (2009a) Structured association patterns and their energetic benefits in female eastern grey kangaroos, *Macropus giganteus*. Anim Behav 77:839–846. https://doi.org/10.1016/j.anbehav.2008.12.007

Carter AJ, Pays O, Goldizen AW (2009b) Individual variation in the relationship between vigilance and group size in eastern grey kangaroos. Behav Ecol Sociobiol 64:237–245. https://doi.org/10.1007/s00265-009-0840-4

Chase ID, Tovey C, Spangler-Martin D, Manfredonia M (2002) Individual differences versus social dynamics in the formation of animal dominance hierarchies. Proc Natl Acad Sci U S A 99:5744–5749. https://doi.org/10.1073/pnas.082104199

Chong VK, Fung HF, Stinchcombe JR (2018) A note on measuring natural selection on principal component scores. Evol Lett. https://doi.org/10.1002/evl3.63

Clary D, Skyner LJ, Ryan CP, et al (2014) Shyness-boldness, but not exploration, predicts glucocorticoid stress response in richardson's ground squirrels (*Urocitellus richardsonii*). Ethology 120:1101–1109. https://doi.org/10.1111/eth.12283

Colléter M, Brown C (2011) Personality traits predict hierarchy rank in male rainbowfish social groups. Anim Behav 81:1231–1237. https://doi.org/10.1016/j.anbehav.2011.03.011

Conradt L, Roper TJ (2005) Consensus decision making in animals. Trends Ecol. Evol. 20:449–456

Coslovsky M, Richner H (2011) Predation risk affects offspring growth via maternal effects. Funct Ecol 25:878–888. https://doi.org/10.1111/j.1365-2435.2011.01834.x

Coulson G (1997) Repertoires of social behaviour in captive and free-ranging grey kangaroos, *Macropus giganteus* and *Macropus fuliginosus* (Marsupialia: Macropodidae). J Zool 242:119–130. https://doi.org/10.1111/j.1469-7998.1997.tb02933.x

Creel S, Schuette P, Christianson D (2014) Effects of predation risk on group size, vigilance, and foraging behavior in an African ungulate community. Behav Ecol 25:773–784. https://doi.org/10.1093/beheco/aru050

Dall SRX, Houston AI, McNamara JM (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. Ecol Lett 7:734–739. https://doi.org/10.1111/j.1461-0248.2004.00618.x

Dannock RJ, Blomberg SP, Goldizen AW (2013) Individual variation in vigilance in female eastern grey kangaroos. Aust J Zool 61:312–319. https://doi.org/10.1071/ZO12122

David M, Auclair Y, Cézilly F (2011) Personality predicts social dominance in female zebra finches, Taeniopygia guttata, in a feeding context. Anim Behav 81:219–224. https://doi.org/10.1016/j.anbehav.2010.10.008

de Villemereuil P, Morrissey MB, Nakagawa S, Schielzeth H (2018) Fixed-effect variance and the estimation of repeatabilities and heritabilities: issues and solutions. J Evol Biol 31:621–632. https://doi.org/10.1111/jeb.13232

Dingemanse NJ (2017) The role of personality research in contemporary behavioral ecology: A comment on Beekman and Jordan. Behav Ecol 28:624–625. https://doi.org/10.1093/beheco/arx027

Dingemanse NJ, Bouwman KM, van de Pol M, et al (2012) Variation in personality and behavioural plasticity across four populations of the great tit Parus major. J Anim Ecol 81:116–126. https://doi.org/10.1111/j.1365-2656.2011.01877.x

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010a) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Dingemanse NJ, Kazem AJNN, Ré Ale D, et al (2010b) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Dingemanse NJ, Réale D (2005) Natural selection and animal personality. Behaviour 142:1159–1184. https://doi.org/10.1163/156853905774539445

Dingemanse NJ, Wolf M (2013) Between-individual differences in behavioural plasticity within populations: Causes and consequences. Anim Behav 85:1031–1039. https://doi.org/10.1016/j.anbehav.2012.12.032

East ML, Hofer H (2010) Social environments, social tactics and their fitness consequences in complex mammalian societies. In: Social Behaviour: Genes, Ecology and Evolution. Cambridge University Press, pp 360–390

Edwards AM, Best EC, Blomberg SP, Goldizen AW (2013) Individual traits influence vigilance in wild female eastern grey kangaroos. Aust J Zool 61:332–341. https://doi.org/10.1071/ZO13025

Favreau F-R, Goldizen AW, Pays O (2010) Interactions among social monitoring, anti-predator vigilance and group size in eastern grey kangaroos. Proc R Soc B Biol Sci 2089–2095. https://doi.org/10.1098/rspb.2009.2337

Favreau F-R, Pays O, Fritz H, et al (2015) Predators, food and social context shape the types of vigilance exhibited by kangaroos. Anim Behav 99:109–121. https://doi.org/10.1016/j.anbehav.2014.11.001

Favreau FR, Goldizen AW, Fritz H, Pays O (2018) Food supply fluctuations constrain group sizes of kangaroos and in turn shape their vigilance and feeding strategies. Anim Behav 135:165–176. https://doi.org/10.1016/j.anbehav.2017.11.020

Fögen T (2014) Animal Communication. Oxford University Press

Foster W (2017) Safety in numbers. J Exp Biol 220:4551–4553. https://doi.org/10.1242/jeb.160721

Foster WA, Treherne JE (1981) Evidence for the dilution effect in the selfish herd from fish predation on a marine insect

Found R, St Clair CC (2016) Behavioural syndromes predict loss of migration in wild elk. Anim Behav 115:35–46. https://doi.org/10.1016/j.anbehav.2016.02.007

Garamszegi LZ, Møller AP (2017) Partitioning within-species variance in behaviour to within- and between-population components for understanding evolution. Ecol Lett 20:599–608. https://doi.org/10.1111/ele.12758

Gélin U, Coulson G, Festa-Bianchet M (2015) Heterogeneity in reproductive success explained by individual differences in bite rate and mass change. Behav Ecol 27:1–7. https://doi.org/10.1093/beheco/arv209

Guay P-J, McLeod EM, Cross R, et al (2013) Observer effects occur when estimating alert but not flight-initiation distances. Wildl Res 40:289. https://doi.org/10.1071/WR13013

Hadfield J (2018) MCMCglmm Course Notes

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. J Stat Softw 33:1–22

Heathcote CF (1987) Grouping of eastern grey kangaroos in open habitat. Aust Wildl Res 14:343–8

Jacob A, Nusslé S, Britschgi A, et al (2007) Male dominance linked to size and age, but not to "good genes" in brown trout (Salmo trutta). BMC Evol Biol 7:207. https://doi.org/10.1186/1471-2148-7-207

Jarman PJ (1987) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Jarman PJ, Coulson G (1989) Dynamics and adaptiveness of grouping in macropods. In: Grigg G, Jarman P, Hume I (eds) Kangaroos, wallabies and rat-kangaroos. Surrey Beatty & Sons Pty Ltd, New South Wales, pp 527–547

Jolles JW, Boogert NJ, Sridhar VH, et al (2017) Consistent individual differences drive collective behaviour and group functioning of schooling fish. Curr Biol 27:2862-2868.e7. https://doi.org/10.1016/j.cub.2017.08.004

Jolles JW, King AJ, Killen SS (2020) The role of individual heterogeneity in collective animal behaviour. Trends Ecol Evol 35:278–291. https://doi.org/10.1016/j.tree.2019.11.001

King WJ, Festa-Bianchet M, Coulson G, Goldizen AW (2017) Long-term consequences of mother-offspring associations in eastern grey kangaroos. Behav Ecol Sociobiol 71:77. https://doi.org/10.1007/s00265-017-2297-1

King WJ, Forsyth DM, Coulson G, Festa-Bianchet M (2015) Adoption in eastern grey kangaroos: a consequence of misdirected care? PLoS One 10:e0125182. https://doi.org/10.1371/journal.pone.0125182

King WJ, Wilson ME, Allen T, et al (2011) A capture technique for free-ranging eastern grey kangaroos (*Macropus giganteus*) habituated to humans. Aust Mammal 33:47–51. https://doi.org/10.1071/AM10029

Mackay AE, Forsyth DM, Coulson G, Festa-Bianchet M (2018) Maternal resource allocation adjusts to timing of parturition in an asynchronous breeder. Behav Ecol Sociobiol 72:7. https://doi.org/10.1007/s00265-017-2419-9

Menz CS, Goldizen AW, Blomberg SP, et al (2017) Understanding repeatability and plasticity in multiple dimensions of the sociability of wild female kangaroos. https://doi.org/10.1016/j.anbehav.2017.01.012

Michelangeli M, Chapple DG, Wong BBM (2016) Are behavioural syndromes sex specific? Personality in a widespread lizard species. Behav Ecol Sociobiol 70:1911–1919. https://doi.org/10.1007/s00265-016-2197-9

Mitchell DJ, Dujon AM, Beckmann C, Biro PA (2020) Temporal autocorrelation: A neglected factor in the study of behavioral repeatability and plasticity. Behav Ecol 31:222–231. https://doi.org/10.1093/beheco/arz180

Møller AP (2008) Flight distance and blood parasites in birds. Behav Ecol 19:1305–1313. https://doi.org/10.1093/beheco/arn074

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Montiglio PO, Ferrari C, Réale D (2013) Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. Philos. Trans. R. Soc. B Biol. Sci. 368

Moreno-Opo R, Trujillano A, Margalida A (2020) Larger size and older age confer competitive advantage: dominance hierarchy within European vulture guild. Sci Rep 10:1–12. https://doi.org/10.1038/s41598-020-59387-4

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:no-no. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Nussey DH, Wilson AJ, Brommer JE (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. J Evol Biol 20:831–844. https://doi.org/10.1111/j.1420-9101.2007.01300.x

Patrick SC, Weimerskirch H (2014) Personality, foraging and fitness consequences in a long lived seabird. PLoS One 9:e87269. https://doi.org/10.1371/journal.pone.0087269

Petelle MB, Martin JGA, Blumstein DT (2015) Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (*Marmota flaviventris*). J Evol Biol 28:1840–1848. https://doi.org/10.1111/jeb.12700

Pillinger R (2013) Random slope models. In: Cent. Multilevel Model. | Univ. Bristol. http://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html#covar. Accessed 14 Nov 2020

Planas-Sitjà I, Nicolis SC, Sempo G, Deneubourg J-L (2018) The interplay between personalities and social interactions affects the cohesion of the group and the speed of aggregation. PLoS One 13:e0201053. https://doi.org/10.1371/journal.pone.0201053

Pollak E (2001) Quantitative Trait. In: Encyclopedia of Genetics. pp 1597–1598

Rankin CH, Abrams T, Barry RJ, et al (2009) Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92:135–138. https://doi.org/10.1016/j.nlm.2008.09.012

Réale D, Dingemanse NJ (2010) Personality and individual social specialisation. In: Szekely T, Moore AJ, Komdeur J (eds) Social Behaviour: genes, ecology and evolution, 1st edn. Cambridge University Press, Cambridge, pp 417–441

Rieucau G, Giraldeau LA (2011) Exploring the costs and benefits of social information use: An appraisal of current experimental evidence. Philos. Trans. R. Soc. B Biol. Sci. 366:949–957

Rudin FS, Tomkins JL, Simmons LW (2017) Changes in dominance status erode personality and behavioral syndromes. Behav Ecol 28:270–279. https://doi.org/10.1093/beheco/arw151

Šárová R, Špinka M, Stěhulová I, et al (2013) Pay respect to the elders: Age, more than body mass, determines dominance in female beef cattle. Anim Behav 86:1315–1323. https://doi.org/10.1016/j.anbehav.2013.10.002

Sauter R, Held L (2016) Quasi-complete separation in random effects of binary response mixed models. J Stat Comput Simul 86:2781–2796. https://doi.org/10.1080/00949655.2015.1129539

Scheiner SM (2002) Selection experiments and the study of phenotypic plasticity. J Evol Biol 15:889–898. https://doi.org/10.1046/j.1420-9101.2002.00468.x

Sih A (2017) Insights for behavioral ecology from behavioral syndromes: a comment on Beekman and Jordan. Behav Ecol 28:627–628. https://doi.org/10.1093/beheco/arx032

Sih A, Bell AM, Johnson JC, Ziemba RE (2004) Behavioral Syndromes: An Integrative Overview. Q Rev Biol 79:241–277. https://doi.org/10.1086/422893

St-Hilaire É, Réale D, Garant D (2017) Determinants, selection and heritability of docility in wild eastern chipmunks (*Tamias striatus*). Behav Ecol Sociobiol 71:88. https://doi.org/10.1007/s00265-017-2320-6

Stamps J, Groothuis TGG (2010) The development of animal personality: Relevance, concepts and perspectives. Biol Rev 85:301–325. https://doi.org/10.1111/j.1469-185X.2009.00103.x

Tang A, Crawford H, Morales S, et al (2020) Infant behavioral inhibition predicts personality and social outcomes three decades later. Proc Natl Acad Sci U S A 117:9800–9807. https://doi.org/10.1073/pnas.1917376117

Thompson RF (2010) Habituation. In: Encyclopedia of Behavioral Neuroscience. Elsevier Inc., pp 1–5

Toni P, Forsyth DM, Festa-Bianchet M (2020) Forage availability and maternal characteristics affect costs of reproduction in a large marsupial. Oecologia 193:97–107. https://doi.org/10.1007/s00442-020-04653-5

Urszán TJ, Garamszegi LZ, Nagy G, et al (2018) Experience during development triggers between-individual variation in behavioural plasticity. J Anim Ecol 87:1264–1273. https://doi.org/10.1111/1365-2656.12847

Villemereuil P de, Schielzeth H, Nakagawa S, Morrissey M (2016) General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics 204:1281–1294. https://doi.org/10.1534/genetics.115.186536

Weston MA, Mcleod EM, Blumstein DT, Guay PJ (2012) A review of flight-initiation distances and their application to managing disturbance to Australian birds. Emu - Austral Ornithol. 112:269–286

Wheat RE, Wilmers CC (2016) Habituation reverses fear-based ecological effects in brown bears (*Ursus arctos*). Ecosphere 7:. https://doi.org/10.1002/ecs2.1408

White SJ, Pascall DJ, Wilson AJ (2020) Towards a comparative approach to the structure of animal personality variation. Behav Ecol 31:340–351. https://doi.org/10.1093/beheco/arz198

Wilson AJ (2008) Why h2 does not always equal VA/VP? J. Evol. Biol. 21:647–650

Wilson AJ, Morrisey MB, Adams MJ, et al (2011) Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, <i>Cervus elaphus<i>. J Evol Biol 24:772–783. https://doi.org/10.1111/j.1420-9101.2010.02212.x

Wolf M, McNamara JM (2013) Adaptive between-individual differences in social competence. Trends Ecol. Evol. 28:253–254

Supplementary Material 1

Extra details of methods

This supplementary information provides extra information on methods. This document is divided into three sections: (1) **Experiment setting up and extra technical details**, (2) **Behavioural traits description and selection** and (3) **MCMCglmm codes for univariate models, and (4) MCMCglmm codes for bivariate models.**

1 Experiment setting up and technical details

1.1 Study area

We carried out experiments at Wilsons Promontory National Park, VIC, Australia (Figure S1.1). The area is mostly grassy-open and flat. A road delimits the west side of the study area, and there is a visitor walk track from the road through to the middle of the study area, in the southwestern border (Figure S1.1).

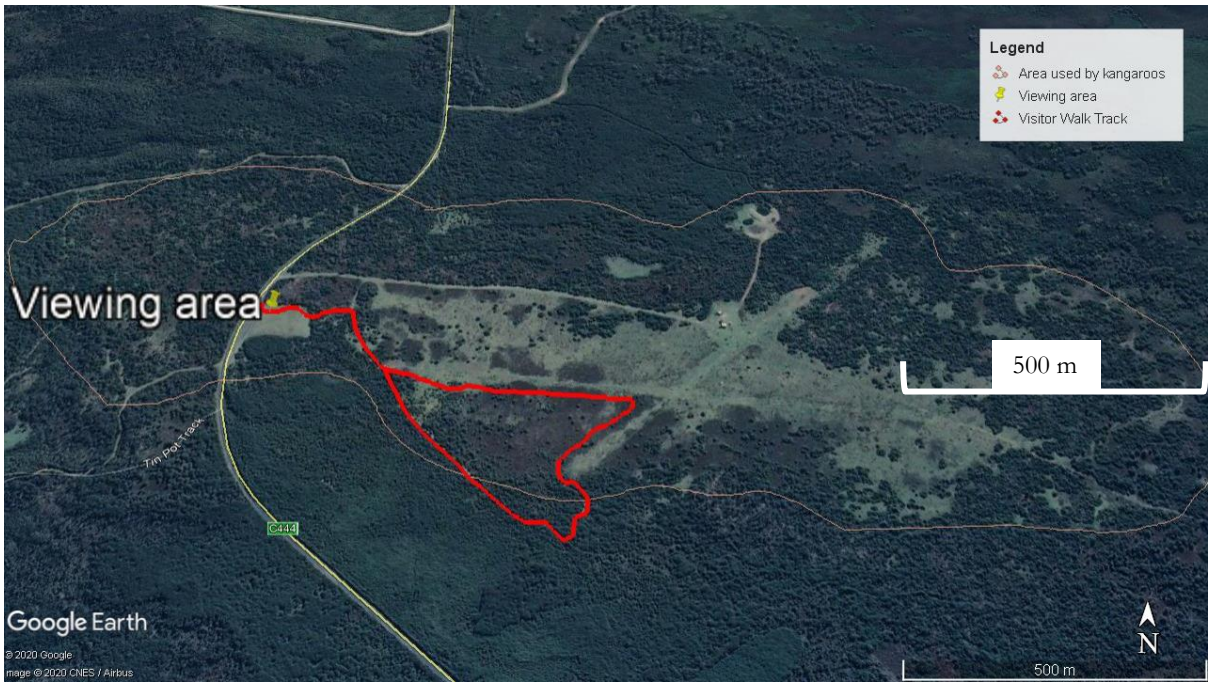


Figure S1.1 Aerial image of the main study site at Wilsons Promontory National Park.

Image © 2020 CNES/Airbus © Google. The area most used by kangaroos is delimited by the maroon line, Visitor Walk Track in red, Viewing Area near the road, in yellow.

1.2 Responses to an Approaching Human

Approaching human (or flight initiation distance (FID) test) is one of the most common experiments to test individual behavioural differences in the wild (Strong et al., 2017). Tests began when individuals in the group were feeding. We started from a distance of 30 m to avoid disturbing animals, using a tripod to mark the initial point and walked steadily and directly towards a target individual until it took flight. Then, we measured the distance walked by the

observer (in m) using a range finder and calculated the FID as 30 m minus the distance walked from the tripod (Figure S1.2). We also calculated the distance displaced using the initial distance from the tripod, the final distance from the tripod, and the angle between the initial and final points. We tested individuals in group sizes of one up to nine; for those in groups, we aimed for an individual in the peripheral area of the group.

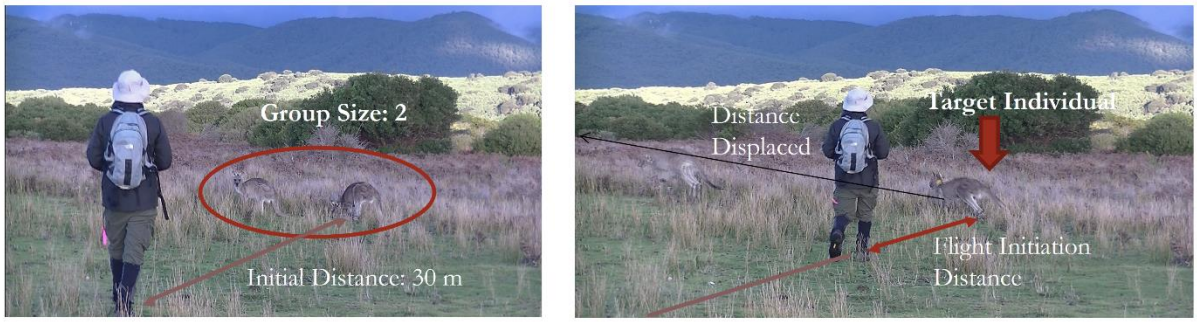


Figure S1.2. The approaching human experiment

1.3 Responses to a Foreign Object

We used a remote-controlled (RC) car as a 'foreign object' because kangaroos did not seem to react to stationary visual objects. Also, it made it easier to test a target group of individuals instead of waiting for them to spontaneously interact with or avoid an object left in their proximity at the observation sampling time. We chose the 'Rock Crawler' model (HSP® 1/10 Pangolin Electric 4WD RTR RC Rock Crawler) because we could easily control the speed and it was capable of traversing irregular surfaces. The RC car also emitted a noise while moving, which was useful in attracting the kangaroos' attention. The set-up order we followed was: install the tripod with the camera attached about 30 m from a target group, place the RC car approximately 15 – 20 m from the target group, measure initial RC car distances and bearings to individuals, return to the tripod, identify the individuals, measure initial observer distances and bearings to individuals, start video recording, drive the RC car parallel to the main axis of the group, and follow the individuals in the video for three minutes (Figure S1.3). We stopped

moving the object when all individuals became vigilant, i.e. raised their heads from feeding. In a few trials, we could not keep all the individuals in the video for the complete three minutes, so we coded their behaviour for as long as we could and used percentages out of the length of time observed rather than out of three minutes. Distances and bearings were measured using a range finder (Bushnell®4x20, Truth With Clear Shot) and a compass (Suunto® KB-14/360R G), respectively.

All trials were recorded using a JVC® GZ-R550BU camcorder, set on a tripod. We used a Bushnell®4x20, Truth With Clear Shot, range finder to measure distances, in metres. We used a combination of measures with the range finder and a Suunto® KB-14/360R G compass to measure angles of displacement, and hence to calculate the distance individuals moved in response to each of the trials. We used the software Solomon Coder beta 17.03.22 (https://solomoncoder.com/) to code individual behaviour from the videos of the foreign object and novel sound trials. We aimed for "blind" scoring by viewing the videos in black and white, which reduced the chances of recognizing individuals by their collar or tag colour. During the video coding, individuals were assigned an A-Z generic ID from left to right. Identities were reassigned after video scoring.

Figure S1.3 The foreign object experiment

1.4 Responses to Novels Sounds

We created thirteen 20-sec playback files of different instrumental pieces (classical or rock, involving instruments such as string, brass, and/or percussion). We selected sections with little variation in sound amplitude. We used the Audacity® 2.1.2 program (http://audacityteam.org/) to apply a high-pass filter (100 Hz cut-off, 36 dB roll-off) and normalise the peak amplitude to -5 dB, removing DC offset (centre on 0.0 vertically), normalising stereo channels independently. Maximum sound amplitude was set to 49 ± 1 dBA at 1 metre using a Digitech® QM-1589 compact digital sound level meter (high-level range, A-frequency weighting, and fast time weighting). We chose this maximum amplitude during trial testing because it led to more variation in response in comparison with lower or higher amplitudes. In the field, the observer placed a black Bluetooth speaker (Braven® BRV-X Waterproof) at approximately 20 m from a target group, facing the sound projector towards the individuals. Experimental set-up (Figure S1.4) was similar to that of the foreign object test, with the observer and video camera stationed about 25 m from the group. Playback tracks were played once from a smartphone (Samsung® Galaxy S5), and the target group was videotaped for 3 minutes. Video coding was carried out as above.

Figure S1.4 The novel sounds experiment

1.5 Defining Group Size

We used the 10-m chain rule (Jarman, 1987) to define a group. This rule includes all individuals within 10 m of at least one other individual in the same group (Figure S1.5). Group sizes: 2 (Group A), 5 (Group B), and 1 (Group C).

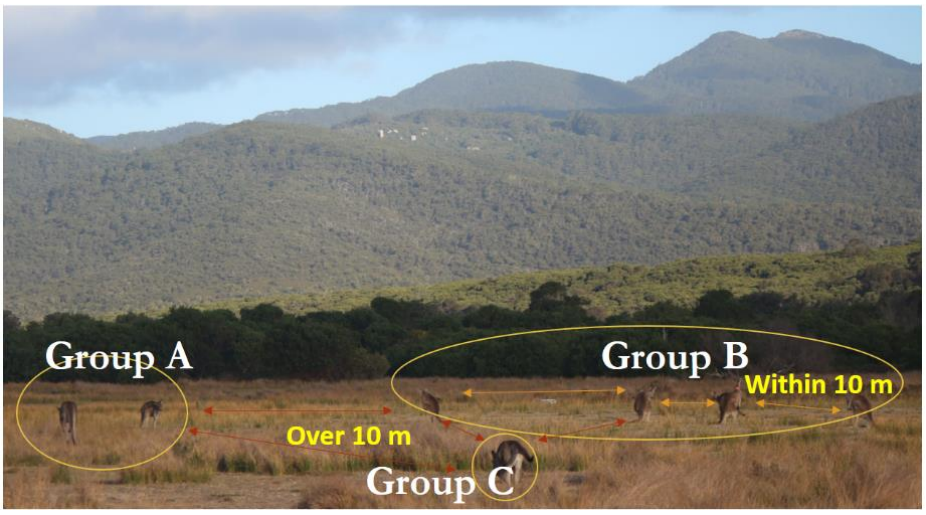


Figure S1.5 Defining group sizes

2 Behavioural trait description and selection

Behavioural trait selection

We measured 16 traits across the three experimental tests (Table S1.1). We retained one trait from the field and another from video coding for each test, except for the FID test, for further analyses. We considered data distribution, how much information the traits contained, and correlation with other traits within the same experiment to select which traits to use. We used the FID trait for the Approaching Human test. We decided to use the binary variable, Hopped Away, for both the Foreign Object and Novel Sounds tests, for consistency. We ran Principal Component Analyses using the R package vegan (Oksanen et al., 2017), and chose the trait with the highest loading in PC1 for traits scored from the videos. The trait selected was thus Duration of Feeding in both tests (Table S1.2). We also performed pairwise Pearson correlations between traits. At this stage, here, of trait selection, we just ran Pearson correlations within "coded from videos" or "recorded in the field" (see Table S1.3) as correlations across more different tests were addressed as parts of the manuscript and addressed used the variance-partitioning approach using a multivariable model. Most traits were correlated (

Table S1.4), indicating that there was no loss of information by reducing the number of traits considered.

In summary, we selected the traits with more information or that had more reliable data distribution for these analyses, i.e., Gaussian or binary family (Supplementary Material 1, Tables S1.1-3) and fitted univariate models for each one using the MCMCglmm R package (Hadfield, 2010). Traits used were: FID-human (using a Gaussian distribution), Hopped-Away-Object (binary), Duration-Feeding-Object (Gaussian), PC1-Object (Gaussian), Hopped-Away-Sound (binary), Duration-Feeding-Object (Gaussian), PC1-Sound (Gaussian), and Dominance (binary) (Table 1).

Table S2.1 Description of the behavioural traits measured in each experimental test. Traits in bold were used in the MCMCglmm analyses presented in the main text.

Test		Behavioural Trait	Definition
Approaching Human		*FID-human (m)*	The distance at which an individual flees when approached by a human
		Distance Displaced- human (m)	Distance moved from the initial to the final location, after being approached by a human
Foreign Object	Recorded in the Field	*Approached-object (0/1)*	1 if the individual approached the car, 0 if not
		Hopped-Away-Object (0/1)	1 if the individual hopped away from the car, 0 if not
		Did Not Displace-object (0/1)	1 if the individual did not displace, 0 if it reacted
	Coded from Videos	*Duration-Feeding-Object (%)*	Percentage of foraging time relative to the total length of the video sample
		Duration of Vigilance-object (%)	Percentage of vigilance time relative to the total length of the video sample
		Duration of Displacement-object (%)	Percentage of displacement time relative to the total length of the video sample
		Latency from First Vigilance to Feeding-object (sec)	Number of seconds an individual took to return to foraging after being vigilant for the first time, following exposure to the remote-controlled car
Novel Sounds	Recorded in the Field	*Hopped-Away-Sound (0/1)*	1 if the individual hopped away after being exposed to playback sounds, 0 if not
		Did Not Displace-sound (0/1)	1 if the individual did not displace, 0 if it reacted
		Distance Displaced-sound (m)	The distance moved from the initial to the final location, after being exposed to a playback sound
	Coded from Videos	*Duration-Feeding-Object (%)*	Percentage of foraging time relative to the total length of the video sample
		Duration of Vigilance-sound (%)	Percentage of vigilance time relative to the total length of the video sample (mean 176.8 sec)
		Duration of Displacement-sound (%)	Percentage of displacement time relative to the total length of the video sample (mean 176.8 sec)
		Latency from First Vigilance to Feeding-sound (sec)	Number of seconds an individual took to return to foraging after being vigilant for the first time, following exposure to the playback sounds

These are all 16 traits I recorded, however, I only analysed a subset. Of these experimental tests, I analysed five traits (FID, and Hopped Away and Duration Feeding for both object and sounds tests). I also analysed Dominance from the long-term observation file.

Table S1.3 Loadings of the Principal Component Analysis. The most representative traits in each test are highlighted in bold.

PCA Loadings: Object test				
	PC1	PC2	PC3	PC4
Approached object	-0.222	-0.522	-0.528	0.271
Hopped Away	-0.129	0.713		-0.204
Did Not Displace	0.207	-0.459	0.414	-0.665
Duration of Feeding	0.535		-0.157	
Duration of Vigilance	-0.487		0.396	0.207
Duration of Displacement	-0.350		-0.550	-0.624
Latency from First Vigilance to Feeding	-0.495		0.251	
Standard Deviation	1.925	1.233	0.933	0.571
Proportion of Variance	0.530	0.218	0.125	0.047
Cumulative Proportion	0.530	0.748	0.872	0.919
Eigenvalue	3.704	1.521	0.870	0.326
PCA Loadings: Sound test				
	PC1	PC2	PC3	PC4
Hopped Away	-0.400	-0.369	0.165	0.500
Did Not Displace	0.278	0.232	-0.801	0.443
Distance Displaced	-0.396	-0.382	-0.183	0.392
Duration Feeding	0.447	-0.329	0.102	
Duration of Vigilance	-0.393	0.456		
Duration of Displacement	-0.314	-0.411	-0.523	-0.620
Latency from First Vigilance to Feeding	-0.392	0.423		
Standard Deviation	1.925	1.233	0.933	0.571
Proportion of Variance	0.530	0.218	0.125	0.047
Cumulative Proportion	0.530	0.748	0.872	0.919
Eigenvalue	3.704	1.521	0.870	0.326

Table S1.4 Trait correlations within tests and confidence intervals (in parentheses). We used the Pearson (r) method.

	Human			
Recorded in the Field		*Distance Displaced-human*		
	FID-human	0.212 (0.165,0.257)		
	Object			
Recorded in the Field		*Approached-object*	*Did Not Displace-object*	
	Hopped-Away-Object	-0.417 (-0.469,-0.363)	-0.453 (-0.503,-0.400)	
Coded from Videos		*Duration of Vigilance-object*	*Duration of Displacement-object*	*Latency from First Vigilance to Feeding-object*
	Duration-Feeding-Object	-0.845 (-0.863,-0.826)	-0.521 (-0.566,-0.473)	-0.767 (-0.792,-0.739)
	Sounds			
Recorded in the Field		*Distance Displaced-sound*	*Did Not Displace-sound*	
	Hopped-Away-Sound	0.735 (0.698, 0.768)	-0.559 (-0.610,-0.505)	
Coded from Videos		*Duration of Vigilance-sound*	*Duration of Displacement-sound*	*Latency from First Vigilance to Feeding-sound*
	Duration-Feeding-Sound	-0.855 (-0.874,-0.833)	-0.374 (-0.438, -0.306)	-0.809 (-0.835,-0.782)

3 MCMCglmm codes for univariate models

```
# Random Intercept Model####
library(MCMCglmm)

# Setting prior distribution####
prior1<- list(G = list(G1 = list(V =1, nu = 0.002),
G2=list(V = 1, nu = 0.002)),
R = list(V = 1, nu = 0.002))
# Here, we set a random structure (G) for two random variables (Animal ID and
#Group ID, see model below) and a residual variance structure (R)
# If a binary trait, we add 'fix=1' after 'nu = 0.002', in the R structure only

# Checking prior curve
priorsamp =rIW(V = 1, nu = 0.002,n = 1000000)
priorsamp=priorsamp[priorsamp<100]
plot(density(priorsamp,from = 0,to=100))#plot prior to see curve – important as check
whether #the prior is not affecting the posterior distribution


# Fitting the model####
Times<-1000 #you can start with lower values here, and increase depending on
autocorrelation (see below)
Model1_DurFeedObject <- MCMCglmm(PercentageFeedingObject ~ 1+
                                 AgeCat+
                                 Sex+
                                 Temperature+
                                 East+
                                 GroupSize+
                                 GroupAssociates+
                                 TestNumber+
                                 Year+
                                 DifferenceDays+
                                InitialDistaceRooxCar,
                                 random = ~ Animal.ID+ GroupID,
                                 data = dtname,
                                 nitt=1300*Times,thin=1*Times,burnin=300*Times, prior = prior1,
                                 verbose = T, rcov=~units,family="gaussian" )

plot(Model1_DurFeedObject$VCV) #random structute density plots
autocorr(Model1_DurFeedObject$Sol)
autocorr(Model1_DurFeedObject$VCV) ## columns that indicate autocorrelation have
first lag as '1'. You want the all values across the consecutive lags to be below 0.1. Look which
lag is <0.1, and use the lag value as the 'times' object – that way, you will multiply nitt, thin, and
burnin for that value.
summary(Model1_DurFeedObject) # outputs
save(Model1_DurFeedObject,file=" Model1_DurFeedObject.Rdata")
```

```
# Repeatability estimation####

# Individual Repeatability
DurFeedObject_IRepeatability <- Model1_DurFeedObject $VCV[, "Animal.ID"]/
  (Model1_DurFeedObject $VCV[,"Animal.ID"] +
Model1_DurFeedObject$VCV[,"GroupID"] + Model1_DurFeedObject$VCV[, "units"])

hist(DurFeedObject_IRepeatability)
posterior.mode(DurFeedObject_IRepeatability)
HPDinterval(DurFeedObject_IRepeatability)
save(DurFeedObject_IRepeatability,file="DurFeedObject_IRepeatability.Rdata")

# Group Repeatability
DurFeedObject_GRepeatability <- Model1_DurFeedObject $VCV[, "GroupID"]/
  (Model1_DurFeedObject $VCV[,"Animal.ID"] +
Model1_DurFeedObject$VCV[,"GroupID"] + Model1_DurFeedObject$VCV[, "units"])

hist(DurFeedObject_GRepeatability)
posterior.mode(DurFeedObject_GRepeatability)
HPDinterval(DurFeedObject_GRepeatability)
save(DurFeedObject_GRepeatability,file="DurFeedObject_GRepeatability.Rdata")
```

 # Random Slope Model####

```
prior2<- list(G = list(G1 = list(V=diag(2), n = 2),  #structure required to fit an
interaction
G2=list(V = 1, nu = 0.002)),
R = list(V = 1,nu = 0.002)) #fix=1

Times<-10000 #check comments above
Model4_ DurFeedObject <- MCMCglmm(PercentageFeedingObject ~ 1+
                            AgeCat+
                            Sex+
                            Temperature+
                            East+
                            GroupSize+
                            GroupAssociates+
                            TestNumber+
                            Year+
                            DifferenceDays+
                          InitialDistaceRooxCar,
                  random = ~ us(1+TestNumber): Animal.ID +  GroupID,
                  data = dtname,
                  nitt=1300* Times,thin=1* Times,burnin=300* Times, prior = prior2,
                  verbose = T, rcov=~units)

plot(Model4_ DurFeedObject $Sol)
plot(Model4_ DurFeedObject $VCV)
autocorr(Model4_ DurFeedObject $Sol)
autocorr(Model4_ DurFeedObject $VCV) #check comments above
HPDinterval(Model4_ DurFeedObject $VCV)
summary(Model4_ DurFeedObject )
```

4 MCMCglmm codes for bivariate models

 # Data preparation####

- Make sure R reads your factors as characters (ie., special attention to Animal ID or Year whose levels tend to be numbered, and R automatically reads it as numeric variables)
- Before merging files, make sure to remove NAs from the variables – also, a hot tip is to have a variable called 'Data File' before merging, so after the different files are merged, it is easier to do data cleaning if required
- After merging the files, variables that are not common across different responses will receive NAs – transform those NAs to '0', and we will set the models in the way those 0 values will not be used. ATTENTION! Do this only for explanatory variables.

 # Setting Prior####

```
Prior3 <- list(G = list(G1 = list(V = diag(2), n = 2), #diag is used to estimate a covariance across
                        the #random variable listed first in the model structure below (ie., Animal ID)
                              G2=list(V=diag(2),n=2)), #ditto
                              R = list(V = diag(2), n = 2)) #Attention, here, if one of the responses is
                        categorical, #you should list it second in the model below, and include 'fix=2'
                        here (see above). If #both responses are categorical, you should use 'fix=1', and
                        the residual variance of #both will be set to 1 (you read this as, if you use 'fix=2':
                        fix residual variance to 1 from #the second listed response, or if you use 'fix=1':
                        fix residual variance to 1 from the first #listed response
```

 # Fitting a bivariate model####

```
Times<-1000
ModelBiv_FIDxDurFeedObject<-MCMCglmm(cbind(FlightInitDist, PercentageFeedingObject)
~
                      trait - 1+
                       trait:(AgeCat + Sex + Temperature + East + GroupSize + RelFemAlong +
                              TestNumber + DifferenceDays + Year)+
                         at.level (x = trait, level = 2): InitialDistaceRooxCar,
                        random = ~us(trait):Animal.ID + us(at.level(x=trait,level=2)):GroupID,
                        rcov = ~idh(trait):units, family = c("gaussian","gaussian"),
                        data = data.merged5, prior = prior2.2FIDxPFeedC, verbose = TRUE,
                        nitt=1300*Times,thin=1*Times,burnin=300*Times)
```

Here, we are estimating trait variance-covariance across Animal ID, trait variance for the second listed #response only across Group ID, and residual variance only for both responses. If both traits were #originally from the same data file, you could also estimate residual covariances by using: "rcov = #~us(trait):units".

```
plot(ModelBiv_FIDxDurFeedObject$Sol)
plot(ModelBiv_FIDxDurFeedObject$VCV)
posterior.mode(ModelBiv_FIDxDurFeedObject$VCV)
HPDinterval(ModelBiv_FIDxDurFeedObject$VCV)
autocorr(ModelBiv_FIDxDurFeedObject$VCV)
summary(ModelBiv_FIDxDurFeedObject)

save(ModelBiv_FIDxDurFeedObject,file=" ModelBiv_FIDxDurFeedObject.RData")
```

References

Jarman, P.J., 1987. Group size and activity in eastern grey kangaroos. Anim. Behav. 35, 1044–1050. doi:10.1016/S0003-3472(87)80161-6

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2017. vegan: Community Ecology Package.

Strong, J.S., Weladji, R.B., Holand, Ø., Røed, K.H., Nieminen, M., 2017. Personality and fitness consequences of flight initiation distance and mating behavior in subdominant male reindeer (Rangifer tarandus). Ethology 123, 484–492. doi:10.1111/eth.12619

Supplementary Material 2

Extra details of Results

This supplementary information provides complete outputs of the (generalised) linear mixed effect univariate models.

Table S2.1: Flight Initiation Distance. Full details of output from linear mixed effects model of FID, with Gaussian errors.

'Model 1': Random Intercepts model with all fixed effects					
Iterations = 300001:1299001		Thinning interval = 1000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-2.496	-8.634	3.823	972.6	0.426
Age (Older)	0.532	-0.086	1.218	1000	0.112
Sex (Male)	1.471	0.676	2.357	1000	0.002
Temperature	0.008	-0.035	0.050	1000	0.704
Easting	0.001	0.001	0.002	979.1	<0.001
Group Size	-0.271	-0.449	-0.091	995	0.008
Group Associates					
- male courting female	-1.455	-2.668	-0.164	904.1	0.024
- mother with LPY	0.841	0.151	1.492	1000	0.026
- 1YO with mother	-0.032	-0.840	0.828	1000	0.948
- mother with SPY/MPY	0.469	-0.500	1.453	1131	0.318
- mother with 1YO	0.603	-0.208	1.343	1000	0.144
Test Number	-0.307	-0.443	-0.151	1000	<0.001
Year (2018)	0.680	0.186	1.247	1000	0.012
Year (2019)	1.182	0.303	2.272	1000	0.034
Observer (WMC)	-1.104	-1.882	-0.250	1000	0.002
Number of Days from Previous Test	0.001	0.000	0.002	1000	0.004
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	9.282	7.407	11.47	1000	
Residual	9.321	8.62	10.07	1358	

'Model 2': Random Intercepts model with *age* + *sex* + *test number* + *number of days from previous test*					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	6.42962	5.772	7.129	1000	<0.001
Age (Older)	0.529	-0.053	1.087	1000	0.078
Sex (Male)	1.132	0.308	2.055	1000	0.018
Test Number	-0.087	-0.168	-0.006	1000	0.032
Number of Days from Previous Test	0.002	0.001	0.003	1000	<0.001

Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	10.54	8.513	12.75	912.6
Residual	9.369	8.705	10.07	840.6

'Model 3': Random Intercepts model with *test number + number of days from previous test*					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	7.719	7.231	8.188	1000	<0.001
Test Number	-0.150	-0.21587	-0.086	1000	<0.001
Number of Days from Previous Test	0.002	-0.000	0.004	1000	0.092

Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	10.74	8.506	12.87	969.2
Residual	9.491	8.819	10.23	722.6

'Model 4': Random Slopes and Intercepts model, with all fixed effects				
Fixed effects are as in 'Model 1', hence not repeated here				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	14.189	10.828	17.635	1000
Individual x Test Number (Slope)	0.168	0.098	0.238	1000
Individual Intercept-Test Number Slope (Covariance)	-1.014	-1.509	-0.533	1000
Residual	8.607	7.907	9.350	1000

Here, we show the complete variables/level results from the model. Definitions: 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. Males 1.5 m less aproachable than females on average and animals at 1000 m to the east were less approachable by 1m than individuals at 0 m. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017), Observer (CD). Sample sizes: 1642 observations on 292 individuals.

Table S2.2: Hopped-Away-Object. Full details of output from the generalised linear mixed effects model of the binary trait of whether or not an individual hopped away in the 'Object' test.

'Model 1': Random Intercept with all fixed effects					
Iterations = 3000001:12990001		Thinning interval = 10000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-3.935	-13.880	7.832	1000	0.420
Age (Older)	0.686	-0.547	1.769	1031	0.252
Sex (Male)	0.291	-0.507	1.123	1000	0.482
Temperature	-0.067	-0.165	0.022	1000	0.152
Easting	0.001	-0.001	0.002	1000	0.282
Group Size	-0.323	-0.677	-0.034	1000	0.036
Associates:					
- male courting female	0.150	-1.262	1.595	911	0.870
- mother with LPY	0.252	-0.638	1.160	1000	0.594
- 1YO with mother	0.898	-0.664	2.411	1000	0.264
- mother with SPY/MPY	-0.093	-2.052	1.805	957	0.926
- mother with 1YO	0.415	-0.803	1.710	1000	0.496
Test Number	-0.256	-0.518	-0.010	1107	0.046
Year (2018)	1.132	0.005	2.217	1000	0.040
Year (2019)	1.122	-1.127	3.388	1000	0.294
Number of Days from Previous Test	0.002	0.000	0.004	871	0.066
Intiial distance to the stimulus	-0.031	-0.120	0.054	885	0.494
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	1.648	0.003	3.462	1000	
Group	11.020	5.623	16.920	1000	
Residual	Fixed to 1				

'Model 2': Random Intercept with *age + sex + test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	1.049	0.000	2.417	700
Group	8.874	5.144	13.430	747
Residuals	Fixed to 1			

'Model 3': Random Intercept with *test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp

Individual	1.000	0.001	2.256	1000
Group	8.513	4.821	12.500	903.7
Residual	Fixed to 1			

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	24.137	1.851	66.605	93
Individual x Test Number (Slope)	1.132	0.074	2.859	104
Individual Intercept-Test Number Slope (Covariance)	-4.419	-11.987	0.141	109
Group (Intercept)	30.57	8.887	67.92	101.2
Residual	Fixed to 1			

Here, we show the complete variables/level results from the model. This model is binomial with estimates on the latent scale (but back transformed to data-scale in Table 2 (i.e., the repeatability estimates)). Definitions: 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 930 observations on 240 individuals.

.

Table S2.3: Duration-Feeding-Object. Full details of output from linear mixed effects model of duration of feeding in the 'Object' test, with Gaussian errors.

'Model 1': Random Intercept with all fixed effects					
Iterations = 1500001:6495001		Thinning interval = 5000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	41.500	-3.559	92.170	1000	0.080
Age (Older)	-10.900	-16.880	-4.854	1000	<0.001
Sex (Male)	0.443	-4.410	4.702	1000	0.834
Temperature	0.309	-0.106	0.678	1000	0.124
Easting	0.000	-0.006	0.006	1000	0.888
Group Size	1.817	0.685	3.121	1000	<0.001
Associates:					
- male courting female	-6.735	-14.170	1.285	636	-6.735
- mother with LPY	0.505	-4.354	5.538	1000	0.505
- 1YO with mother	-5.305	-13.630	1.938	1187	-5.305
- mother with SPY/MPY	7.363	-3.191	17.900	971	7.363
- mother with 1YO	0.502	-5.265	6.384	875	0.502
Test Number	2.634	1.390	3.932	1108	2.634
Year (2018)	2.032	-2.649	6.919	1000	2.032
Year (2019)	1.607	-8.048	12.510	1000	1.607
Number of Days from Previous Test	-0.011	-0.020	-0.002	886	-0.011
Intiial distance to the stimulus	-0.310	-0.748	0.075	1000	-0.310
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	48.820	9.910	94.050	1000	
Group	117.600	63.460	168.800	1000	
Residual	436.200	380.800	501.100	904	

'Model 2': Random Intercept with *age* + *sex* + *test number* + *number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	49.260	8.787	90.740	885
Group	135.000	83.300	185.400	899
Residuals	428.000	376.900	490.800	1000

'Model 3': Random Intercept with *test number* + *number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp

Individual	57.82	18.22	99.83	1000
Group	143.900	92.680	195.500	1000
Residual	426.400	367.300	478.000	1265

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	238.306	123.069	376.240	907
Individual x Test Number (Slope)	6.927	1.923	12.920	1000
Individual Intercept-Test Number Slope (Covariance)	-38.626	-66.062	-14.150	1000
Group (Intercept)	108.100	51.670	166.700	1000
Residual	393.5	339	454.2	1000

Here, we show the complete variables/level results from the model. Definitions: 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 930 observations on 240 individuals

Table S2.4: PC1-Object. Full details of output from linear mixed effects model of principal component 1 in the 'Object' test, with Gaussian errors.

'Model 1': Random Intercept with all fixed effects					
Iterations = 1500001:6495001		Thinning interval = 5000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.433	-4.551	1.703	1000	0.422
Age (Older)	-0.662	-1.048	-0.283	1000	0.002
Sex (Male)	0.099	-0.249	0.421	1338	0.548
Temperature	0.019	-0.007	0.044	1000	0.15
Easting	0.000	0.000	0.001	1000	0.646
Group Size	0.127	0.056	0.211	1000	0.002
Associates:					
- male courting female	-0.469	-1.047	0.056	1000	0.094
- mother with LPY	-0.005	-0.324	0.339	1000	0.996
- 1YO with mother	-0.139	-0.719	0.355	1282	0.588
- mother with SPY/MPY	0.492	-0.164	1.203	1000	0.172
- mother with 1YO	0.072	-0.342	0.503	1000	0.766
Test Number	0.183	0.099	0.286	1000	<0.001
Year (2018)	0.195	-0.167	0.536	1000	0.252
Year (2019)	0.051	-0.607	0.660	1000	0.876
Number of Days from Previous Test	-0.001	-0.001	0.000	1000	0.024
Intiial distance to the stimulus	-0.003	-0.032	0.026	837	0.818
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	0.452	0.243	0.684	1292	
Group	0.417	0.199	0.629	1136	
Residual	1.884	1.628	2.156	1000	

'Model 2': Random Intercept with *age + sex + test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	0.421	0.217	0.639	1270
Group	0.468	0.274	0.680	882
Residuals	1.879	1.631	2.133	1000

'Model 3': Random Intercept with *test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp

Individual	0.476	0.269	0.707	903
Group	0.490	0.254	0.704	1000
Residual	1.884	1.641	2.147	1142

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	2.022	1.283	2.895	1347
Individual x Test Number (Slope)	0.129	0.077	0.191	1000
Individual Intercept-Test Number Slope (Covariance)	-0.398	-0.606	-0.206	1000
Group (Intercept)	0.355	0.117	0.597	1417
Residual	1.566	1.306	1.807	1000

Here, we show the complete variables/level results from the model. Definitions: 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 930 observations on 240 individuals. Comparing these results with the ones from the Object raw traits, above, individual repeatability was higher for PC1-Object (18%) than it had been for Hopped-Away-Object (5%) or Duration-Feeding-Object (8%). The among-group proportion of PC1-Object (16%) was more similar to the Duration-Feeding-Object (19%) than to Hopped-Away-Object (41%).

Table S2.5: Hopped-Away-Sound. Full details of output from the generalised linear mixed effects model of the binary trait of whether or not an individual hopped away in the 'Sound' test.

'Model 1': Random Intercept with all fixed effects					
Iterations = 1500001:6495001		Thinning interval = 5000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	234.563	-842.070	1339.481	57	0.656
Age (Older)	7.549	-59.086	72.084	48	0.852
Sex (Male)	3.385	-51.307	56.337	94	0.884
Temperature	3.393	-7.321	14.110	40	0.526
Easting	-0.014	-0.171	0.118	48	0.830
Group Size	2.979	-22.124	27.599	115	0.820
Associates:					
- male courting female	-47.911	-133.517	28.802	44	0.282
- mother with LPY	-23.262	-91.201	42.657	79	0.466
- 1YO with mother	-21.546	-114.003	76.575	58	0.650
- mother with SPY/MPY	67.077	-133.355	267.766	45	0.516
- mother with 1YO	-137.334	-235.593	-46.794	38	0.002
Test Number	-24.572	-43.236	-8.560	136	0.006
Year (2018)	104.020	16.117	207.738	69	0.030
Year (2019)	-90.134	-295.581	132.576	35	0.414
Number of Days from Previous Test	0.073	-0.040	0.214	48	0.208
Intiial distance to the stimulus	-15.923	-23.882	-8.199	63	<0.001
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	10738.0	4779.0	17214.0	25	
Group	58876.0	32814.0	79483.0	14	
Residual	Fixed to 1				

'Model 2': Random Intercept with *age + sex + test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	2132.0	362.0	5900.0	6
Group	11016.0	1978.0	26452.0	4
Residuals	Fixed to 1			

'Model 3': Random Intercept with *test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	2.083	0.000	9.352	102
Group	39.320	12.260	90.460	107
Residual	Fixed to 1			

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	16238.0	951.0	29308.8	10
Individual x Test Number (Slope)	365.0	0.3	1055.3	15
Individual Intercept-Test Number Slope (Covariance)	-2263.0	-5269.9	573.2	17
Group (Intercept)	44092.0	11915.0	73366.0	5
Residual	Fixed to 1			

Definitions: Here, we show the complete variables/level results from the model. This model is binomial with estimates on the latent scale (but back transformed to data-scale in Table 2 (i.e., the repeatability estimates)) 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 657 observations on 215 individuals .

Table S2.6: Duration-Feeding-Sound. Full details of output from linear mixed effects model of duration of feeding in the 'Sound test, with Gaussian errors.

'Model 1': Random Intercept with all fixed effects					
Iterations = 1500001:6495001		Thinning interval = 5000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)					
Age (Older)	-4.527	-67.980	58.670	1000	0.878
Sex (Male)	1.261	-4.568	7.658	1000	0.722
Temperature	-1.433	-6.177	3.191	1000	0.566
Easting	-0.273	-0.964	0.329	950	0.442
Group Size	0.009	0.000	0.017	1000	0.052
Associates:					
- male courting female	-7.770	-16.060	0.254	911	0.068
- mother with LPY	0.884	-4.659	6.438	1000	0.734
- 1YO with mother	0.185	-7.208	9.223	1111	0.970
- mother with SPY/MPY	-8.713	-23.230	6.132	1000	0.268
- mother with 1YO	2.546	-4.027	9.718	904	0.468
Test Number	2.979	1.370	4.777	1000	0.004
Year (2018)	-9.090	-15.800	-2.553	1000	0.008
Year (2019)	5.192	-8.213	17.450	1000	0.430
Number of Days from Previous Test	-0.008	-0.019	0.003	1000	0.136
Intiial distance to the stimulus	-0.326	-0.983	0.401	1000	0.368
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	17.160	0.001	55.740	722	
Group	217.700	145.500	296.000	1000	
Residual	387.100	317.800	453.100	888	

'Model 2': Random Intercept with *age + sex + test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	17.920	0.001	57.130	1000
Group	237.900	157.600	306.600	1000
Residuals	385.6	316.6	443.3	1000

'Model 3': Random Intercept with *test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp

Individual	25.160	0.000	68.190	1000
Group	230.500	164.200	304.900	1000
Residual	384.200	313.400	444.300	1000

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	10.880	0.112	58.895	1029
Individual x Test Number (Slope)	1.470	0.171	3.734	1000
Individual Intercept-Test Number Slope (Covariance)	-1.013	-7.824	4.127	1000
Group (Intercept)	218.700	151.100	299.200	1000
Residual	386.1	331.8	448.1	1031

Definitions: Here, we show the complete variables/level results from the model. 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 657 observations on 215 individuals.

Table S2.7: PC1-Sound. Full details of output from linear mixed effects model of principal component 1 in the 'Sound test, with Gaussian errors.

'Model 1': Random Intercept with all fixed effects					
Iterations = 1500001:6495001		Thinning interval = 5000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)					
Age (Older)	3.514	-1.483	8.754	1000	0.196
Sex (Male)	-0.066	-0.459	0.322	1000	0.718
Temperature	0.090	-0.175	0.411	1000	0.566
Easting	0.015	-0.043	0.064	961	0.538
Group Size	0.000	-0.001	0.000	1000	0.21
Associates:					
- male courting female	0.322	-0.230	0.791	889	0.228
- mother with LPY	-0.016	-0.396	0.298	1000	0.932
- 1YO with mother	-0.212	-0.757	0.270	1130	0.454
- mother with SPY/MPY	0.621	-0.306	1.537	1000	0.218
- mother with 1YO	-0.322	-0.769	0.095	830	0.154
Test Number	-0.211	-0.313	-0.094	1000	<0.001
Year (2018)	0.966	0.468	1.510	1000	<0.001
Year (2019)	-0.503	-1.492	0.601	1000	0.356
Number of Days from Previous Test	0.001	0.000	0.001	1000	0.114
Intiial distance to the stimulus	-0.025	-0.068	0.019	1000	0.278
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Individual	0.056	0.000	0.176	1000	
Group	2.040	1.569	2.536	710	
Residual	1.342	1.141	1.560	1000	

'Model 2': Random Intercept with *age + sex + test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual	0.054	0.000	0.170	1000
Group	2.246	1.741	2.725	1000
Residuals	1.338	1.140	1.572	1000

'Model 3': Random Intercept with *test number + number of days from previous test*				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp

Individual	0.074	0.000	0.218	856
Group	2.195	1.757	2.764	1000
Residual	1.339	1.118	1.559	1000

'Model 4': Random Slope with all fixed effects				
Fixed effects estimates are similar to above				
Random effects				
	post.mean	l-95% CI	u-95% CI	eff.samp
Individual (Intercept)	0.514	0.171	0.943	1000
Individual x Test Number (Slope)	0.101	0.055	0.161	838
Individual Intercept-Test Number Slope (Covariance)	-0.126	-0.271	-0.010	801
Group (Intercept)	1.952	1.506	2.440	1181
Residual	1.140	0.931	1.343	1000

Definitions: Here, we show the complete variables/level results from the model. 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 'LPY': large pouch young, "SPY/MPY": small or medium pouch young, 1YO: 1-year-old juvenile. The reference levels are: Age (yearling), Sex (female), Associates (none), Year (2017). Sample sizes: 657 observations on 215 individual. Comparing these results with the ones from the Sound raw traits, above, there was no support for individual repeatability in PC1-Sound, consistent with no indication of individual repeatability for the observed traits. Considering the among-group proportions of variance, the PC1-Sound proportion of 59% was more similar to Hopped-Away-Sound (61%) than to Duration-Feeding-Sound (36%), which is different from the pattern observed in the Object traits.

Table S2.8: Dominance. Full details of output from the generalised linear mixed effects model of the binary trait of whether or not an individual won a dominance encounter.

'Model 1': Random Intercept with all fixed effects					
Iterations = 15000001:64950001		Thinning interval = 50000		Sample size = 1000	
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.693	-7.819	3.723	1000	0.564
Sex Code					
- 2: female-female	-1.566	-6.083	2.087	1126	0.468
- 3: male-female	-0.099	-4.467	4.582	1000	0.936
- 4: female-male	1.868	-2.944	6.667	1000	0.448
Age Difference					
- 1YO-2YO	-6.577	-15.247	2.063	1000	0.138
- 1YO-Older	-457.218	-1011.685	-14.046	4.55	<0.001
- 2YO-1YO	14.419	5.877	24.506	1000	<0.001
- 2YO-2YO	0.605	-5.222	6.004	1000	0.86
- 2YO-Older	-10.536	-17.511	-3.353	1000	<0.001
- Older-1YO	395.594	52.233	678.654	7.54	<0.001
- Older-2YO	10.892	4.277	17.934	1000	<0.001
- Older-Older	1.917	-4.180	7.505	1000	0.496
Random effects variance components					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Focal	63.380	31.900	103.600	1000	
Opponent	56.990	29.840	88.530	1000	
Year	0.1907	0.0002411	0.8151	521.6	
Residual	Fixed to 1				

Definitions: Here, we show the complete variables/level results from the model. This model is binomial with estimates on the latent scale (but back transformed to data-scale in Table 2 (i.e., the repeatability estimates)). 'post.mean' stands for posterior distribution mean and 'eff.samp' for effective sample sizes. 1YO: 1-year-old juvenile., etc. The reference levels are: Sex Code (1: male-male), Age Difference (1YO-1YO). Sample sizes: 1461 observations on 218 individuals.

Chapter III.

Whose personality determines group-level behaviour in a wild marsupial?

Authorship

Weliton Menário Costa[1], Timothee Bonnet[1], Marco Festa-Bianchet[2,1], Loeske E.B. Kruuk[1]

1 Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT, Australia

[2] Département de biologie, Université de Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke, Québec, J1K 2R1, Canada

Abstract

Group-level behaviour, defined as the average of group members' responses, arises from individuals' behaviour, and hence can be affected by individual trait variation. However, group cohesion requires that all individuals behave similarly. Some studies have shown that 'personality' contributes to group-level behaviour – but whose personality matters, and how important is personality compared to other factors? Analyses of behavioural data from experimental trials on over 250 wild eastern grey kangaroos over three years revealed that the average personality of group members better predicted the group-level duration of feeding or whether or not the group moved away, compared to extreme personalities (highest or lowest score) or older individuals, suggesting that group decisions are 'democratic'. The between-group variance explained by the personality of the individuals in the group was higher than the variance explained by group structure, such as adult sex ratio and proportion of elder individuals, or spatial-temporal variation, and hence, was the main driver of group-level responses. In addition, individual personality shaped group composition, with evidence for assortment based on individual personalities, over and above effects of spatial structure.

Introduction

Studies of animal 'personality', individual repeatable differences in behaviour (Wolf and Weissing 2012; Dall and Griffith 2014), now provide clear evidence that personality plays a role in social interactions (Montiglio et al. 2013; Planas-Sitjà et al. 2018), and also that personality affects group-level behaviour (Jolles et al. 2017, 2020), defined here as the average of group members' responses. But how do different personalities within a group determine group-level behaviour? Some studies have found that key individuals may exert a disproportionate influence on group-level decisions (King and Cowlishaw 2009; Strandburg-Peshkin et al. 2018), for example, extreme personality types (Brown and Irving 2014; Planas-Sitjà 2020) or older individuals (Wolf and Krause 2014; Allen et al. 2020), with some individuals being leaders and others followers. Experience and motivation shape leader-follower interactions in sticklebacks (*Gasterosteus aculeatus*) (Webster 2017), and knowledge of the environment was important in maintaining a leadership role in Australian steel blue sawfly (*Perga affini*) (Hodgkin et al. 2017). Older individuals lead collective movements in killer whales (*Orcinus orca*) (Brent et al. 2015), African savannah elephants (*Loxodonta africana*) (McComb et al. 2011; Allen et al. 2020), rhesus macaques (*Macaca mulatta*) (Sueur and Petit 2008), and bonobos (*Pan paniscus*) (Tokuyama and Furuichi 2017), and influence forage decisions in groups of rhesus macaque (Lee and Teichroeb 2016). Differences in personality and maturity level can be associated with differences in experience, motivation, or knowledge, and, hence, can unevenly influence group-level responses (Wolf and Krause 2014; Allen et al. 2020). To maintain cohesiveness, groups require a consensus decision (Conradt and Roper 2005). If individual members of the group have conflicting preferences, cohesiveness will impose some costs (Couzin and Krause 2003). Different personalities may conflict on what the best outcome should be (Gavrilets et al. 2016) as they often differ in experience and information acquisition or use (Coppens et al. 2010; Wolf and Krause 2014). If all individuals adjust to a consensus decision, they will share costs (Conradt and Roper 2007). If a consensus is not reached, the group will likely split (Merkle et al. 2015).

Group demographic structure and environment can also contribute to differences between groups in the behaviour of individuals in them (Jolles et al. 2020), but there have been few attempts to quantify how much variance these variables explain in comparison to personality. Interactions between individuals within a group drive group-level outcomes (Couzin et al. 2005). Variation in specific individual traits can be intimately linked with group-level responses as they can alter group structure, function, and behaviour (Jolles et al. 2020). Differences in sex ratio, for example, can affect the structure, and hence, the function of a group (Kappeler 2017; Cox et al. 2019). Individual variability in several traits may affect group-level behaviour (Dreiss et al. 2010; Jolles et al. 2020) but no studies, to our knowledge, have quantified the relative amount of variance explained by these different factors in the context of group decisions.

If individuals tend to assort with others of similar traits (Massen and Koski 2014; Jolles et al. 2018), variance in individual characteristics could affect social organization. Individuals of similar personalities could be observed in the same groups as a result of active choice or a passive outcome of fission-fusion dynamics (Muller et al. 2018). If individuals are grouped based on their personalities over and above any spatial or temporal structure, personality will contribute to group composition, and the group members will influence each other's behaviour, affecting mean group-level behaviour (Wolf et al. 2011; Farine et al. 2015; Herbert-Read 2017). Some studies found that personality could shape qualitatively or quantitatively the social environment. Group members had similar personalities in a cichlid (*Neolamprologus pulcher*) (Schürch et al. 2010), network structure was underpinned by personality differences in great tits (*Parus major*) (Aplin et al. 2013), and groups differed in personality in stickleback (*Gasterosteus aculeatus*) (Herbert-Read 2017). Some studies focused on the role of social interactions in the ontogeny of personality differences (Bengston and Jandt 2014), but few have explored, in the wild, how personality affects social structure (Aplin et al. 2013; Díaz López 2020). Disentangling the different causes of behavioural variation, especially in regards to the interplay between group

versus individual effects on behaviour, can be challenging, but species living in fission-fusion dynamics prove helpful as group membership changes frequently over time (Couzin 2006).

We studied a wild population of eastern grey kangaroos (*Macropus giganteus*) in Victoria, Australia, which has been the subject of long-term monitoring (Montana et al. 2020), to test, firstly, whether individual personality contributed to group-level behavioural response and, secondly, whether individuals assorted based on personality. We tested whether group behaviour was affected by the average personality of group members, extreme personalities within the group, or the average personality of elder individuals. We reasoned that extreme personalities might be important if group members tended to copy the behaviour of the first individual to react to a stimulus. We expected that older kangaroos may play a leadership role in group decisions, as reported in several other species (Merkle et al. 2015; Brent et al. 2015; Allen et al. 2020). Eastern grey kangaroos live in fission-fusion dynamics (Best et al. 2013), and have been shown to have personalities (Best et al. 2015). In our study population, among-group differences in behaviour were larger than individual differences in various traits (Chapter II). Thus, this species provides an excellent opportunity to dissect how individual differences in behaviour may lead to differences in the behaviour of groups. In Chapter II, I found flight initiation distance, FID, (Weston et al. 2012) to be highly repeatable and to co-vary with other behavioural traits across individuals. Therefore, here, we used individual's mean FID as a proxy index for individual personality (Sih et al. 2004; Réale et al. 2007), estimated from our repeated observations on individuals across multiple time points. First, we measured two group-level responses ('All Hopped Away' and 'Mean Feeding Duration') across three experimental tests (response to an approaching human, response to a foreign object, and response to novel sounds). We used group summaries of the individual FIDs measures of group members as predictor variables. Then, we used group composition data from population surveys to test for among-group differences in personality. We applied powerful and novel statistical methods (Froy et al. 2016; Hadfield 2018) to address four questions:

1. Whose personality matters most in the group: that of all individuals, or extreme personalities?
2. Does the personality of elder individuals influence group-level responses more than that of other individuals?
3. How important is personality compared to other factors for group-level response variance?
4. Are individuals grouped based on their personalities over and above spatial structure?

Methods

Study population

We conducted experiments and behavioural observations using a wild but habituated population of eastern grey kangaroos at Wilsons Promontory National Park, VIC (38º57" S, 146º17" E), in which individuals have been artificially marked (King et al. 2011). The programme monitors 120-140 adult females, 40-80 adult males, and 30-80 juveniles every year, where the term 'adult' refers to individuals at least three years of age. All monitored individuals are ear-tagged with a unique colour combination. Females weighing over 18 kg and males weighing over 38 kg receive a unique plastic collar. Most individuals are recaptured once a year (Mackay et al. 2018), and markings are replaced if necessary on recapture. Surveys are conducted regularly during winter/spring (August-November) and autumn (March) each year, when multiple observers conduct a systematic recording of group locations and composition (Montana et al. 2020). We used behavioural data from experimental tests and group composition observations from March 2017 to April 2019. There are two parts to this chapter, and they use slightly different data sets.

Part 1: feeding and hopped group-level responses

In Part 1, we addressed three questions using feeding and hopped responses variables of group members. We assessed the behaviour of individuals in groups through three experimental tests, between August 2017 and April 2019: approaching human test (Approach test), response to a foreign object (Object test), and response to novel sounds (Sound test). We used the 10-metre-chain rule to determine groups (Jarman 1987; King et al. 2015a), defining a group as 2 or more individuals aged at least 12 months within the chain rule (range of 2-10 individuals), but in the third question we also include single individuals. In the Approach test, an observer walked directly towards a target individual and stopped when it took flight. The target individual's flight

initiation response (FID) was recorded in metres (Weston et al. 2012), and we also noted whether the non-target individuals hopped away or not. In the Object test, we moved a remote-controlled car in front of a target group for approximately 15 seconds and recorded the group member responses for 180 seconds. However, if the animal left, the trial was terminated before (mean duration of 174.6 sec, range 18.4 – 180 sec). The Sound test was similar to the Object test. We played a 20-second instrumental music playback and recorded individuals' responses for a mean of 176.8 sec (range 34 – 180 sec). More details of the experimental designs are available as Supplementary Materials.

We summarised the behaviour of group members into two group-level responses that were used in the analyses below. In the Approach test, we measured the 'hop away' response of the group members, following the measure of the flight initiation distance (FID) of a target individual. The individual-level hopped away response of group members was summarized at the group level as an 'All Hopped Away' response, defined as whether all group members hopped away or not (binary). Like the above, in the Object and Sound tests, we also measured the hopped away response of group members and summarised them as group-level All Hopped Away responses to the two respective tests. In addition, we also measured group members' duration of feeding (% of time) in these tests, summarised as a group-level 'Mean Feeding Duration' response, which was defined as the average duration of feeding across all group members. We combined the data sets from these three tests to analyse the group-level responses together, thus in the end considering two traits:

- All Hopped Away (binary): all group members hopped away (or not) in response to either an approaching human, an approaching remote-controlled car, or instrumental music playbacks;
- Mean Feeding Duration: average of the group members' feeding duration following an approaching remote-controlled car or instrumental music playbacks.

Summaries of personalities of individuals in each group

Our previous analyses (Chapter II) found measures of individuals' flight initiation distance, FID, to be repeatable (with repeatability of 48.7%), indicating consistent individual differences or personalities. As discussed in Chapter II, behavioural measures like FID might over-estimate the ecological importance of personality. Still, FID is a reliable measure of individual variability, so we used it here as a proxy for personality and tested its relevance for social factors such as group-level behaviour (Part 1) and group assorting (Part 2, see below). We therefore used the average measures of FID from repeated tests on individuals across multiple time points to generate an individual's 'intrinsic' FID. This individual-level FID was calculated from the mean of the repeated raw measures of FID in each test, tested on average 5.63 (SD ±2.91) times per individual across different social groups throughout three years; we refer to this mean value as 'individual FID' from here on. Then, we used the individual FIDs to generate four group-level summary statistics of the personalities within each group:

- *Mean-FID*: average of all group members' FIDs;
- *Maximum-FID*: the largest value of the individual FIDs for the group, indicating the most reactive personality in the group;
- *Minimum-FID*: the lowest value of the individual FIDs for the group, indicating the least reactive personality in the group;
- *Elders' mean FID*: the average of the individual FIDs of group members aged at least four years.

In the Approach test All Hopped Away response, when calculating these group statistics, we excluded the individual FID of the target individual as the stimulation of target versus non-target individuals was different. In the other two tests, we used the individual FID of all group members because we targeted the stimuli towards the whole group. For testing the role of elders' personality, we used the mean FID of all individuals in an 'elder category' instead of the

personality of the very eldest individual because we did not have the exact age for individuals who were first captured as adults. Hence, we could not identify the eldest individual within the groups in most cases: 68% of groups had at least one individual whom we knew to be four or more years, but did not have an exact age for. However, the mean of elder individuals would always include the personality of the eldest. For the groups for which we did know precisely all ages within the elder category (32% of groups), the average age was 7.7 (SD ±1.7, range 4-11), and 78% were at least seven years old at the time of the observations. Nevertheless, the 'elder' category inevitably included some relatively young adults (e.g., 4 or 5 year-olds) in some cases. For this calculation, we included individuals repeatedly across different years as they aged.

Group structure summaries

We considered individuals' attributes, such as their sex, age, and presence of group associates, to create summary indices for group structure. Group members' sex was summarised as '*adult sex ratio*', considering the number of adult males and adult females in the group. The age structure of the group was summarised as '*proportion of 2yo+*'. The category 2yo+ includes juveniles aged two years (24-36 months) and adults. We combined juveniles with adults because our previous analyses indicate they behave more similarly to the adult class than to 1-year-olds (aged 13-24 months). 1-year-olds are mostly unweaned and more closely associated with their mothers than are 2-year-olds (King and Goldizen 2016). The level of association amongst group members was summarised as a '*proportion of associates*'. We considered pairs of 'associates' as more durable bonds within a social group compared to a general fission-fusion association. This included a mother with her offspring (of any stage, from small pouch young to 1-year old, prior to weaning), a 1-year-old with its mother, or a male courting a female (see Chapter II for further discussion). We then estimated the *proportion of associates* in the group as the proportion of individuals who were part of such a pair. Finally, we estimated '*group size'* as the number of group members of at least one year of age. The four indices of group structure were therefore defined as:

- *Adult sex ratio*: number of adult males divided by the group's total number of adult individuals.
- *Group size*: the total number of individuals aged at least 12 months in the group.
- *Proportion of 2yo+*: the number of individuals aged at least 24 months divided by the group size.
- *Proportion of associates*: the number of individuals in the group who also had some form of more durable association within the group, divided by the group size (so, for example, if a group consisted of a mother and young, and two other non-associated individuals, its value would be 0.5).

Spatial-temporal and experimental factors

We accounted for spatial and temporal variation between groups by considering the location along the east axis (*easting*) and *date*. We did not use northing because most of the variation in space use and habitat type occurs from west to east in our study area. We also considered the experimental factors *stimulus proximity* and *test type*. The *stimulus proximity* factor was a measure of the proportion of the number of individuals in the group at a similar/closer distance to the stimulus compared to the number of individuals at greater distances. *Test type* was a three-level factor to indicate the experiment used to measure the hopped away and feeding duration responses, as described above: i.e. Approach, Object or Sounds. We therefore considered the following spatial-temporal and experimental factors:

- *Easting*: group location along the east axis (metres).
- *Date*: date when the experiment was conducted.
- *Stimulus proximity*: number of individuals at a similar/closer distance to the stimulus divided by the group size (including individuals at a similar distance and those at a greater distance).

- *Test type:* a three-level factor (approach, object, and sound) in All Hopped Away and a two-level factor (object and sound) in Mean Feeding Duration.

Statistical analysis

In Part 1, we used our feeding/hopped group-level responses and the group summaries of personality, group structure, and spatial-temporal and experimental factors to address three questions: (1) Whose personality matters most in the group: that of all individuals, or extreme personalities? (2) Do elder individuals influence group-level responses more than other individuals? (3) How important is personality compared to other factors for group-level response variance? In the analyses in question 1, we used groups with at least two individuals (190 groups from the Approach test, 289 from Object test, and 197 from Sound test), and subset these observations for groups with at least one elder in question 2 (178 groups from the Approach test, 270 from Object test, and 151 from Sound test). In the analyses in question 3, we used all groups, including those with single individuals (sample size as described in Table 1).

We fitted GLMMs using a Bayesian Monte Carlo Markov Chain framework in the statistical package MCMCglmm (Hadfield 2010) to address our three questions. We specified priors for random and residual effects distribution, using default inverse-Wishart prior distributions (Hadfield 2018). We used a Gaussian family to model the trait of Mean Feeding Duration, and a categorical family with a logit link function to model the binary trait All Hopped Away. The residual variance was fixed at 1 for the All Hopped Away (binary) trait. The sample size of iterations stored was 1000 for all models, and we specified 2.6×10^4 iterations, 6×10^3 burning, and 2×10^2 thinning intervals. All models were checked for possible autocorrelations in the posterior distribution of the random variables and run to ensure autocorrelation values were below 0.1 (Hadfield 2018). Parameter estimates were presented as the posterior mean with 95% credible intervals. We considered statistical support for predictor variables if the 95% credible intervals did not overlap zero. We fitted the fixed effects of personality summaries (*mean-FID*, *maximum-FID*, or *minimum-FID*), group structure summaries (*adult sex ratio*, *proportion of 2yo+,*

proportion of associates, and *group size*, see definitions above), spatial structure (*easting*), and experimental variation (*stimulus proximity* and *test type*), and we fitted the random effect of *date* to account for temporal variation. We used a random intercept model with all fixed effects standardised and mean centred.

Question 1: Whose personality matters most in the group: average or extremes?

In question 1, we determined which personality summary (*mean-FID*, *maximum-FID*, or *minimum-FID*, see above) better predicted group-level responses. *Mean-FID* tested whether all group members contributed similarly with their personalities for the group-level response, and *maximum-FID* and *minimum-FID* tested whether the mean group-level response was better predicted by the most or the least responsive individual, respectively. As these personality summaries were highly correlated within groups, we fitted a model for each of them, resulting in three models for each of the two response traits (Mean-Feeding-Duration and All-Hopped-Away). The other predictors included in the models (see above) were not substantially correlated with each other (Pearson correlations <0.20), so we included them all together. We used the Deviance Information Criterion (DIC) to select the best personality summary in predicting group-level responses, with the best models having the smallest DIC. We initially ran models using data from the three different experimental tests (Approaching Human, Object, and Sounds) separately, but the results were very similar (Table S6), so we combined the group-level responses from different tests for simplicity.

Question 2: Do older individuals influence group-level responses more than other individuals?

In question 2, we compared the importance of two personality summaries: *all individuals' mean-FID* and *elders' mean-FID*. These predictors were also highly correlated, so we fitted them in separate models and compared the DIC values, as above. We chose the model with the smallest DIC value.

Question 3: How important is personality compared to other factors for group-level response variance?

We used two methods to quantify the relative amounts of variance explained by fixed effects in the models of the two group-level responses (Grömping 2007). The first method measures the change in marginal R^2 (ΔR^2), and the second a partial R^2 (Grömping 2007). Both methods quantify the relative importance of the variance explained by each fixed effect, but their meaning is slightly different. The first method excludes the covariance between the fixed effect of interest with the other predictors as it is assessed by the change in marginal R^2 when we drop the fixed effect of interest from the model. The second method considers the covariances of all fixed effects together, estimating the variance of all fixed together; see below for details of calculation. Therefore, the second approach is more comprehensive and allows us to sum up the variance of different predictors into classes. However this method can only be conducted for Gaussian traits. We estimated both ΔR^2 and partial R^2 for the Mean Feeding Duration trait, but we were restricted to the first method (ΔR^2) only for All Hopped Away.

To estimate a change in marginal R^2 (ΔR^2), we standardised and mean-centred the fixed effects. Marginal R^2 was estimated by dividing the variance of the fixed effects by the total variance – and we did that for a model with all variables in it, and again every time we dropped a predictor. The difference in marginal R^2 of a model after we dropped a predictor compared to the model with all predictors indicated the proportion of variance explained by that predictor alone (Froy et al. 2016). Calculations were made on the data scale for the Gaussian trait and the link scale for the binary trait (Nakagawa and Schielzeth 2013) (see code in Supplementary Materials).

Then, to estimate partial R^2, we standardised and mean-centred both fixed effects and the response variable (which is not possible for binary traits). When the response variable and the respective fixed effects are standardised and mean-centred, the parameter estimates are equal to partial correlation coefficients: the intercept was approximately 0, and all fixed effect estimates were between -1 and 1 (Grömping 2007). The square of the fixed effect parameters gave us a partial R^2 for each predictor. Partial R^2 is additive because we consider the correlation between all

fixed effects while estimating a variance component for each predictor – unlike ΔR^2, above. We could then sum the fixed effect variances into six classes: individual personality (*mean-FID*), group structure (*adult sex ratio* + *proportion of 2yo+ proportion of associates* + *group size*), spatial variation (*easting*), temporal variation (*date*), experimental variation (*stimulus proximity* + *test type*), and residual variation (*residuals*), to assess the relative importance of the different types of variable.

Part 2: among-group variation in individual's intrinsic FID

We used the individual's intrinsic FID and the group composition data to address a fourth question: (4) Are individuals grouped based on their personalities over and above spatial structure? Part 2 was a bit different because here the response variable was the individual's intrinsic FID, and the test was whether there was between-group variation. Individual's intrinsic FID was the individual FID-BLUP (best linear unbiased predictor (Robinson 1991)) from a mixed-effect linear model to account for age, habituation, and location effects (details of the model as in Chapter II), so slightly different from individual FID or mean FID that used raw behavioural measures, defined above. We fitted *age category* to correct for differences in behaviour between yearlings and 2yo+ individuals; *test number* and *days* to account for habituation over time, *year* to account for any temporal variation between years, *easting* to correct for spatial structure effects, and *observer* to correct for any observer effect (see Chapter II for more details on these variables). Then, we extracted the FID BLUP for each individual, which was treated as a cleaner measure of 'individual personality' after accounting for other intrinsic and extrinsic effects on FID. We merged the individual's intrinsic FID data to the group composition data from the population survey (different data set from above) for this part of the analyses.

Group composition data

We used records of group composition observations from detailed censuses noting the IDs of the associated individuals, their location, and group size; multiple observers conducted

these. Also different from above, we define a group as 2 or more individuals aged at least 24 months within the chain rule (group size range 2-20). We assigned a unique *group identity* for each group, including unique identities for groups of the same composition (i.e. same individuals) observed at different time points. We used group composition data from 2017-2019 and focused on foraging groups with at least two individuals, which makes up the majority of observations of groups (58.5%), and only involves excluding courtship, fighting males, mother-young, and single-individual type of groups. We conducted these exclusions because the association meaning in these groups was potentially different from those of foraging groups. We did not consider pairs of mothers with their <24-month young as they were usually associated with and behave similarly to each other (see Chapter IV). If the mother-offspring pair was grouped with other individuals, we considered that group but removed the <24-month young from the *group composition* and also did not count the young for the *group size*. The final data set consisted of 4,339 groups (mean group size 4.1 (±2.7 SD) individuals; Table 1).

In summary, we had a data file with the following variables:

- Individual's intrinsic FID (response variable): individual FID corrected for age, habituation and location effects.
- *Group size*: number of individuals aged at least 24 months in the group.
- *Group identity*: unique ID for groups of the same composition.

Statistical analysis

We fitted GLMMs as described above, using the Gaussian family for intrinsic FID. We fitted the fixed effect of *group size* in some models and the random effect of *group identity* in all models. We quantified the proportion of variance explained by group identity to test whether individuals in different groups differed in their personalities (in the form of FID). For this, we estimated '*among-group proportion*' (or 'group repeatability') by dividing the variance among groups by total variance. We considered statistical support for among-group proportion if the 95%

credible intervals were different from zero. In this question, we used a random intercept model with non-standardised fixed effects.

Question 4: Are individuals grouped based on their personalities over and above spatial structure?

We combined individual's intrinsic FID to independent group composition observations, which allowed us to test for consistent differences in personality over and above spatial structure among groups. Here, we present this in the form of among-group proportion in FID. We estimated among-group proportion in six models using different subsets of group size data as follows:

- All groups sizes (all-group model).
- Groups of 2 individuals only (2-individual group model).
- Groups of 3 individuals only (3-individual group model).
- Groups of 4 individuals only (4-individual group model).
- Groups of 2-4 individuals (smaller-group model).
- Groups of 5-20-individuals (larger-group model).

We compared the among-group proportion to see whether the importance of personality for group composition varied across groups of different sizes. In all-group, smaller-group, and larger-group class models, we used data from groups with different number of individuals in them, so we fitted the fixed effect of *group size*. With this, we tested whether different personalities tended to be found grouped with more or fewer conspecifics.

Table 1. Sample sizes and summary statistics for the complete behavioural data in eastern grey kangaroos in Victoria, Australia.

Variables	Data files		
	Part 1 (groups as used in question 3)		Part 2
	Object + Sound tests	Object + Sound + Approach tests	Group Composition observations
Trait value	Mean Feeding Duration (%): 48.333 (±21.142)	All Hopped Away (binary): 0.425 (±0.495)	Individual's Intrinsic FID: -0.112 (±2.461)
Sample size	619 groups (366 from Object test, and 253 from Sound test))	1,320 groups (366 from Object test, 253 from Sound test, and 701 from Approach test))	4,339 groups (14,712 individual observations)
Group size	2.651 (±1.469)	2.489 (±1.129)	4.070 (±2.7)
Adult sex ratio	0.224 (±0.313)	0.183 (±0.295)	0.213 (±0.409)
Proportion of 2yo+	0.872 (±0.203)	0.822 (±0.228))	-
Proportion of elders (4yo+)	0.650 (±0.338)	0.649 (±0.373)	-
Proportion of associates	0.603 (±0.429)	0.508 (±0.433)	-

In Part 1 (shaded in grey), we define a group as 2 or more individuals aged at least 12 months within the chain rule in questions 1-2, and that gives a group size range of 2-10, but in the analysis in question 3 we also include single individuals (sample sizes and summary statistics presented here are as in question 3). In Part 2, we define a group as 2 or more individuals aged at least 24 months within the chain rule, and that gives a group size range of 2-20. We show the mean followed by the standard deviation in brackets. '-' refers to no information.

Results

We present below the results addressing our four questions. The main results are summarised in Table 2 (for *personality summary* model selection, questions 1 and 2), Table 3 (for *variance relative importance of predictors*, question 3), and Table 4 (for *among-group proportion*, question 4); further details are available as Supplementary Materials (Tables S1-S6).

1. Whose personality matters most in the group: average or extreme personalities?

We found mean-FID, maximum-FID, and minimum-FID to be strongly correlated with each other within groups (Pearson correlations: mean-FID-maximum-FID: 0.88, mean-FID-minimum-FID: 0.84, and maximum-FID-minimum-FID: 0.50). All these FID summaries affected group-level responses significantly, with negative effects on Mean Feeding Duration and positive effects on All Hopped Away (Table 2A and Tables S1-S2): groups with higher FID values had lower feeding durations and higher chances of hopping away. Mean-FID was a better predictor of the group-level responses than maximum-FID or minimum-FID, having the smallest model DIC (Table 2A). Considering the non-standardised values (Table 3A), groups with a mean-FID of 5 m would feed for 11.1% longer than groups with a mean-FID of 10 m (Figure 1A), and the mean-FID of groups that all hopped away was 7.5 m compared to 6.6 m for groups that did not (Figure 1B).

2. Do elders influence group-level responses more than other individuals?

All individuals' mean-FID and elders' mean-FID were highly correlated within groups (Pearson correlation: 0.84), and both predictors contributed significantly to group-level responses. As above, the mean-FID of all individuals was a better predictor than the elders' mean-FID (Table 2B and Tables S3-S4). There was therefore no indication that the personality of elders was more important than that of all individuals.

3. How important is personality compared to other factors for mean group-level response variance?

Mean-FID significantly predicted both group-level responses (Table 3A), but there was also support for spatial variation effects on the Mean Feeding Duration and for group demographic structure effects on All Hopped Away. Groups observed 1 km east of the western edge of the study area spent 8% less time feeding compared to groups at the western edge (Table 3A). Groups tested in the Object trial spent on average 10% less time feeding than groups tested with sounds.

When we tested for the relative magnitude of variance explained by each fixed effect variable, we found mean-FID to explain more variance than did fixed effects reflecting group demographic structure or spatial variation: mean-FID was associated with ΔR^2 of 5.60% in Mean Feeding Duration and 6.16% in All Hopped Away, larger than any other variable (Table 3A). When we summed up the variance explained by all the group demographic structure variables in the Mean Feeding Duration (Table 3B), their total was still only 1.19%, less than the 5.67% for individual personality – which therefore remained the most important predictor of trait variance. The results from the two statistical methods used to quantify fixed effect variance (Table 3A versus Table 3B) were similar. Still, the second method was more robust as it allowed for a correlation between all predictors and estimated 95% credible intervals. The variance attributed to experiment variation was relatively high, especially for All Hopped Away (26.30%), but that is not unexpected as we had combined traits from different experiments.

4. Are individuals grouped based on their personalities over and above spatial structure?

In the final question, we considered models with individuals' intrinsic FID as the response variable, to test whether groups differed in intrinsic FIDs. Groups of two kangaroos were the most frequent, and then the frequency decreased as group size increased (Figure 2B). We found no support for the fixed effect of group size (95% credible intervals overlapped zero,

Table S5), which indicates that different personalities were not found in groups of different sizes. We found support for among-group proportion estimates in intrinsic FID: 13.14% of the total variance in FID was due to differences amongst groups (Figure 2B and Table S5). Smaller groups were more frequent than larger groups (Figure 2B), and smaller groups had higher among-group proportion estimates (Figure 2A-B): 16.76% and 6.65% when comparing group size 2-4 to groups of 5-20 individuals, respectively. Groups of two individuals had the highest among-group proportion (22.6%).

Table 2. Results from the different models testing which FID summaries better predicted group-level responses (Mean Feeding Duration and All Hopped Away) in eastern grey kangaroos, questions 1-2.

(A) Question 1: Whose personality matters most in the group?			
Trait	Personality summaries	Model DIC	Table (Supplementary Materials)
Mean Feeding Duration	**Mean-FID**	**4239.321**	S1 (A)
	Maximum-FID	4245.852	S1 (B)
	Minimum-FID	4247.599	S1 (C)
All Hopped Away	**Mean-FID**	**753.2209**	S2 (A)
	Maximum-FID	760.7052	S2 (B)
	Minimum-FID	754.6496	S2 (C)
(B) Question 2: Do elders influence group-level responses more than other individuals?			
Trait	Personality summaries	Model DIC	Table
Mean Feeding Duration	**Mean-FID (all individuals)**	**3932.491**	S3 (A)
	Mean-FID (Elders only)	3939.355	S3 (B)
All Hopped Away	**Mean-FID (all individuals)**	**696.9772**	S4 (A)
	Mean-FID (Elders only)	699.3102	S4 (B)

*(**A**) shows a comparison between mean-FID, maximum-FID, and minimum-FID (Question 1), reflecting the influence of all individuals' personalities and of the most or the least responsive individuals, respectively. (**B**) shows a comparison between all individuals' mean-FID versus elders' mean-FID (Question 2). We standardised and mean-centred our predictors (see above), so that parameter estimates from different models were on the same scale. We chose the models (in bold) with the smallest Deviance Information Criterion (DIC) values.*

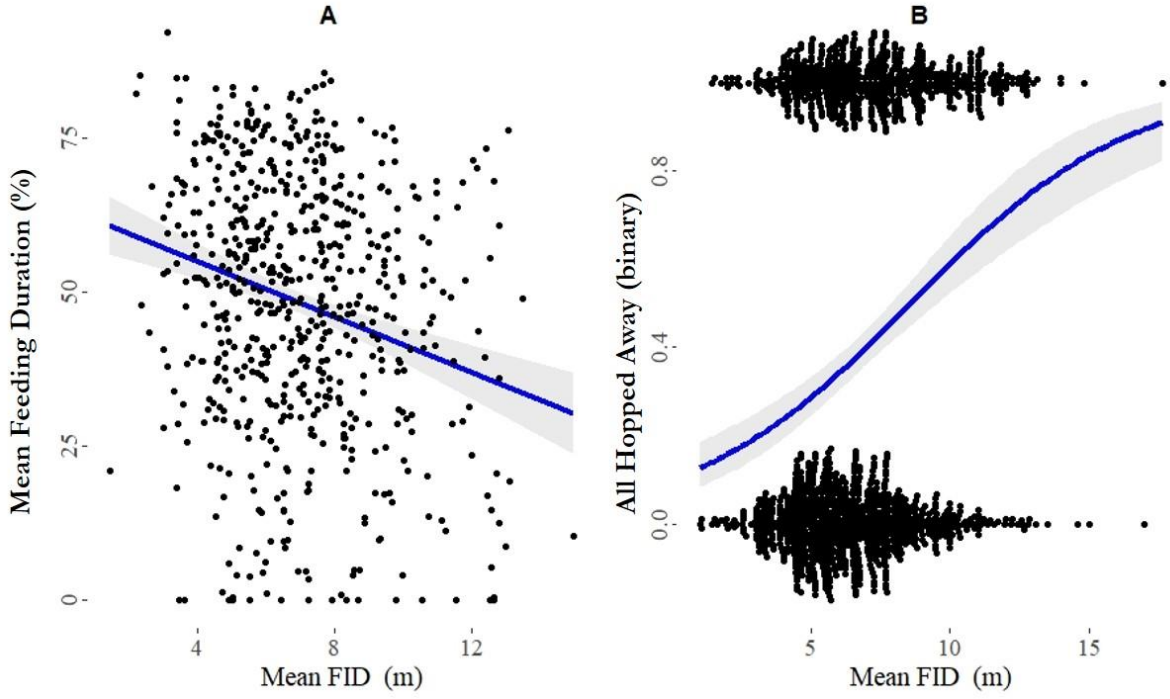


Figure 1. Relationship between mean FID and group-level responses, question 3.

*(**A**) Mean Feeding Duration (n: 619 groups), and (**B**) All Hopped Away (n: 1320 groups). Dots show the raw data (non-standardised), blue lines show the regression curve, and shade areas indicate the 95% credible intervals.*

Table 3. Results from the two methods used to quantify the relative importance of fixed effects (change in marginal R^2, or ΔR^2, and partial R^2), question 3.

(A) Parameter estimates from a model in the original/data scale, followed by a change in marginal R^2 (ΔR^2), estimated in models using standardised predictors

Variable		Mean Feeding Duration		All Hopped Away	
		Estimate (95% CI)	ΔR^2(%)	Estimate (95% CI)	ΔR^2(%)
		Marginal R^2 = 12.49%		Marginal R^2 =56.40%	
Fixed Effect	Intercept	9.120 (-34.112, 52.872)	.	1.238 (-2.305, 4.723)	.
	Mean-FID (m)	-2.227 (-2.983, -1.432)*	-5.60	0.253 (0.191, 0.327)*	-6.16
	Adult Sex Ratio	-3.634 (-9.027, 1.740)	-0.21	-0.468 (-0.983, 0.056)	-1.23
	Group Size	0.953 (-0.264, 2.203)	-2.41	-0.389 (-0.551, -0.249)*	-5.11
	Proportion of Adults	-3.335 (-10.814, 5.953)	-0.98	0.404 (-0.292, 1.105)	-2.96
	Proportion of Associates	0.830 (-4.164, 4.554)	-0.49	0.586 (0.174, 0.927)*	-3.44
	Easting	0.008 (0.002, 0.014)*	-1.57	0.000 (-0.001, 0.000)	-1.72
	Stimulus proximity	-3.686 (-12.095, 5.780)	-0.63	1.303 (0.875, 1.697)*	-6.58
	Test type (Sound)	(base)	-4.33	-2.276 (-2.816, -1.850)*	-26.30
	Test type (Object)	-10.359 (-13.854, -6.464)*		-1.429 (-1.861, -1.014)*	
Random Effect (variance)	Date	3.245 (0.000, 16.220)	.	0.189 (0.000, 0.512)	.
	Residual	400.6 (352.5, 446.3)	.	Fixed to 1	.

(B) Partial R^2 (summed up into classes), estimated from a model using standardised predictors and response

Predictor Class	Mean Feeding Duration	All Hopped Away
	Partial R^2 (%)	
Individuals' personality	5.67 (1.79, 9.66)	|
Group demographic structure	1.19 (0.04, 3.15)	|
Spatial variation	1.11 (0.000, 3.78)	|
Experimental variation	4.44 (2.44, 11.03)	|
Temporal variation	0.12 (0.01, 4.68)	|
Residual	87.75 (78.31, 98.83)	|

*Here, we show the estimates and their respective 95% credible intervals on the original data scale, but marginal R^2 and partial R^2 were estimated from models with standardised and mean-centred predictors (as in Tables S1-S4). (**A**) shows the fixed and random effect estimates, in the original data scale, and the $\perp R^2$ (as a %) estimated from separated models that used standardized and mean-centred predictors. (**B**) shows the partial R^2 summed up into classes (see main text) estimated from a model using standardised and mean-centred predictors and response. The posterior mean is followed by the respective 95% credible intervals. Base levels: Test type ('Sound' for Mean*

Feeding Duration and 'Approach' for All Hopped Away). Bold indicates the highest predictor variance, comparing individuals' personality, group structure factors, and spatial variation (shaded in grey). '' indicates statistical support. '.' means not estimated. '|' means not estimable. Codes for estimating Marginal or Partial R^2 are available as supplementary materials.*

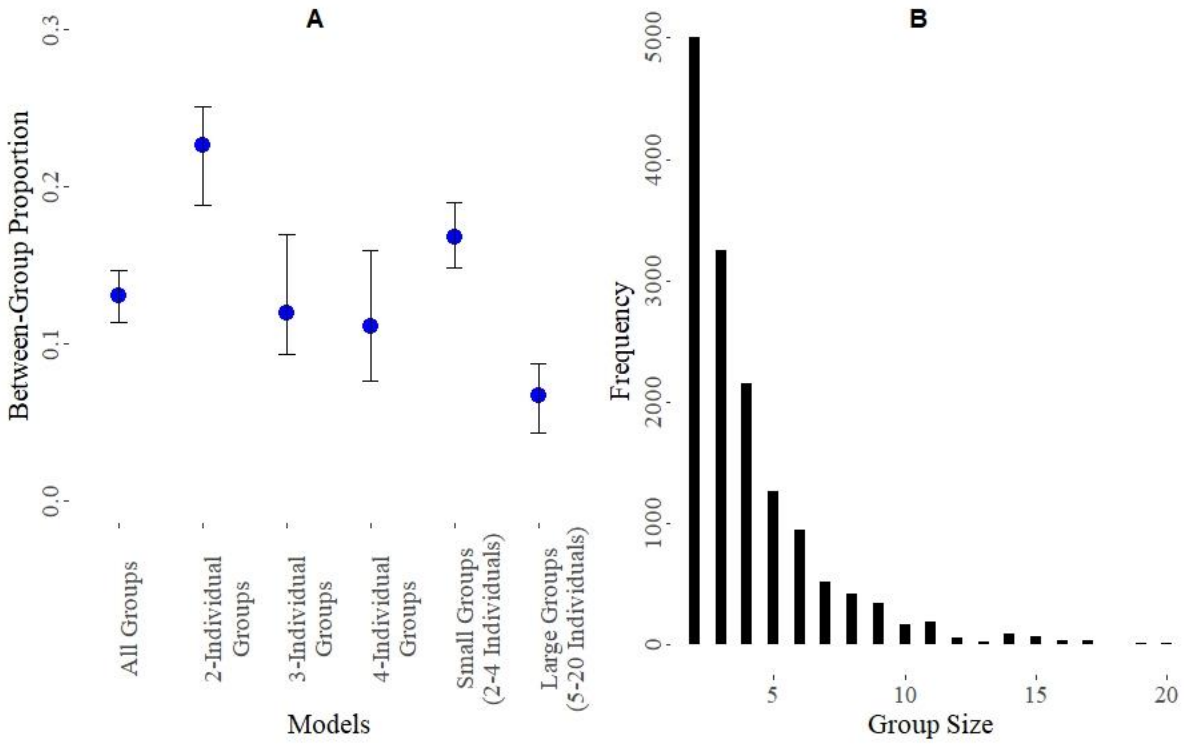


Figure 2. Summary of among-group proportion estimates from models considering different subsets of group size, over and above spatial-temporal structure effects, question 4.

*See Table S5 for more details. (**A**) among-group proportion estimates across models considering groups of different sizes: all groups (group size 2-20, n: 4339), 2-individual group data only (n: 2085 groups), 3-individual groups (n: 1079), 4-individual groups (n: 539), smaller groups (group size 2-4, n: 3703), and larger groups (group size 5-20, n: 636). Blue dots indicate the posterior mode, and bars show the 95% credible intervals. (**B**) foraging group size-frequency from August 2017-April 2019 (excluding groups with an individual alone). Each bar shows the frequency for each group size, starting from 2-individual groups on the left.*

Discussion

Our analyses indicated that the average of all group members' FID was the best predictor of group-level responses, and that this was a better predictor than the most or the least responsive individuals' mean FID, or the average of elders' FIDs. The results therefore indicate a lack of leadership within eastern grey kangaroo groups. We found support for the effects of mean FID and easting on Mean Feeding Duration, indicating the importance of personality and location. Mean FID, group size, and proportion of associates significantly affected All Hopped Away, indicating the importance of personality and group structure. When considering the relative importance of variance explained by the predictors, we found that the average group members' FID was more important than group demographic structure or spatial-temporal variation for the total variances of the group-level responses. Finally, we found evidence for between-group proportion estimates in FID, over and above spatial-temporal structure, indicating that personality contributes to group composition.

Our results suggest that eastern grey kangaroos contributed 'democratically' (Conradt and Roper 2007) to the two group-level responses analysed here, with less support for a disproportionate influence of extreme personalities (highest or lowest score) or elder individuals. This differs somewhat from conclusions in other studies. Extreme personalities had a disproportionate influence on group-level responses in two species in laboratory conditions (Brown and Irving 2014; Planas-Sitjà 2020). Group exploratory behaviour in feral guppies (*Poecilia reticulata*) is not affected by an average group members' personality score but is predicted by both the lowest individual activity score and the highest sociability score (Brown and Irving 2014). Aggregation behaviour in cockroaches (*Periplaneta americana*) seems to be driven by highly exploratory individuals rather than by other personality types (Planas-Sitjà 2020). Some mammalian studies have found that elder individuals tend to undertake leadership roles in collective movement (Sueur and Petit 2008; McComb et al. 2011; Brent et al. 2015; Tokuyama and Furuichi 2017; Allen et al. 2020). These studies have not always considered the average of

group members' personality the same way as we did here, and with our results, extreme and elders are still important, but just not as important than the average of all group members' personalities. In similar results to our study, Tonkean macaques (*Macaca tonkeana*) of all ages were found to equally influence group-level decisions to move (Sueur and Petit 2008).

In Chapter II, we found that group differences were more substantial than individual differences in behaviour, suggesting that the group social environment explained more behavioural variance than individual personality. Here, we sought to understand the factors that would best contribute to group-level responses, and hence, be the main factors driving group behavioural differences. We compared the relative magnitude effect of group members' personalities versus group structure factors, such as adult sex ratio, age proportion, and the proportion of close associates, and spatial-temporal structure. We found that the average personality of group members accounted for a greater magnitude of the variance than these other variables, indicating that variation in personality was the central factor underpinning group behavioural differences. This finding highlights the substantial ecological importance of personality to sociality (Sih et al. 2004; Réale et al. 2007). Individuals have personalities, but when they experience the personality expression of the others in the group, they may adjust behaviourally to each other (King and Cowlishaw 2007; Stevens et al. 2016). Thus, their behaviour will be more consistent with the other group members' behaviour at the time of the observation than with its future or former behaviour as they change social groups. The average personality of group members is the baseline for group-level behaviour, so individuals in different groups having different personalities may be the main factor associated with group behavioural differences. These results are particularly critical in eastern grey kangaroos, as their group composition changes every 9 minutes on average for groups with at least an adult female (King and Goldizen 2016).

We found that individual personality also had effects on group composition, observed via consistent among-group differences in personality across three years, over and above spatial-

temporal structure. Individuals can be predisposed to assort with others that behave similarly (Massen and Koski 2014; Jolles et al. 2018), with assortment resulting from active social choices or a passive outcome of fission-fusion dynamics (Muller et al. 2018). As different personalities may acquire or use information about the environment differently (Coppens et al. 2010; Wolf and Krause 2014), having divergent preferences or needs (Gavrilets et al. 2016), more dissimilar personalities may tend to be previously separated into different groups over time (Merkle et al. 2015). Previous studies have shown that kangaroo social structure is not random (Carter et al. 2009) but also is not strongly affected by kinship (Best et al. 2014; King et al. 2015b), with associations correlated with spatial structure (Best et al. 2014). Our study extends that knowledge, showing that personality differences significantly influence eastern grey kangaroos' social organisation. To date, to our knowledge, only two other species studies have been studied to investigate the effect of personality on association patterns in the wild: great tits (Aplin et al. 2013) and bottlenose dolphins (*Tursiops truncates*) (Díaz López 2020). In the laboratory, among-group differences in personality have been reported in sticklebacks (Herbert-Read 2017; Jolles et al. 2017, 2018), and personality has also been shown to underpin network metrics in Trinidadian guppies (*Poecilia reticulata*) (Croft et al. 2009) and sticklebacks (Pike et al. 2008). Personality also contributes to social bonding in captive chimpanzees (*Pan troglodytes*) (Massen and Koski 2014). Here, we have not found support individuals of different personalities to be found in groups of different sizes. This conclusion differs from those from a study in female eastern grey kangaroos in a population in southern Queensland, where more reactive females preferred larger groups (Best et al. 2015). However, it is worth noting that this other study was based on 51 unmarked individuals and the overall effect was dependent on a single shy female foraging in a larger group (Best et al. 2015, Figure 2).

We found smaller groups to have stronger among-group variance in personality compared to larger groups (Figure 2). It is possible that this result is an outcome of our definition of a 'group' (i.e., we defined a group based on individuals positioned within 10 m of

each other (Jarman 1987)), or highlights a difference in the ecological meaning of a smaller versus a larger group. By the 10-m-chain rule (Jarman 1987) group definition, some individuals will be relatively closer or more distant from each other in larger groups. Still, we would consider all those individuals to be associated. As individuals assort with others of similar personality, it is possible that, in larger groups, individuals are more strongly associated with those immediately close to them. We were not able to test whether the nearest neighbours were more similar in personality, so this remains an open and interesting question. Alternatively, larger groups are much less frequent and can occur as a result of disturbance – hence, they could be seen as an antipredator response themselves (Creel et al. 2014). Therefore, under disturbance, the ecological meaning of social associations within a group could change. A higher number of associates may be more important than specific individual characteristics ('safety in numbers or dilution effect' (Foster 2017)). Future studies using social networks, confronting analysis on smaller groups against complete data will be beneficial to understand these issues better.

In conclusion, this study addressed key questions about the processes shaping group-level behaviour in a wild marsupial. We found that the average group members' personality scores were the best predictor of group-level responses and also showed that personality variation played an even more important role in shaping group differences in behaviour than group structure or spatial-temporal structure variables. We also showed that personality contributed to association patterns, with individuals grouped with others with similar behavioural tendencies, over and above spatial-temporal structure. Further studies using social networks or other quantitative statistical tools to test the relationship between individual attributes, association preferences, and social dynamics will contribute to understanding social structure in eastern grey kangaroos.

References

Allen CRB, Brent LJN, Motsentwa T, et al (2020) Importance of old bulls: leaders and followers in collective movements of all-male groups in African savannah elephants (Loxodonta africana). Sci Rep 10:. https://doi.org/10.1038/s41598-020-70682-y

Aplin LM, Farine DR, Morand-Ferron J, et al (2013) Individual personalities predict social behaviour in wild networks of great tits (Parus major). Ecol Lett 16:1365–1372. https://doi.org/10.1111/ele.12181

Bengston SE, Jandt JM (2014) The development of collective personality: The ontogenetic drivers of behavioral variation across groups. Front. Ecol. Evol. 2:81

Best EC, Blomberg SP, Goldizen AW (2015) Shy female kangaroos seek safety in numbers and have fewer preferred friendships. Behav Ecol 26:639–646. https://doi.org/10.1093/beheco/arv003

Best EC, Dwyer RG, Seddon JM, Goldizen AW (2014) Associations are more strongly correlated with space use than kinship in female eastern grey kangaroos. Anim Behav 89:1–10. https://doi.org/10.1016/j.anbehav.2013.12.011

Best EC, Seddon JM, Dwyer RG, Goldizen AW (2013) Social preference influences female community structure in a population of wild eastern grey kangaroos. Anim Behav 86:1031–1040. https://doi.org/10.1016/j.anbehav.2013.09.008

Brent LJN, Franks DW, Foster EA, et al (2015) Ecological knowledge, leadership, and the evolution of menopause in killer whales. Curr Biol 25:746–750. https://doi.org/10.1016/j.cub.2015.01.037

Brown C, Irving E (2014) Individual personality traits influence group exploration in a feral guppy population. Behav Ecol 25:95–101. https://doi.org/10.1093/beheco/art090

Carter AJ, Macdonald SL, Thomson VA, Goldizen AW (2009) Structured association patterns and their energetic benefits in female eastern grey kangaroos, *Macropus giganteus*. Anim Behav 77:839–846. https://doi.org/10.1016/j.anbehav.2008.12.007

Conradt L, Roper T. (2007) Democracy in animals: the evolution of shared group decisions. Proc R Soc B Biol Sci 274:2317–2326. https://doi.org/10.1098/rspb.2007.0186

Conradt L, Roper TJ (2005) Consensus decision making in animals. Trends Ecol. Evol. 20:449–456

Coppens CM, de Boer SF, Koolhaas JM (2010) Coping styles and behavioural flexibility: towards underlying mechanisms. Philos Trans R Soc B Biol Sci 365:4021–4028. https://doi.org/10.1098/rstb.2010.0217

Couzin ID (2006) Behavioral ecology: social organization in fission–fusion societies. Curr Biol 16:R169–R171. https://doi.org/10.1016/J.CUB.2006.02.042

Couzin ID, Krause J (2003) Self-organization and collective behavior in vertebrates. Adv Study Behav 32:1–75. https://doi.org/10.1016/S0065-3454(03)01001-5

Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. Nature 433:513–516. https://doi.org/10.1038/nature03236

Cox JA, Cusick JA, DuVal EH (2019) Manipulated sex ratios alter group structure and cooperation in the brown-headed nuthatch. Behav Ecol 30:883–893. https://doi.org/10.1093/beheco/arz030

Creel S, Schuette P, Christianson D (2014) Effects of predation risk on group size, vigilance, and foraging behavior in an African ungulate community. Behav Ecol 25:773–784. https://doi.org/10.1093/beheco/aru050

Croft DP, Krause J, Darden SK, et al (2009) Behavioural trait assortment in a social network: patterns and implications. Behav Ecol Sociobiol 63:1495–1503. https://doi.org/10.1007/s00265-009-0802-x

Dall SRX, Griffith SC (2014) An empiricist guide to animal personality variation in ecology and evolution. Front Ecol Evol 2:1–7. https://doi.org/10.3389/fevo.2014.00003

Díaz López B (2020) When personality matters: personality and social structure in wild bottlenose dolphins, Tursiops truncatus. Anim Behav 163:73–84. https://doi.org/10.1016/j.anbehav.2020.03.001

Dreiss AN, Cote J, Richard M, et al (2010) Age- and sex-specific response to population density and sex ratio. Behav Ecol 21:356–364. https://doi.org/10.1093/beheco/arp198

Farine DR, Montiglio PO, Spiegel O (2015) From individuals to groups and back: the evolutionary implications of group phenotypic composition. Trends Ecol. Evol. 30:609–621

Foster W (2017) Safety in numbers. J Exp Biol 220:4551–4553. https://doi.org/10.1242/jeb.160721

Froy H, Walling CA, Pemberton JM, et al (2016) Relative costs of offspring sex and offspring survival in a polygynous mammal. Biol Lett 0–3. https://doi.org/10.1098/rsbl.2016.0417

Gavrilets S, Auerbach J, Van Vugt M (2016) Convergence to consensus in heterogeneous groups and the emergence of informal leadership. Sci Rep 6:. https://doi.org/10.1038/srep29704

Grömping U (2007) Estimators of relative importance in linear regression based on variance decomposition. Am Stat 61:139–147

Hadfield J (2018) MCMCglmm Course Notes

Herbert-Read JE (2017) Social behaviour: the personalities of groups. Curr Biol 27:R1015–R1017. https://doi.org/10.1016/j.cub.2017.07.042

Hodgkin LK, Symonds MRE, Elgar MA (2017) Leadership through knowledge and experience in a social sawfly. Anim Behav 134:177–181. https://doi.org/10.1016/j.anbehav.2017.10.017

Jarman PJ (1987) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Jolles JW, Boogert NJ, Sridhar VH, et al (2017) Consistent individual differences drive collective behavior and group functioning of schooling fish highlights. Curr Biol 27:2862-2868.e7. https://doi.org/10.1016/j.cub.2017.08.004

Jolles JW, King AJ, Killen SS (2020) The role of individual heterogeneity in collective animal behaviour. Trends Ecol Evol 35:278–291. https://doi.org/10.1016/j.tree.2019.11.001

Jolles JW, Laskowski KL, Boogert NJ, Manica A (2018) Repeatable group differences in the collective behaviour of stickleback shoals across ecological contexts. Proc R Soc B Biol Sci 285:20172629. https://doi.org/10.1098/rspb.2017.2629

Kappeler PM (2017) Sex roles and adult sex ratios: Insights from mammalian biology and consequences for primate behaviour. Philos. Trans. R. Soc. B Biol. Sci. 372

King AJ, Cowlishaw G (2009) Leaders, followers and group decision-making. Commun. Integr. Biol. 2:147–150

King AJ, Cowlishaw G (2007) When to use social information: the advantage of large group size in individual decision making. Biol Lett 3:137–139. https://doi.org/10.1098/rsbl.2007.0017

King WJ, Forsyth DM, Coulson G, Festa-Bianchet M (2015a) Adoption in eastern grey kangaroos: a consequence of misdirected care? PLoS One 10:e0125182. https://doi.org/10.1371/journal.pone.0125182

King WJ, Garant D, Festa-Bianchet M (2015b) Mother-offspring distances reflect sex differences in fine-scale genetic structure of eastern grey kangaroos. Ecol Evol 5:2084–2094. https://doi.org/10.1002/ece3.1498

King WJ, Goldizen AW (2016) Few sex effects in the ontogeny of mother-offspring relationships in eastern grey kangaroos. Anim Behav 113:59–67. https://doi.org/10.1016/j.anbehav.2015.12.020

King WJ, Wilson ME, Allen T, et al (2011) A capture technique for free-ranging eastern grey kangaroos (*Macropus giganteus*) habituated to humans. Aust Mammal 33:47–51. https://doi.org/10.1071/AM10029

Lee HC, Teichroeb JA (2016) Partially shared consensus decision making and distributed leadership in vervet monkeys: older females lead the group to forage. Am J Phys Anthropol 161:580–590. https://doi.org/10.1002/ajpa.23058

Mackay AE, Forsyth DM, Coulson G, Festa-Bianchet M (2018) Maternal resource allocation adjusts to timing of parturition in an asynchronous breeder. Behav Ecol Sociobiol 72:7. https://doi.org/10.1007/s00265-017-2419-9

Massen JJM, Koski SE (2014) Chimps of a feather sit together: Chimpanzee friendships are based on homophily in personality. Evol Hum Behav 35:1–8. https://doi.org/10.1016/j.evolhumbehav.2013.08.008

McComb K, Shannon G, Durant SM, et al (2011) Leadership in elephants: the adaptive value of age. Proc R Soc B Biol Sci 278:3270–3276. https://doi.org/10.1098/rspb.2011.0168

Merkle JA, Sigaud M, Fortin D (2015) To follow or not? How animals in fusion-fission societies handle conflicting information during group decision-making. https://doi.org/10.1111/ele.12457

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Montiglio PO, Ferrari C, Réale D (2013) Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. Philos. Trans. R. Soc. B Biol. Sci. 368

Muller Z, Cantor M, Cuthill IC, Harris S (2018) Giraffe social preferences are context dependent. Anim Behav 146:37–49. https://doi.org/10.1016/j.anbehav.2018.10.006

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R 2 from generalized linear mixed-effects models. Methods Ecol Evol 4:133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Pike TW, Samanta M, Lindström J, Royle NJ (2008) Behavioural phenotype affects social interactions in an animal network. Proc R Soc B Biol Sci 275:2515–2520. https://doi.org/10.1098/rspb.2008.0744

Planas-Sitjà I (2020) Personality variation improves collective decision-making in cockroaches. Behav Processes 177:. https://doi.org/10.1016/j.beproc.2020.104147

Planas-Sitjà I, Nicolis SC, Sempo G, Deneubourg J-L (2018) The interplay between personalities and social interactions affects the cohesion of the group and the speed of aggregation. PLoS One 13:e0201053. https://doi.org/10.1371/journal.pone.0201053

Réale D, Reader SM, Sol D, et al (2007) Integrating animal temperament within ecology and evolution. Biol Rev 82:291–318. https://doi.org/10.1111/j.1469-185X.2007.00010.x

Robinson GK (1991) That BLUP is a good thing: The estimation of random effects. Stat Sci 6:15–32. https://doi.org/10.1214/ss/1177011926

Schürch R, Rothenberger S, Heg D (2010) The building-up of social relationships: behavioural types, social networks and cooperative breeding in a cichlid. Philos Trans R Soc B Biol Sci 365:4089–4098. https://doi.org/10.1098/rstb.2010.0177

Sih A, Bell A, Johnson JC (2004) Behavioral syndromes: An ecological and evolutionary overview. Trends Ecol Evol 19:372–378. https://doi.org/10.1016/j.tree.2004.04.009

Stevens JR, Leca J-B, Duboscq julieam J, et al (2016) Social information transmission in animals: lessons from studies of diffusion. https://doi.org/10.3389/fpsyg.2016.01147

Strandburg-Peshkin A, Papageorgiou D, Crofoot MC, Farine DR (2018) Inferring influence and leadership in moving animal groups. Philos. Trans. R. Soc. B Biol. Sci. 373

Sueur C, Petit O (2008) Shared or unshared consensus decision in macaques? Behav Processes 78:84–92. https://doi.org/10.1016/j.beproc.2008.01.004

Tokuyama N, Furuichi T (2017) Leadership of old females in collective departures in wild bonobos (Pan paniscus) at Wamba. Behav Ecol Sociobiol 71:1–10. https://doi.org/10.1007/s00265-017-2277-5

Webster MM (2017) Experience and motivation shape leader–follower interactions in fish shoals. Behav Ecol 28:77–84. https://doi.org/10.1093/beheco/arw133

Weston MA, Mcleod EM, Blumstein DT, Guay PJ (2012) A review of flight-initiation distances and their application to managing disturbance to Australian birds. Emu - Austral Ornithol. 112:269–286

Wolf M, Krause J (2014) Why personality differences matter for social functioning and social structure. Trends Ecol. Evol. 29:306–308

Wolf M, Van Doorn GS, Weissing FJ (2011) On the coevolution of social responsiveness and behavioural consistency. Proc R Soc B Biol Sci 278:440–448. https://doi.org/10.1098/rspb.2010.1051

Wolf M, Weissing FJ (2012) Animal personalities: consequences for ecology and evolution. Trends Ecol Evol 27:452–461. https://doi.org/10.1016/j.tree.2012.05.001

Supplementary Materials

1. Experimental tests

1.1 Responses to an Approaching Human

Approaching human (Approach test) is one of the most common experiments to test individual behavioural differences in the wild (Strong et al., 2017). Trials began when individuals in the group were feeding. We started from a distance of 30 m to avoid disturbing animals, using a tripod to mark the initial point and walked steadily and directly towards a target individual until it took flight. Then, we measured the distance walked by the observer (in m) using a range finder (Bushnell ® 4x20, Truth With Clear Shot) and calculated the flight intiation distance (FID) as 30 m minus the distance walked from the tripod (Figure S1.2). We also recorded whether non-target individuals responded by hopping away following the test on the target individual. We tested individuals in group sizes of one up to ten; for those in groups, we aimed for an individual in the peripheral area of the group. FID was used as a proxy for personality as it is highly repeatable and covaries with other traits across individuals (Chapter II). Hopped away (from non-target individuals) was used as a group response in this manuscript.

Figure S1.2 The Approaching human experiment

1.2 Responses to a Foreign Object

We used a remote-controlled (RC) car (HSP® 1/10 Pangolin Electric 4WD RTR RC Rock Crawler) as a 'foreign object' because kangaroos did not seem to react to stationary visual

objects (Object test). The RC car also emitted a noise while moving, which was useful in attracting the kangaroos' attention. We installed the tripod with the camera attached about 30 m from a target group, placed the RC car approximately 15 – 20 m from the target group, identified the individuals, measure initial distances using a range finder, then drove the RC car parallel to the main axis of the group and followed the individuals in the video for three minutes (Figure S1.3). We stopped moving the object when all individuals became vigilant, i.e. raised their heads from feeding. In a few trials, we could not keep all the individuals in the video for the complete three minutes, so we coded their behaviour for as long as possible. We noted whether individuals hopped away or not, in the field, and used videos to measure the duration of feeding. We then used percentages out of the length of time observed. We used the software Solomon Coder beta 17.03.22 (https://solomoncoder.com/) to code individual behaviour from the videos of the foreign object and novel sound trials. We aimed for "blind" scoring by viewing the videos in black and white, which reduced the chances of recognising individuals by their collar or tag colour. During the video coding, individuals were assigned an A-Z generic ID from left to right. Identities were reassigned after video scoring

Figure S1.3 The Foreign Object experiment

1.3 Responses to Novel Sounds

We played back at kangaroos thirteen 20-sec samples of different instrumental pieces (classical or rock, involving instruments such as string, brass or percussion) (Sound test). We selected sections with little variation in sound amplitude. We used the Audacity® 2.1.2 program (http://audacityteam.org/) to apply a high-pass filter (100 Hz cut-off, 36 dB roll-off) and normalise the peak amplitude to -5 dB, removing DC offset (centre on 0.0 vertically), normalising stereo channels independently. Maximum sound amplitude was set to 49 ± 1 dBA at 1 metre using a Digitech® QM-1589 compact digital sound level meter (high-level range, A-frequency weighting, and fast time weighting). We chose this maximum amplitude during trial testing because it led to more variation in response in comparison with lower or higher amplitudes. In the field, the observer placed a black Bluetooth speaker (Braven® BRV-X Waterproof) at approximately 20 m from a target group, facing the sound projector towards the individuals. The experimental set-up (Figure S1.4) was similar to that of the foreign object test, with the observer and video camera stationed about 25 m from the group. Playback tracks were played once from a smartphone (Samsung® Galaxy S5), and the target group was videotaped for 3 minutes. We noted whether individuals hopped away or not, in the field, and recorded the duration of feeding (as a percentage) from videos (as above).

Figure S1.4 The Novel Sounds experiment

1.4 Defining Group Size

We used the '10-m chain rule' (Jarman, 1987) to define a group. This rule includes all individuals within 10 m of at least one other individual in the same group (Figure S1.5). Group sizes: 2 (Group A), 5 (Group B), and 1 (Group C).

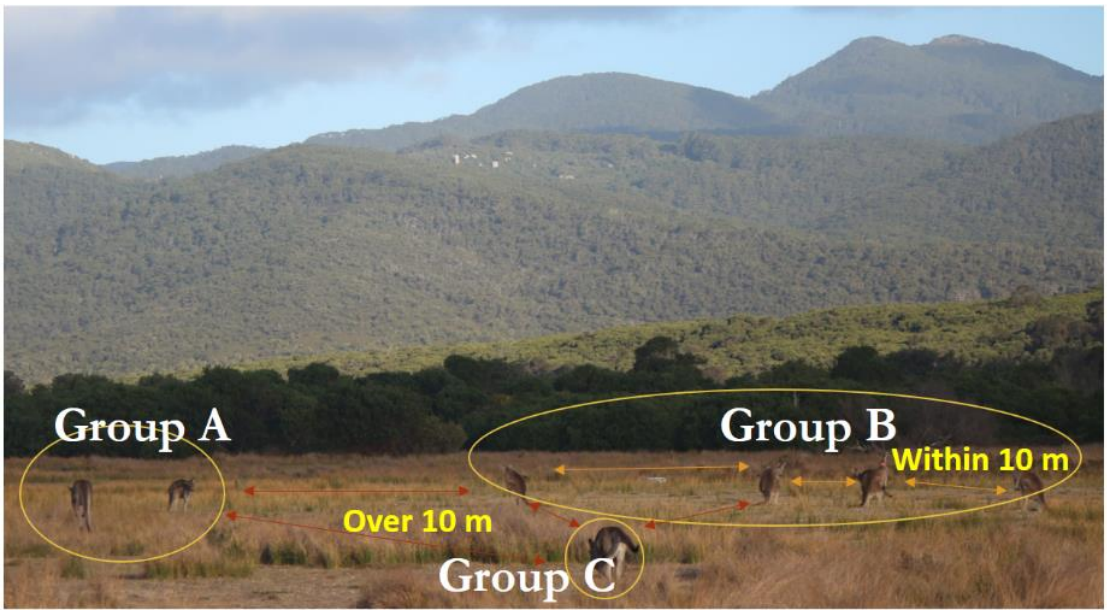


Figure S1.5 Defining group sizes

Data preparation (for statistics in Question 3)

For estimating the change in Marginal R^2 (ΔR^2), standardise and mean centre all fixed effects, and to estimate Partial R^2 standardise and mean centre fixed effects and the response. If a fixed effect is categorical, transform it into numeric (e.g., make it binary as we did for *test*)

#Make categorical variable binary

```
dataA$TestBi_Ob<-ifelse(dataA$Test=="dtObj",1,0)
dataA$TestBi_So<-ifelse(dataA$Test=="dtSound",1,0)
# dataA$TestBi_Fi<-ifelse(dataA$Test=="dtFID",1,0) # you don't need to do this for all levels, as one #will be automatically the base-level, so do not need to include it as a binary trait in the model
```

#Standardise and mean centre variables

```
dataA$Stdzd_GroupSize<-(dataA$GroupSize - mean(dataA$GroupSize, na.rm=TRUE))/sd(dataA$GroupSize, na.rm=TRUE)
```

Codes for estimating Marginal R^2

```
# Fit a model in MCMC  glmm (standardised fixed effects, but no response)

library(MCMCglmm)

# Model with all fixed effects####

prior2 <- list(G = list(G1 = list(V = 1, nu = 0.002)),
          R = list(V = 1,nu=0.002)) # fix=1 if the trait is binary


MCMCGroup_DuraFeed_Mean_Question3_2_allfixedeff <- MCMCglmm(AvGDurFeed ~ 1+
                                                  Stdzd_GMeanFID +
                                                  Stdzd_AduSexRatio +
                                                  Stdzd_GroupSize +
                                                  Stdzd_PropGAdu +
                                                  Stdzd_PropGAssoci +
                                                  Stdzd_Easting +
                                                  Stdzd_PropExposure+
                                                  Stdzd_TestBi,
                                                random = ~Date,
                                                data = dataA,

                                          nitt=1300*times,thin=1*times,burnin=300*times,
                                          family="gaussian", prior = prior2,
                                                verbose = T, rcov=~units)

FixedEffctVar<-vector(length = 1000)
for(i in 1:1000)
{
  FixedEffctVar[i]<- var(predict (MCMCGroup_DuraFeed_Mean_Question3_2_allfixedeff, it =
i))
}

Sol_DurFeed2_all<-
(FixedEffctVar/var(dataA[complete.cases(dataA$AvGDurFeed),"AvGDurFeed"])) *100
posterior.mode(as.mcmc(Sol_DurFeed2_all))
HPDinterval(as.mcmc(Sol_DurFeed2_all))

# Model removing a predictor (FID in this case)####
MCMCGroup_DuraFeed_Mean_Question3_2_MinusFID <- MCMCglmm(AvGDurFeed ~ 1+
                                                  # Stdzd_GMeanFID +
                                                  Stdzd_AduSexRatio +
                                                  Stdzd_GroupSize +
                                                  Stdzd_PropGAdu +
                                                  Stdzd_PropGAssoci +
                                                  Stdzd_Easting +
                                                  Stdzd_PropExposure+
                                                  Stdzd_TestBi,
                                                random = ~Date,
                                                data = dataA,
```

```
                                nitt=1300*times,thin=1*times,burnin=300*times,
                                family="gaussian", prior = prior2,
                                        verbose = T, rcov=~units)

FixedEffctVar<-vector(length = 1000)
for(i in 1:1000)
{
  FixedEffctVar[i]<- var(predict(MCMCGroup_DuraFeed_Mean_Question3_2_MinusFID, it =
i))
}

Sol_DurFeed2_MinusFID<-
(FixedEffctVar/var(dataA[complete.cases(dataA$AvGDurFeed),"AvGDurFeed"])) *100
posterior.mode(as.mcmc(Sol_DurFeed2_MinusFID))
HPDinterval(as.mcmc(Sol_DurFeed2_MinusFID))

# Estimate the difference in Marginal R2
posterior.mode(as.mcmc(Sol_DurFeed2_all)) -
posterior.mode(as.mcmc(Sol_DurFeed2_MinusFID))
```

Codes for estimating Partial R^2

```
# Fit a model in MCMC (standardised response and fixed effects)

library(MCMCglmm)
prior2 <- list(G = list(G1 = list(V = 1, nu = 0.002)),
               R = list(V = 1,nu=0.002)) # fix=1 if the trait is binary

times<-200

MCMCGroup_DuraFeed_GeneralStats_Mean_Question3 <-
                                              MCMCglmm(Stdzd_AvGDurFeed ~ 1+
                                            Stdzd_GMeanFID +
                                            Stdzd_AduSexRatio +
                                            Stdzd_GroupSize +
                                            Stdzd_PropGAdu +
                                            Stdzd_PropGAssoci +
                                            # Stdzd_Temperature +
                                            Stdzd_Easting +
                                            Stdzd_PropExposure+
                                            Stdzd_TestBi,
                                    random = ~Date,
                                    data = dataA,
                                    nitt=1300*times,thin=1*times,burnin=300*times,
family="gaussian", prior = prior2,
                                    verbose = T, rcov=~units)
save(MCMCGroup_DuraFeed_GeneralStats_Mean_Question3,file="MCMCGroup_DuraFeed_
GeneralStats_Mean_Question3_Final.rdataA")
summary(MCMCGroup_DuraFeed_GeneralStats_Mean_Question3)

# Fixed effect proportions (all fixed effects)
Sol_DurFeed<-MCMCGroup_DuraFeed_GeneralStats_Mean_Question3$Sol^2
summary(Sol_DurFeed)
HPDinterval(Sol_DurFeed)
posterior.mode(rowSums(Sol_DurFeed[,-1])+DateProportion+ResidualProportion)
posterior.mode(Sol_DurFeed)

# Personality proportion (group mean-FID)
posterior.mode (as.mcmc(Sol_DurFeed[,2])*100)
HPDinterval(as.mcmc(Sol_DurFeed[,2])*100)
# Group demographic structure proportion
posterior.mode(as.mcmc(rowSums(Sol_DurFeed[,3:6])*100))
HPDinterval(as.mcmc(rowSums(Sol_DurFeed[,3:6])*100))

# Location proportion (east)
posterior.mode (as.mcmc(Sol_DurFeed[,7])*100)
HPDinterval(as.mcmc(Sol_DurFeed[,7])*100)

# Experimental proportion (stimulus exposure + test)
posterior.mode (as.mcmc(rowSums(Sol_DurFeed[,8:9])*100))
HPDinterval(as.mcmc(rowSums(Sol_DurFeed[,8:9])*100))
```

```
# Temporal effect proportion (date)
DateProportion<-
MCMCGroup_DuraFeed_GeneralStats_Mean_Question3$VCV[,"Date"]/var(dataA[complete.c
ases(dataA$AvGDurFeed),"Stdzd_AvGDurFeed"])
posterior.mode(DateProportion*100)
HPDinterval(DateProportion*100)

# Residual proportion
ResidualProportion<-
MCMCGroup_DuraFeed_GeneralStats_Mean_Question3$VCV[,"units"]/var(dataA[complete.c
ases(dataA$AvGDurFeed),"Stdzd_AvGDurFeed"])
posterior.mode(ResidualProportion*100)
HPDinterval(ResidualProportion*100)
```

Table S1. Mean Feeding Duration model selection (Question 1)

(A)'Question 1' (Mean Feeding Duration): Mean FID					
DIC = 4239.321					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	48.991	47.291	50.953	901	<0.001
Mean-FID (m)	-5.044	-7.134	-3.209	888.4	<0.001
Adult Sex Ratio	-1.275	-2.970	0.610	1000	0.152
Group Size	0.826	-0.887	2.558	1359.4	0.378
Proportion of 2yo+	-0.473	-2.168	1.354	1143.7	0.598
Proportion of Associates	0.727	-1.440	3.031	1000	0.508
Easting	3.003	0.981	4.841	862.9	0.004
Stimulus Proximity	-0.892	-2.641	0.856	1000	0.340
Test (Object)	-5.506	-7.848	-3.208	1000	<0.001
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	14.470	0.000	43.830	147.3	
Residual	342.700	294.300	395.500	428.6	

(B) 'Question 1' (Mean Feeding Duration): Maximum FID					
DIC: 4245.852					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	49.087	47.230	50.755	1000	<0.001
Maximum-FID (m)	-4.223	-6.060	-2.367	1000	<0.001
Adult Sex Ratio	-1.181	-2.883	0.664	1000	0.200
Group Size	1.865	0.178	3.754	1000	0.038
Proportion of 2yo+	-0.512	-2.263	1.294	1000	0.558
Proportion of Associates	0.353	-1.573	2.578	1000	0.754
Easting	1.996	0.356	4.134	2034	0.038
Stimulus Proximity	-0.751	-2.507	1.032	1000	0.400
Test (Object)	-5.214	-7.167	-2.994	1060	<0.001
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	9.390	0.000	37.210	299.2	
Residual	349.700	302.600	399.700	829.4	

(C) 'Question 1' (Mean Feeding Duration): Minimum FID					
DIC: 4247.599					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	49.094	47.218	50.869	849.9	<0.001
Minimum-FID (m)	-4.109	-6.317	-2.310	667.8	<0.001

Adult Sex Ratio	-1.572	-3.196	0.448	1000	0.110
Group Size	0.367	-1.287	2.381	1000	0.704
Proportion of 2yo+	-0.678	-2.278	1.154	1000	0.468
Proportion of Associates	0.390	-1.642	2.564	799.6	0.706
Easting	2.723	0.638	4.827	1000	0.004
Stimulus Proximity	-0.469	-2.406	1.468	906.5	0.582
Test (Object)	-5.077	-7.420	-2.939	1000	<0.001
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	19.420	0.000	49.940	293.3	
Residual	344.800	294.300	398.200	708.3	

Results from the different models testing what group-summary statistics for personality best predicted 'Mean Feeding Duration' (Question 1). We mean-centred and standardised our predictors so that parameter estimates from different models were on the same scale. We chose the models with the smallest Deviance Information Criterion (DIC) values. Sample size: 486 observations. Base level: test (Sound). Model settings: Iterations = 60001:259801, Thinning interval = 200, and sample size = 1000. Definitions: post.mean (posterior mean), l-95% CI (lower-95% credible interval), u-95% CI (upper-95% CI), and eff.samp (effective sample size). ***(A)*** *shows results for a model with mean-FID,* ***(B)*** *with maximum-FID, and* ***(C)*** *with minimum-FID.*

Table S2. All Hopped Away model selection (Question 1)

(A) 'Question 1' (All Hopped Away): Mean FID					
DIC: 753.2209					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.314	-1.583	-1.038	1000	<0.001
Mean-FID	0.522	0.233	0.745	843.6	<0.001
Adult Sex Ratio	-0.049	-0.300	0.166	1000	0.688
Group Size	-0.206	-0.456	0.045	1000	0.110
Proportion of 2yo+	0.070	-0.169	0.293	1000	0.600
Proportion of Associates	0.219	-0.074	0.496	1000	0.118
Easting	-0.061	-0.302	0.203	1000	0.634
Stimulus Proximity	0.241	-0.082	0.541	1000	0.126
Test (Sound)	-0.246	-0.587	0.081	1000	0.156
Test (FID-non-target)	0.302	-0.058	0.662	1267.4	0.094
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	0.584	0.001	1.292	463.7	
Residual	Fixed to 1				

(B) 'Question 1' (All Hopped Away): Maximum FID					
DIC: 760.7052					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.295	-1.587	-1.013	831.9	<0.001
Maximum-FID	0.383	0.165	0.627	1127.4	<0.001
Adult Sex Ratio	-0.038	-0.278	0.186	1000	0.766
Group Size	-0.299	-0.562	-0.067	1000	0.018
Proportion of 2yo+	0.063	-0.164	0.277	1000	0.574
Proportion of Associates	0.256	0.007	0.529	1000	0.048
Easting	0.057	-0.189	0.279	1000	0.642
Stimulus Proximity	0.224	-0.039	0.526	1006.2	0.110
Test (Sound)	-0.209	-0.507	0.111	1000	0.192
Test (FID-non-target)	0.334	0.021	0.670	1000	0.046
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	0.528	0.000	1.276	433.2	
Residual	Fixed to 1				
(C) 'Question 1' (All Hopped Away): Minimum FID					
DIC: 754.6496					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.301	-1.592	-1.037	690.1	<0.001

Minimum-FID	0.525	0.273	0.794	1000	<0.001
Adult Sex Ratio	-0.025	-0.226	0.221	1000	0.856
Group Size	-0.125	-0.394	0.116	1000	0.352
Proportion of 2yo+	0.085	-0.140	0.312	1000	0.492
Proportion of Associates	0.233	-0.057	0.508	1000	0.126
Easting	-0.078	-0.333	0.179	1000	0.558
Stimulus Proximity	0.207	-0.090	0.481	943.5	0.172
Test (Sound)	-0.212	-0.538	0.074	1000	0.162
Test (FID-non-target)	0.255	-0.103	0.544	1000	0.148

Random effects (variance components)				
	post.mean	l-95% CI	u-95% CI	eff.samp
Date	0.555	0.000	1.297	407.7
Residual	Fixed to 1			

Results from the different models testing what group-summary statistics for personality best predicted 'All Hopped Away' (Question 1). We mean-centred and standardised our predictors so that parameter estimates from different models were on the same scale. We chose the models with the smallest Deviance Information Criterion (DIC) values. Sample size: 664 observations. Base level: test (Object). Model settings: Iterations = 60001:259801, Thinning interval = 200, and Sample size = 1000. Definitions: post.mean (posterior mean), l-95% CI (lower-95% credible interval), u-95% CI (upper-95% CI), and eff.samp (effective sample size). ***(A)*** *shows results for a model with mean-FID,* ***(B)*** *with maximum-FID, and* ***(C)*** *with minimum-FID.*

Table S3. Mean Feeding Duration – all individuals versus elders (Question 2)

(A) 'Question 2' (Mean Feeding Duration): All Members' Mean FID					
DIC: 3932.491					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	48.082	45.990	50.388	1000	<0.001
Mean-FID	-4.810	-6.788	-2.796	1000	<0.001
Adult Sex Ratio	-1.664	-3.535	0.066	1296	0.066
Group Size	0.710	-0.928	2.190	1000	0.354
Proportion of 2yo+	-0.253	-2.119	1.626	1000	0.764
Proportion of Associates	0.823	-1.133	2.833	1000	0.408
Easting	3.319	1.179	5.205	1000	0.002
Stimulus Proximity	-1.289	-4.227	1.453	1000	0.366
Test (Sound)	5.160	3.050	6.925	1000	<0.001
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	13.85	0.000	41.630	242	
Residual	344.500	296.500	394.400	637.9	
(B) 'Question 2' (Mean Feeding Duration): Elders' Mean FID					
DIC: 3939.355					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	47.936	45.597	50.192	1000	<0.001
Elders' Mean FID	-3.572	-5.306	-1.834	1000	<0.001
Adult Sex Ratio	-1.784	-3.465	0.155	1204.2	0.058
Group Size	1.074	-0.378	2.752	695.6	0.188
Proportion of 2yo+	-0.239	-2.222	1.556	1000	0.828
Proportion of Associates	0.442	-1.790	2.350	718.9	0.666
Easting	2.193	0.234	4.136	1096.5	0.028
Stimulus Proximity	-0.710	-3.727	1.884	1000	0.604
Test (Sound)	4.784	2.739	6.883	1000	<0.001
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	13.850	0.000	41.630	242	
Residual	13.850	0.000	41.630	242	

Results from the different models testing what group-summary statistics for personality best predicted 'Mean Feeding Duration' (Question 2). We mean-centred and standardised our predictors so that parameter estimates from different models were on the same scale. We chose the models with the smallest Deviance Information Criterion (DIC) values. Sample size: 421 observations. Base level: test (Object). Model settings: Iterations = 60001:259801, Thinning interval = 200, and Sample size = 1000. Definitions: post.mean (posterior mean), l-95% CI (lower-95% credible interval), u-95% CI (upper-95% CI), and eff.samp (effective sample size). ***(A)*** *shows results for a model with all individuals' mean-FID,* ***(B)*** *with elders' mean-FID. These models use a different subset on the model in Table S1, hence, (A) are slightly different between the tables.*

Table S4. All Hopped Away – all individuals versus elders (Question 2)

(A) 'Question 2' (All Hopped Away): All Members' Mean FID					
DIC: 696.9772					
Fixed effects	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.302	-1.575	-0.996	756.7	<0.001
Mean-FID	0.592	0.340	0.863	716.6	<0.001
Adult Sex Ratio	-0.072	-0.317	0.199	1000	0.604
Group Size	-0.185	-0.466	0.070	1000	0.190
Proportion of 2yo+	0.169	-0.061	0.416	1000	0.176
Proportion of Associates	0.155	-0.103	0.434	902.7	0.270
Easting	0.044	-0.195	0.278	1000	0.734
Stimulus Proximity	-0.086	-0.349	0.159	1000	0.532
Test (Sound)	0.258	-0.043	0.577	1000	0.112
Test (FID-non-target)	-0.276	-0.595	0.056	1000	0.098
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	0.503	0.000	1.328	280.6	
Residual	Fixed to 1				
(B) 'Question 2' (All Hopped Away): Elders' Mean FID					
DIC: 699.3102					
Fixed Effects					
	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
(Intercept)	-1.299	-1.610	-1.016	577.7	<0.001
Elders' Mean FID	0.510	0.266	0.759	1000	<0.001
Adult Sex Ratio	-0.070	-0.320	0.171	1000	0.618
Group Size	-0.206	-0.438	0.046	1110.8	0.108
Proportion of 2yo+	0.177	-0.097	0.397	1000	0.170
Proportion of Associates	0.200	-0.060	0.505	1000	0.170
Easting	0.022	-0.215	0.291	1000	0.878
Stimulus Proximity	0.019	-0.229	0.252	1000	0.886
Test (Sound)	0.240	-0.076	0.571	1000	0.136
Test (FID-non-target)	-0.233	-0.571	0.070	1000	0.156
Random effects (variance components)					
	post.mean	l-95% CI	u-95% CI	eff.samp	
Date	0.544	0.000	1.342	288.7	
Residual	Fixed to 1				

Results from the different models testing what group-summary statistics for personality best predicted 'All Hopped Away' (Question 2). We mean-centred and standardised our predictors so that parameter estimates from different models were on the same scale. We chose the models the smallest Deviance Information Criterion (DIC) values. Sample size: 611 observations. Base level: test (Object). Model settings: Iterations = 60001:259801, Thinning interval = 200, and Sample size = 1000. Definitions: post.mean (posterior mean), l-95% CI (lower-95% credible interval), u-95% CI (upper-95% CI), and eff.samp (effective sample size). ***(A)*** *shows results for a model*

with all individuals' mean-FID, ***(B)*** *with elders' mean-FID. These models use a different subset on the model in Table S2, hence, (A) are slightly different between the tables.*

Table S5. Among-group proportion estimates in the individual's intrinsic FID (Question 4)

(A)	**Models**		
	2-Individual Groups Only	3-Individual Groups Only	4-Individual Groups Only
Intercept	-0.145 (-0.233, -0.070)	-0.096 (-0.189, -0.005)	-0.139 (-0.243, -0.019)
Among-group variance	1.326 (1.130, 1.559)	0.765 (0.556, 1.030)	0.738 (0.493, 1.017)
Residual variance	4.643 (4.415, 4.859)	5.211 (4.929, 5.519)	5.360 (4.971, 5.716)
Among-group proportion	**22.62% (18.81%, 25.13%)**	**11.94% (9.33%, 16.96%)**	**11.12% (7.60%, 15.92%)**
(B)	All Groups (2-20 Individuals)	Smaller Groups (2-4 Individuals)	Larger Groups (5-20 Individuals)
Intercept	0.150 (-0.235, -0.063)	-0.164 (-0.374, 0.027)	0.084 (-0.158, 0.348)
Group Size	0.011 (-0.009, 0.033)	0.013 (-0.055, 0.091)	-0.016 (-0.050, 0.016)
Among-group variance	0.788 (0.681, 0.897)	1.017 (0.876, 1.149)	0.386 (0.258, 0.542)
Residual variance	5.261 (5.102, 5.393)	4.989 (4.822, 5.161)	5.713 (5.442, 5.970)
Among-group proportion	**13.14% (11.30%, 14.66%)**	**16.76% (14.87%, 18.99%)**	**6.65% (4.27%, 8.72%)**

Summary of among-group proportions in individual's intrinsic FID (in bold), over and above spatial structure or temporal effects. (A) shows the set of results of group-size specific models: 2-individual group data only (n: 5008 observations, 2085 Group IDs), 3-individual groups (n: 3249 observations, 1079 Group IDs), and 4-individual groups (n: 2152 observations, 539 Group IDs). (B) shows the results of combined group sizes: all groups (group size 2-20, n: 14712 observations, 4339 Group IDs), smaller groups (group size 2-4, n: 10409 observations, 3703 Group IDs), and larger groups (group size 5-20, n: 4203 observations, 636 Group IDs). No shading colour shows fixed effects, and grey colour shows random effect variance (with 'Among-group proportion' showing the proportion of total variance attributed to among-group variance, in %). Values are the posterior mean followed by 95% credible intervals for fixed effect and variance estimates. We show the posterior model and respective 95% CI for among-group proportion.

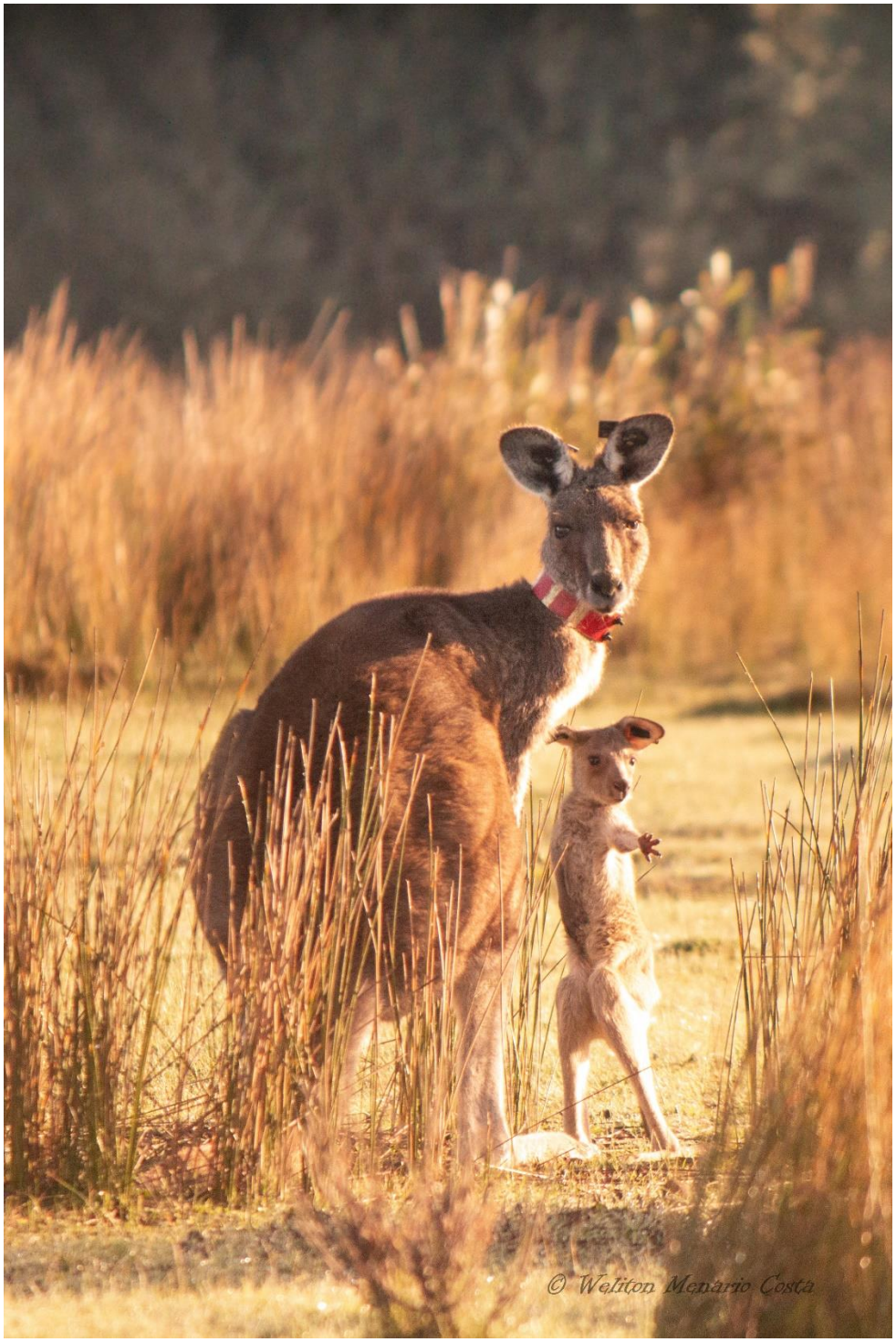
Table S6. Model selection using data from different experiments separately

Test	Trait	Model	Group Members Considered	FID Summary Statistics	Estimate of Standardised FID (95%CI)	Model DIC
Object test (275 groups)	Mean Feeding Duration-Object	OFeed_M1	all	**Mean**	**-2.176 (-3.378, -1.118)***	**2382.470**
		OFeed_M2		Maximum	-1.143 (-1.861, -0.328)*	2386.937
		OFeed_M3		Minimum	**-2.364 (-3.701, -1.202)***	**2382.601**
	All Hopped Away-Object	OHop_M1	all	**Mean**	**0.306 (0.127, 0.489)***	**332.964**
		OHop_M2		Maximum	0.025 (-0.019, 0.066)	581.015
		OHop_M3		Minimum	0.051 (0.026, 0.115)	580.2877
Sound test (189 groups)	Mean Feeding Duration-Sounds	SFeed_M1	all	**Mean**	**-2.279 (-3.573, -0.781)***	**1677.604**
		SFeed_M2		Maximum	-1.143 (-2.010, -0.245)*	1681.037
		SFeed_M3		Minimum	-1.923 (-3.372, -0.468)*	1681.633
	All Hopped Away-Sounds	SHop_M1	all	**Mean**	**0.248 (0.047, 0.466)***	**190.2946**
		SHop_M2		Maximum	0.016 (-0.025, 0.06)	396.4095
		SHop_M3		Minimum	0.030 (-0.044, 0.106)	396.3504
Approach test (501 groups)	All Non-Target Hopped Away-Approach	FHop_M1	all non-target	**Mean**	**0.390 (0.260, 0.509)***	**599.0167**
		FHop_M2		Maximum	0.045 (0.008, 0.080)*	1047.574
		FHop_M3		Minimum	0.069 (0.026, 0.110)*	1042.222

Results from the different models used to test what group-summary statistics best-influenced group differences in behaviour (considering data of each experimental test separately). We mean-centred and standardised our predictors so that parameter estimates from different models were on the same scale. We chose the models with the smallest model Deviance Information Criterion (DIC) values. Overall, these results were consistent with what is presented above (Tables S1 & S2, Question 1), when we combined the various data sets.

Chapter IV.

Maternal variance in juvenile behavior and survival in a wild marsupial



Authorship

Weliton Menário Costa[1], Wendy J. King[1,2], Marco Festa-Bianchet[1,2], Loeske E.B. Kruuk[1]

1 Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT, Australia

[2] Département de biologie, Université de Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke, Québec, J1K 2R1, Canada

To be submitted to Behavioral Ecology

Abstract

Individual behavior varies for many reasons, but how early in life are such differences apparent, and are they under selection? We investigated variation in behavior of pouch young, subadults, and their mothers, and associations with early survival in wild eastern grey kangaroos (*Macropus giganteus*). We measured behavior of young while still in the pouch and as subadults and estimated the covariation of these measures with maternal behavior. We found consistent variation between offspring of different mothers, or non-siblings, in early-life behaviour, including levels of pouch young movement and flight initiation distance of subadults, and in subadult survival, indicating similarity among siblings. There was no covariance between the measures of behavior at the two offspring stages, nor between the offspring behavioural traits with subadult survival. However, there was a strong covariance between offspring and mothers' behavior, in the form of flight initiation distance (FID). Of the total repeatability of subadult FID (55.3%), more than two-thirds could be attributed to differences between mothers. Our results indicate that (i) behavioral differences are apparent at a very early stage of development; (ii) repeatability in juvenile behavior is shaped by among-mother differences; and (iii) mothers and offspring exhibit similar behavioral responses. However, (iv) we found no evidence of selection on juvenile or maternal behavioral traits in this wild marsupial.

Introduction

Behavior varies both between and within individuals. Recent interest in repeatable behavioral differences between individuals has generated the field of animal 'personality', and with it substantial research on the causes of consistent differences between individuals (Bell et al. 2009; Dingemanse and Dochtermann 2013; Beekman and Jordan 2017). These differences may be family-specific: for any trait, behavioral or otherwise, offspring of different mothers are likely to differ in phenotype due both to differences in the genes they have inherited, and to maternal effects, which may be genetic or environmental (Mousseau and Fox 1998; Räsänen and Kruuk 2007). For behavioral traits, differences between offspring of different mothers may contribute to repeatable individual variation, and hence to measures of individual personality. The estimation of maternal-level variation involves the same variance-component approach used to estimate the repeatability of behavioral traits from among-individual variance in behavior (Dingemanse and Dochtermann 2013). Thus, if heritable genetic or maternal effects are important sources of individual differences (Taylor et al. 2012; White and Wilson 2019; Zablocki-Thomas et al. 2019), they will play an important role in determining behavioral repeatability. In mammals, maternal effects are especially important for early-life traits, but then typically decline as offspring age (e.g. (Wilson and Réale 2006; Gauzere et al. 2020), thus maternal effects on behavior play an especially important role very early in life. Understanding the causes of variation in behavior is particularly relevant when behavioral differences have fitness consequences (Smith and Blumstein 2008). If variation in behavior is associated with differences in survival or reproduction, it will be shaped by natural selection (Travis and Reznick 2018). Maternal effects on juvenile behavior could therefore have adaptive implications (Wolf and Wade 2001).

Differences between mothers in behavioral traits may also have implications for their offspring. Behavioral traits related to maternal care such as support during aggression or length of association with offspring have been shown to be correlated with offspring size and mass in

chimpanzees (*Pan troglodytes verus*) (Samuni et al. 2020). As examples, maternal behavior is also associated with offspring growth and/or fitness in spotted hyenas (*Crocuta crocuta*) (Höner et al. 2010), grey seals (*Halichoerus grypus*) (Hall et al. 2001), North American red squirrels (*Tamiasciurus hudsonicus*) (Westrick et al. 2020), chimpanzees (Murray et al. 2018; Schülke et al. 2019), and eastern grey kangaroos (King et al. 2017). Correlations between maternal and offspring behavioral traits have also been shown in several mammals: for example, there are mother-offspring similarities in personality traits in North American red squirrels (Taylor et al. 2012) and grey mouse lemurs (*Microcebus murinus*) (Zablocki-Thomas et al. 2019), and in dominance rank in rhesus macaques (*Macaca mulatta*), baboons (*Papio* sp.), vervet monkeys (*Cercopithecus aethiops sabaeus*) (see review: (Maestripieri 2018)), and spotted hyenas (East et al. 2009). Such correlations may be driven by both genetic and non-genetic effects, but there is increasing evidence of heritability of behavioral traits across a range of species (Weiss et al. 2000; Dingemanse et al. 2002; Williamson et al. 2003; Taylor et al. 2012; Dochtermann et al. 2015; Zablocki-Thomas et al. 2019). The similarity between parents and offspring in a trait, measured as the slope of regression of offspring on parent values, can provide an upper limit of the trait's heritability within a population, though the estimate may be inflated by common environment effects (Kruuk and Hadfield 2007; Thomson et al. 2018). Estimation of these parameters in wild populations requires identification and monitoring of parents and offspring and their behavior, which can be challenging in natural conditions (Archard and Braithwaite 2010). As a result, most studies of the heritability of behavior have used either captive-bred individuals or wild animals that were captured and then tested in artificial environments (Dochtermann et al. 2019) (but see (Saastamoinen et al. 2018), on dispersal).

In this study, we investigated variance and covariance in juvenile behavior and survival at two stages of early life in wild eastern grey kangaroos in Victoria, Australia, and how these traits (co)varied between offspring of different mothers and with maternal behavior. Studies of marsupials can offer insights on early juvenile development because of the possibility of

measuring phenotypes of very young offspring in the pouch. Our study population has been the subject of a long-term study (Montana et al. 2020), and previous work has shown that maternal body size and condition are positively correlated with offspring survival, fetal growth, and fecundity (Quesnel et al. 2017, 2018; Mackay et al. 2018). In other marsupials, a study in brushtail possums (*Trichosurus vulpecula*) showed that maternal mass is correlated with age at sexual maturity and with dispersal of sons (Bannister et al. 2019). We present here what is, to our knowledge, the first analysis of the behavior of pouch young, of maternal variation in behavior, and heritability of behavioral traits in marsupials. We also use rare observations of adoption (King et al. 2015a) to compare associations of mother-offspring behavior with biological versus adoptive mothers.

We observed two measures of offspring behavior and one of adult behavior and offspring survival. Offspring behavior was assessed by (i) levels of movement of pouch young in response to handling, at the very early stage of pouch young aged about 7-9 months, or 1-3 months before permanent pouch exit; and (ii) by flight initiation distance (FID) in response to an approaching human of subadults aged 13-35 months. The behavior of the mothers was also assessed with FID, and offspring survival was defined as survival to 21 months. We used a Bayesian Monte Carlo Markov Chain framework to analyze Gaussian and non-Gaussian traits using generalized linear mixed models (GLMMs). We first used a multivariate model of all four traits (three measures of behavior, one of survival) to quantify among-mother variance (or 'maternal repeatability', consistent differences between offspring of different mothers) in offspring behavior and survival, and also to quantify the covariances between the different offspring traits and maternal behavior. We then focused on FID, for which we had repeated measures of both offspring and mothers, and used these firstly to quantify the proportion of offspring repeatability (or 'personality') that was due to differences between mothers, and secondly to estimate the similarity between mother and offspring trait values, and hence obtain an indication of an upper limit of heritability. Finally, considering the 3% of cases of adoption in the study population (King et al. 2015a), we considered similarity of measures on FID of

adoptive young and both their biological and adoptive mothers, which could give more information into genetic versus non-genetic causes of similarity in behavior. We thus addressed four questions:

1. Do offspring of different mothers have different behaviors?
2. Are behavioral traits associated with subadult survival?
3. Do maternal effects contribute to the repeatability of offspring behavior?
4. Are maternal and offspring behaviour correlated, and do these associations differ for biological vs adoptive mothers?

Methods

Study population

We studied a wild population of marked eastern grey kangaroos at Wilsons Promontory National Park, VIC, Australia (38º57'S, 146º17'E). Offspring are predominantly born between November to May, and cohorts are named after the calendar year for January (e.g. an offspring born in December 2017 would be assigned to the 2018 cohort). This study involved individuals born in four cohorts (2016, 2017, 2018 and 2019) and their mothers. Observations took place in 2017, 2018, and 2019 (see below). Subadult Survival rates varied widely across years (Quesnel et al. 2018). Approximately 75 to 85% of kangaroos in the study area are individually marked. Captured individuals are ear-tagged with a unique color combination at first capture; this includes young in the pouch weighing >900 g (King et al. 2011). We considered as adults individuals aged three years and older. In the years of this study (2017, 2018 and 2019), there were 276-336 marked individuals in the population, including 115-140 adult females, about 55 adult males, 51-83 subadults, and 30-55 pouch young. We estimated the birth dates of pouch young based on body size measurements (Poole et al. 1982). Mothers were typically captured from August to November, when pouch young were aged about 6-10 months, when they would make only occasional sorties from the pouch (Poole 1975). Juveniles are usually weaned between the ages of 16 to 19 months (King & Goldizen 2016). Mothers were anaesthetized for capture, but pouch young were not (King et al. 2011). We left the young in the pouch while we measured the mother, and removed the young for only a few minutes to measure and tag it, and to weigh its mother. As described below, we scored the behavior of each pouch young during its short removal from the pouch.

Definitions of traits

We recorded data on four traits: (1) movement of pouch young handled at capture; (2) FID of subadults aged 13-35 months; (3) FID of adult females; and (4) subadult survival. We describe these below; summary statistics and samples sizes for each trait are in Table 1.

(1) Pouch Young Movement (PY Movement)

We assessed pouch young behavior by scoring body movements of pouch young across several stages of being handled at capture. We recorded 1/0 for any presence/absence of movement at eight stages and used the mean of these measures as an individual score between 0 and 1. The eight stages were: extraction from the pouch, placement inside a cloth bag, weighing, taking of three body measurements (foot, hind leg and head length), tagging, and being held inside an observer's jacket while the mother was weighed. For practicality in 2019, we combined responses for foot and hind leg length measurements and for extraction from the pouch and placement inside a cloth bag, to give a total of six stages of handling. The mean total duration of the process was 7.9 (±1.9 SD) min. PY Movement was assessed for the 2017, 2018, and 2019 cohorts, at 6.4 – 10.8 months of age. The average score of PY Movement was 0.46 (±0.50 SD, range 0 - 1). In the analyses of PY Movement, we included age at capture (mean 8.0 ± 0.8 SD) to account for potential age effects.

(2) and (3) Flight Initiation Distance (FID)

We measured FID on (2) subadults and (3) adult females, as the distance to the nearest metre at which a subadult (aged 1 or 2 years, range 13-35 months old) or an adult female moved away when approached (Strong et al. 2017). Mothers were tested on different dates from their young. Offspring were frequently closely associated with their mothers (King et al. 2015b). The observer started at 30 m from a group of kangaroos, walked directly towards a target individual at a constant speed, and stopped when the individual took flight. We tested one individual per group per approach, in group sizes of 1 to 7; for those in groups, the target individual was always the closest to the observer. FID measures were made in 2017-2019. Juveniles from the 2016 and

2017 cohorts thus had their responses measured at ages 1 and 2, whereas juveniles from the 2018 cohort were measured only at age 1, and the 2019 cohort was not tested for FID. We measured the FID of 168 adult females between 2017 and 2019, and included all observations in analyses; of these 168 females, 98 had offspring with PY Movement or Subadult FID measures.

For all measures of FID, we also noted the date, time, geographic position, and group size and composition, when applicable. Geographic positions were recorded with a Garmin® GPSMAP 64 device, in metres. To determine group composition, we followed a '10-m chain rule' (Jarman 1987; King et al. 2015a) whereby individuals were considered within the same group if they were within 10 m of at least one other group member. Group size was the number of individuals aged 1+ within a group. For each individual, we classified group composition by noting an individual's immediate associates, as a five-level factor: (1) no direct associate; (2) with its mother (for 1-year-old young; 2-year-olds were considered sufficiently independent as not to be considered as a direct associate); (3) with early-stage pouch young (aged 6 months or younger, not completely furred); (4) with large pouch young (completely furred young, aged 7 months or older); and (5) with 1-year-old offspring. There were rare cases of females with both a large pouch young and a yearling, so to reduce a parameter in the models, we considered those cases within class 4 (with a large pouch young). Finally, to account for any habituation via multiple testing of FIDs, we also recorded the test number and the number of days since the previous test (see Chapter II for more details on FID testing and comprehensive discussion of these variables), and included these in our models. High values of FID indicated a more distant response to the approaching human.

FID measures were available on more adult females than just the mothers of the offspring considered in this analysis. We analyzed all available data to provide the best estimates of variation in adult female FID. Adult FID measures were made by two observers, so the observer was fitted in the analysis (see below).

(4) Subadult Survival

We monitored Subadult Survival from 6 to 21 months, when most young were weaned (King et al. 2017). Complete survival data were available for cohorts 2016-2018, but we also included young that died before reaching ~15 months of age from the 2019 cohort, as the March 2020 field season was the last before these analyses were carried out. This introduces a negative bias on the 2019 year effect in Subadult Survival, but the additional data points added more precision to the estimation of maternal-level variance and did not change the qualitative results compared to a model excluding data from 2019.

Table 1. Summary statistics (means and standard deviations) and sample sizes for behavioural traits of eastern grey kangaroos.

Category	Trait	Number of Observations	Number of Individuals	Mean # of Observations per Individual	Number of Mothers	Mean # of Young per Mother	Trait Mean
Behavior	Pouch Young Movement (PY Movement) (score)	126	126	1	87	1.4 (±0.6)	0.45 (±0.50)
	Flight Initiation Distance – 1- and 2-year-olds (Subadult FID) (m)	487	106	4.6 (±2.2)	85	1.2 (±0.5)	6.60 (±4.4)
	Flight Initiation Distance – Adult Females (Adult Female FID) (m)	995	168	5.9 (±2.9)	-	-	6.70 (±3.80)
Survival	Subadult Survival (binary)	147	147	1	98	1.5 (±0.6)	0.55 (±0.50)

Pouch Young Movement is a score from 0-1 (see Methods); Flight Initiation Distance is in metres (m); '1- and 2-year-olds' refer to subadults of 13-35 months of age. Subadult Survival scores survival from 6 to 21 months (0 died, 1 survived).

Ethical Note

Captures were undertaken with ethics approval from the Université de Sherbrooke (no.s MFB2008-03, MFB2012-02, MFB2016-01), the University of Melbourne (no.s 0911512.1, 1312902.1) and the Australian National University (no. A2018/02) and research permits from the Victorian Department of Environment, Land, Water & Planning (no.s 10005558, 10007062, 10008630). Behavioral experiments and observations were conducted with animal ethics approval from the Australian National University (no.s A2017/17, A2018/02).

Statistical analysis

We fitted two multivariate generalized linear mixed models (GLMMs) using a Bayesian Monte Carlo Markov Chain framework in the statistical package *MCMCglmm* (Hadfield 2010). Code for the MCMCglmm models is provided in the Supplementary Material.

Model I: Four-trait model to estimate among-mother variances and covariances

Our first model was a four-trait GLMM with response variables of PY Movement, Subadult FID, Adult Female FID (which includes the juveniles' mothers), and Subadult Survival. Although we had repeated measures of PY Movement and Subadult FID on some offspring, to assist fitting of this four-trait complex model, we averaged multiple behavioural measures on each offspring into a single observation. Repeated measures of Subadult FID was considered in Model II below. Each offspring was therefore represented by a single observation for each behavioral trait as the average of the raw data measures. However, each mother could be represented by multiple offspring and also by multiple observations for her FID (average 5.9 observations per individual, range 1-14; Table 1). For all four traits, we therefore fitted Mother ID as a random effect to estimate among-mother variances (or maternal repeatability) and covariances for the offspring traits, and to estimate individual-level repeatability for Adult Female FID. For the offspring traits (PY Movement, Subadult FID, and Subadult Survival), the residual variance represented within-mother among-offspring variance.

For PY Movement, we fitted fixed effects of age at capture (in months), sex, and year (a 3-level factor: 2017, 2018 and 2019). For Subadult FID and Subadult Survival, we fitted fixed effects of sex and year (again, a 3-level factor). For Adult Female FID, we included GPS location along an east-west axis, group size, association type, test number, number of days since the previous test, year, and a two-level observer effect.

The MCMCglmm runs generated posterior distributions of parameter estimates, from which we could estimate posterior means and 95% higher posterior density credible intervals (CIs) for each parameter and also any derived estimates. For each offspring trait, we estimated *maternal repeatability* by dividing the estimate of the variance between mothers by the total variance in the offspring traits (defined as the sum of all the variance components). For Adult Female FID, we estimated *individual repeatability* by dividing among-individual variance in behavior (the Mother ID variance) by the total variance (Bell et al. 2009). Calculations were done on each sample of the posterior distribution to generate posterior distributions of the two

repeatability estimates, from which we estimated posterior means and 95% CIs. As Subadult Survival was a binary trait (0/1), we fitted it specifying the family 'categorical' in MCMCglmm. In a binomial model, residual effects have a variance fixed at 1, but can covary with other random effects. Because parameter estimates from the Subadult Survival model were on the latent (logit) scale, we used the QGicc function in the package QGglmm (Villemereuil et al. 2016; Villemereuil 2018) to back-transform the latent-scale variance estimates and also to calculate repeatability on the original data scale as well as on the latent scale (see equations below). In doing so, we note that the latent-scale residual variance is effectively 'overdispersion' variance on the data scale, with the data-scale residual variance generated by the stochasticity of the binomial sampling (Villemereuil et al. 2016).

Latent scale:

$Maternal\ repeatability = \frac{var(Mother)}{(var(Mother)+var(residual)}$, where var(residual)=1

Data scale:

$$Maternal\ repeatability = \frac{var(Mother)}{(var(Mother) + var(residual) + var(binomial\ sample)}$$

We used Model I to estimate covariances between the four traits. In particular, we estimated maternal-level covariance between the Mother ID effects, for all behavioral responses, to test for maternal-level associations across offspring traits (e.g. do mothers whose offspring have high PY Movement also have high FID?), and for associations between a mother's FID and her offspring's average behavior and Subadult Survival (e.g. do mothers with a high FID have offspring with high PY Movement?). We also fitted residual covariance for offspring traits to test for within-mother offspring-level associations (e.g. for an individual offspring, is relatively high PY Movement associated with relatively high FID?). For the random effect Mother ID, a covariance matrix between traits was set using the *us* variance structure in MCMCglmm. We also used the *us* structure for the residual variance-covariance of the offspring traits, but we could not estimate residual covariance with the Adult Female FID because this was the only trait measured

on adults, and so used the *idh* structure for this term. The model was run for 1300 x 10^3 iterations, with a burn-in of 300 x 10^3, a thinning interval of 1 x 10^3 and an inverse-Wishart prior distribution. We report the posterior distribution mode and 95% CIs for each parameter, considering there to be statistical support for the covariance of random effects if the 95% CIs did not overlap 0 and, for fixed effects, if pMCMC (Particle Markov Chain Monte Carlo) was <0.05.

Model II: Bivariate model of repeated measures of subadult and adult FID

Our second model was a bivariate model for Subadult FID and Adult Female FID to estimate the contribution of among-mother variation to the repeatability of juvenile behavior. For this model, we used multiple measures on each offspring for Subadult FID (average 4.6 observations per individual, range 1-10; Table 1). For the response variable Subadult FID, we fitted random effects of Mother ID and Offspring ID, and for the response variable Adult Female FID, we fitted a random effect of Mother ID. For both traits, we fitted the fixed effects of GPS location along the east-west axis, group size, associate type, test number, the number of days from the previous test, and year. Further, for Subadult FID, we fitted fixed effects of age at observation (as a covariate, measured in years) and sex, and for Adult Female FID, a two-level effect of observer.

As above, we fitted the covariance between the Mother ID effects across the behavioral responses using a *us* variance structure. We used the *idh* structure for both the Offspring ID random effect and the residual effects, as there could be no covariance across the two traits in within-individual effects. The model was run for 1300 x 10^3 iterations, with a burn-in of 300 x 10^3, a thinning interval of 1 x 10^3 and an inverse-Wishart prior distribution. For Subadult FID, we estimated *maternal repeatability* by dividing the variance between mothers (Mother ID variance) by the total variance, and *individual repeatability* by dividing the variance between offspring (Offspring ID variance) by the total variance as above. For Adult Female FID, we estimated *individual repeatability* by dividing the variance between individuals (Mother ID variance) by the

total variance. We only had Gaussian-family traits, so unlike Model I, we did not need to converge repeatability to the data scale in Model II.

In the Subadult FID response, we then estimated *total repeatable* variance between individuals from the variance between mothers plus the additional variance between offspring, and hence a *total repeatability* by dividing this sum by the total variance:

$$Total\ Repeatability = \frac{var(Mother)+var(Offspring)}{var(Mother)+var(Offspring)+var(residual)}.$$

Finally, we estimated the proportion of total repeatability (i.e., offspring plus maternal variance, as in a model without Mother ID the among-mother variance would be attributed to among-offspring variance) that could be ascribed to maternal effects as:

$$Maternal\ Proportion\ of\ Total\ Repeatability = \frac{var(Mother)}{var(Mother)+var(Offspring)}$$

Mother-offspring regression for FID

We estimated the slope of the mother-offspring regression of FID in this wild marsupial population as an indication of the trait heritability, as we did not have a sufficiently informative multi-generational pedigree for a more complex 'animal model' (Kruuk, 2004). Instead, we used the phenotypic covariance between mothers' and offspring's FID estimated here to provide an estimate of the mother-offspring regression slope, and hence (having doubled the estimate) an upper limit on the heritability $upper\_h^2$ (Falconer and Mackay 1996), noting that this may be inflated by maternal or shared environmental effects (Kruuk and Hadfield 2007):

$$upper\_h^2 = 2\ x\ \frac{cov(Mother,Offspring)}{var(Mother)}$$

Models I and II provided estimates of the covariance between mothers and offspring in FID and the variance between mothers, hence an estimate of the regression slope. We opted to use the estimates from Model II to calculate $upper\_h^2$ as it was a more detailed FID model, with repeated measures for all traits. We estimated the heritability, as above, estimated from calculations for each sample of the posterior distribution.

Finally, to examine 6 cases of adoption (King et al. 2015a), we repeated the analysis of the effect of Adult Female FID on Subadult FID using observations on adopted individuals for which we had measured FID for the young and both the biological and adoptive mothers. Despite the very small sample size, this analysis could indicate the relative importance of genetic and non-genetic causes of similarity in behavior. All adoptions occurred before permanent pouch emergence, and the adoptive mother was responsible for all subsequent maternal care until weaning, about 6 – 9 months later (King et al 2015a).

Results

We present below the results from the two models, addressing our four questions. We used Model I for questions 1 and 2 and Model II for questions 3 and 4. The main results are summarised in Table 2 and Figure 1 (Model I, the four-trait model) and in Table 3 and Figure 2 (Model II, focusing on FID); further details are available in Tables S1-S2.

1. Do offspring of different mothers have different behavior?

We found consistent differences between non-siblings in behavior, suggesting maternal repeatability of both PY Movement and Subadult FID (Model I, Table 2b; Table S1; and Figure 1). PY Movement and Subadult FID had maternal-level repeatability of 55.4% and 59.5%, respectively (Table 2c), with credible intervals clearly separated from zero, indicating that the majority of variance in these traits could be ascribed to differences between offspring of different mothers. These results suggest similarities among offspring of the same mother, as variance amongst mothers implies covariances within mothers. For Subadult Survival, maternal repeatability also accounted for the majority of variation (62.9%) on the latent scale (Table 2b). When we back-transformed the GLMM latent-scale estimates for Subadult Survival to proportions on the data-scale, adding the stochastic variation associated with moving from the expected to observed data scale (Nakagawa and Schielzeth 2010; Villemereuil et al. 2016), maternal repeatability continued to be higher than the proportion of residual variance, but the overall maternal repeatability estimates decreased to 4.9%, with credible intervals separated from zero (Table 2c). This decrease also occurred with the latent-scale residual variance (or over-dispersion), which corresponded to 3% on the data-scale. The back-transformation thus illustrates the dramatic change in estimates of proportions depending on whether estimates are considered on latent vs. data scales.

2. Are behavioral traits associated with survival?

Considering the covariances estimated in Model I, there was no support for any associations between offspring or maternal behavioral traits and Subadult Survival, either across mothers (Table 2b, 'Mother ID') or within mothers (Table 2b, 'Residual').

From the estimates of the other covariances in Table 2b, there was also no support for any association of PY Movement and Subadult FID (Table 2b). The covariance between Maternal and Subadult FID was large, however, and we consider this association in more detail with Model II below.

3. Do maternal effects contribute to the repeatability of offspring behavior?

Model II focused on the repeated measures of FID in both offspring and mothers (Table 3). Maternal repeatability was 39.8% for Subadult FID in this model (Table 3b), slightly lower than the estimated in Model I. A further 15.5% of the variation was due to among-individual (within-mother) differences (Table 3b), so overall the 'total repeatability' of Subadult FID was 55.3%. Differences between mothers, therefore, accounted for 75.8% (95% CI: 45.0 - 92.0%;), of the total repeatability of Subadult FID. The individual repeatability of Adult Female FID was 35.5% (Table 3b).

4. Are maternal and offspring behaviour correlated, and do these associations differ for biological vs adoptive mothers?

We found strong associations between measures of FID of mothers and offspring (Figure 2; note that FIDs were estimated in different trials at different times). Models I and II both estimated positive covariances between maternal and offspring FID, though the estimate was slightly higher in Model I (6.46 (95% CI: 3.88, 9.44), Table 2b) versus 5.35 (95% CI: 3.02, 7.55), Table 3a). We used only the Model II estimates to calculate the upper-limit heritability (twice the slope of the parent-offspring regression) as it is a more detailed FID model. The

upper-limit estimate of heritability was 1.97 (95% CI: 1.32, 2.64). The estimate was larger than 1, which should not occur for narrow-sense genetic heritability (Falconer and Mackay 1996); as we discuss below, this indicates that shared common environmental effects inflated the covariances between mothers and offspring.

The supplementary regressions on the FIDs of adopted young (Table S3) returned equivalent heritability estimates (twice the slope) of 0.75 (95% CI: -6.22, 21.02) for the biological mother and 1.26 (95% CI: -16.07, 22.41) for the adoptive mother. Although the adoptive mother estimate was nearly twice that of the biological mother, the values had very large uncertainty reflecting the small sample sizes (n=6), and it was not possible to determine the relative impact of direct genetic effects and maternal environmental effects.

	PY Movement		Subadult FID		Subadult Survival		Adult Female FID	
(a) Fixed effects	Estimate	p-MCMC	Estimate	p-MCMC	Estimate	p-MCMC	Estimate	p-MCMC
Intercept	0.30	0.508	8.2	<0.001	0.87	0.040	1.5	0.660
	(-0.43, 1.12)		(6.7, 9.6)		(0.08, 1.82)		(-5.6, 9.2)	
Sex (Female)	0.09	0.188	0.3	0.666	-0.58	0.240	-	-
	(-0.04, 0.23)		(-1.1, 1.5)		(-1.55, 0.48)			
Year								
-2017	-0.08	0.212	-2.4	0.002	-0.50	0.388	-0.7	0.278
	(-0.21, 0.05)		(-3.6, -0.8)		(-1.68, 0.53)		(-2.0, 0.6)	
-2018	-0.22	0.010	-1.1	0.176	0.58	0.426	-0.6	0.180
	(-0.37, -0.07)		(-2.7, 0.5)		(-0.62, 1.95)		(-1.6, 0.3)	
-2019	(base)		(base)		-34.57	<0.001	(base)	
					(-64.83, -3.08)			
	Variance-covariance matrices							
(b) Random effects	PY Movement		Subadult FID		Subadult Survival		Adult Female FID	
Mother ID								
PY Movement	0.102 (0.068, 0.140)*							
Subadult FID	-0.038 (-0.377, 0.289)		9.930 (4.126, 15.856)*					
Subadult Survival	-0.011 (-0.141, 0.153)		2.078 (-0.419, 5.083)		2.204 (0.305, 5.307)*			
Adult Female FID	-0.024 (-0.258, 0.232)		6.456 (3.882, 9.444)*		1.460 (-0.271, 3.451)		5.601 (3.842, 7.311)*	
Residual								
PY Movement	0.077 (0.056, 0.103)							
Subadult FID	0.010 (-0.215, 0.236)		6.579 (3.874, 9.899)					
Subadult Survival	-0.002 (-0.082, 0.082)		-0.004 (-2.091, 2.153)		Fixed to 1			
Adult Female FID	-		-		-		9.125 (8.303, 9.968)	

(c) Variance proportions

	PY Movement	Subadult FID	Subadult Survival (latent scale)	Subadult Survival (back-transformed)	Adult Female FID
Maternal repeatability	55.4% (45.7%, 66.8%)*	59.5% (37.0%, 80.0%)*	62.9% (36.3%, 87.6%)*	4.9% (1.3%, 12.3%)	38.6% (29.2%, 45.1%)*
Residual (latent scale = overdispersion on data scale) Overdispersion			37.1% (12.4%, 63.7%)	3.0% (1.8%, 4.1%)	
Residual (data scale)	44.6% (33.2%, 54.3%)	40.5% (20.2%, 63.2%)		92.1% (87.7%, 98.7%)	61.4% (54.9%, 70.8%)

Table 2. Summary of the four-trait Model I, analyzing repeatability in the behavior of eastern grey kangaroos, with response variables: Pouch Young (PY) Movement, Subadult Flight Initiation Distance (FID, averaged), Subadult Survival, and Adult Female FID.

Here, I show only the fixed effects of sex and year as they were more important for our questions in this chapter, but see Table S1 for estimates of all variables included in the model. ***(a)*** *fixed effects of sex and year for each response. The 'base' level for sex is male, and for year is 2019 for the behavioral traits and 2016 for Subadult Survival. As discussed in methods above, there were biases on the 2019 year effect for Subadult Survival; hence, that particular effect is irrelevant.* ***(b)*** *random effects: (co)variance estimates from posterior distribution mode and respective 95% CI in parentheses: variances are on the diagonal, and covariances below diagonal.* ***(c)*** *Estimates of individual and maternal-level repeatability followed by the 95% CI in parentheses. '-' indicates that the variable was not fitted in the model. '*' indicates that pMCMC<0.05 for fixed effects or that the credible intervals did not overlap zero for random effects and repeatability estimates (also in bold). The among-mothers variance proportion estimated for Adult Female FID is an among-individual variance and considers all adult females, not only 'mothers'.*

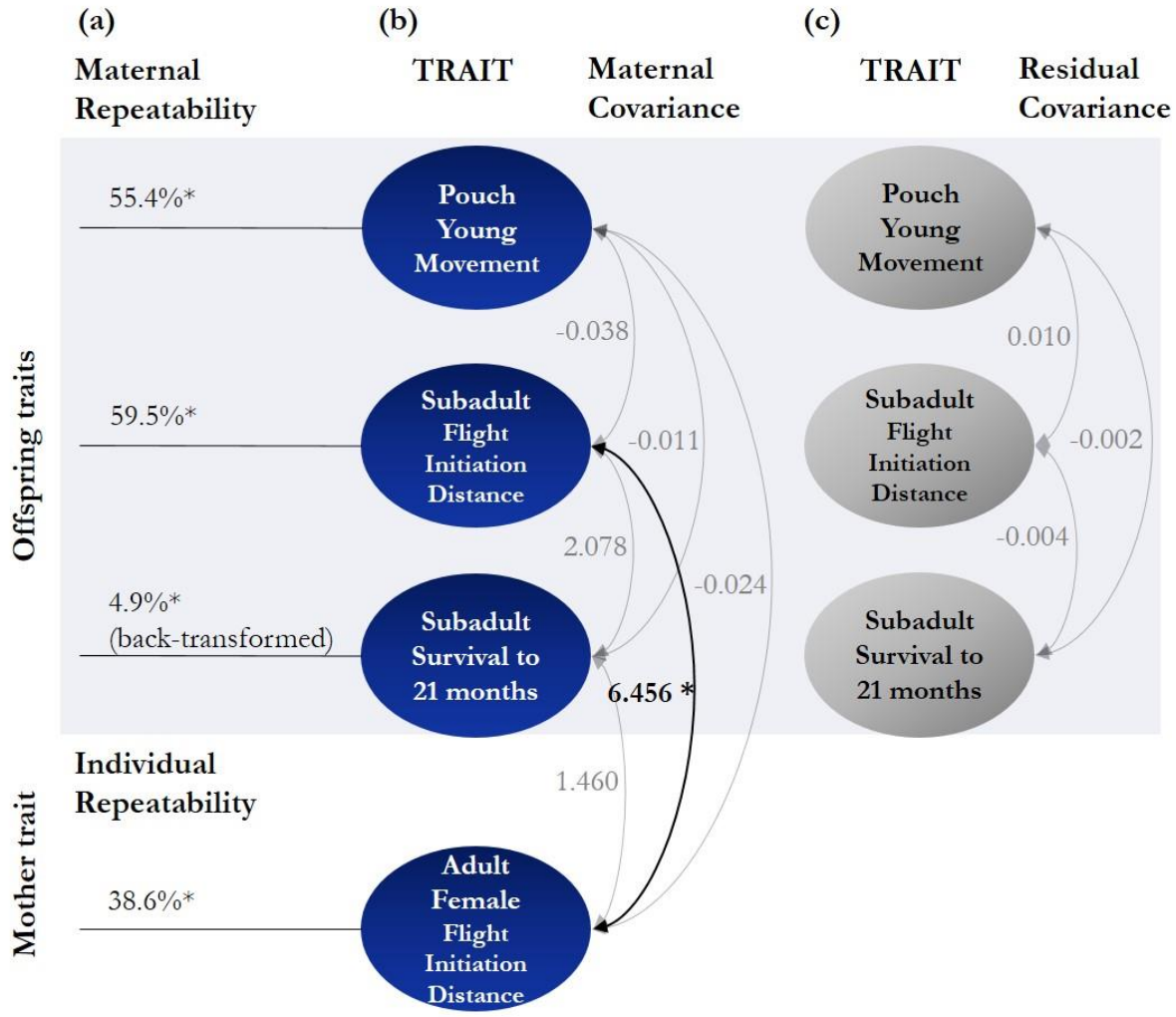


Figure 1. Mother and offspring (co)variances estimated from Model I (Table 2) for eastern grey kangaroos.

(a) *Maternal repeatability estimates for offspring traits, and individual repeatability for Adult Female FID (Table 2c).* ***(b)*** *maternal-level covariances across offspring and mother traits (Table 2b).* ***(c)*** *residual (offspring-within-mother) covariances across offspring traits (Table 2b). '*' indicates statistical support (also, in bold). Arrows link pairs of traits for which covariance estimates are shown.*

Table 3. Summary of the random effect results from the bivariate Model II, using repeated measures of flight initiation distance (FID) on each juvenile and adult female individual eastern grey kangaroo.

	Subadult FID	Adult Female FID
(a) Random effects: variance-covariance		
Offspring ID	2.726 (0.682, 4.884)	-
Mother ID		
Trait: Subadult FID	6.598 (2.748, 10.630)	
Trait: Adult Female FID	5.350 (3.024, 7.549)	5.405 (3.767, 7.167)
Residual	8.326 (7.168, 9.605)	9.137 (8.338, 10.101)
(b) Variance proportions		
Maternal repeatability	39.8% (21.3%, 54.0%)	-
Individual repeatability	15.5% (4.1%, 27.2%)	35.5% (28.5%, 45.0%)
Residual	46.0% (36.6%, 60.0%)	64.5% (55.0%, 71.5%)
Total repeatability	55.3% (40.2%, 62.8%)	35.5% (28.5%, 45.0%)
Maternal proportion of total repeatability	75.8% (45.0 – 92.0%)	-

This table focuses on the random structure, but fixed effect results are presented in Table S2. **(a)** random effect variance-covariance estimates from Model II. **(b)** proportions of total variance contributed by maternal, individual, and residual variances. Total repeatability is defined as the sum of among-mother (maternal repeatability) and among-individual (individual repeatability) proportions. Estimates are followed by the 95% credible intervals (CI) in parenthesis.

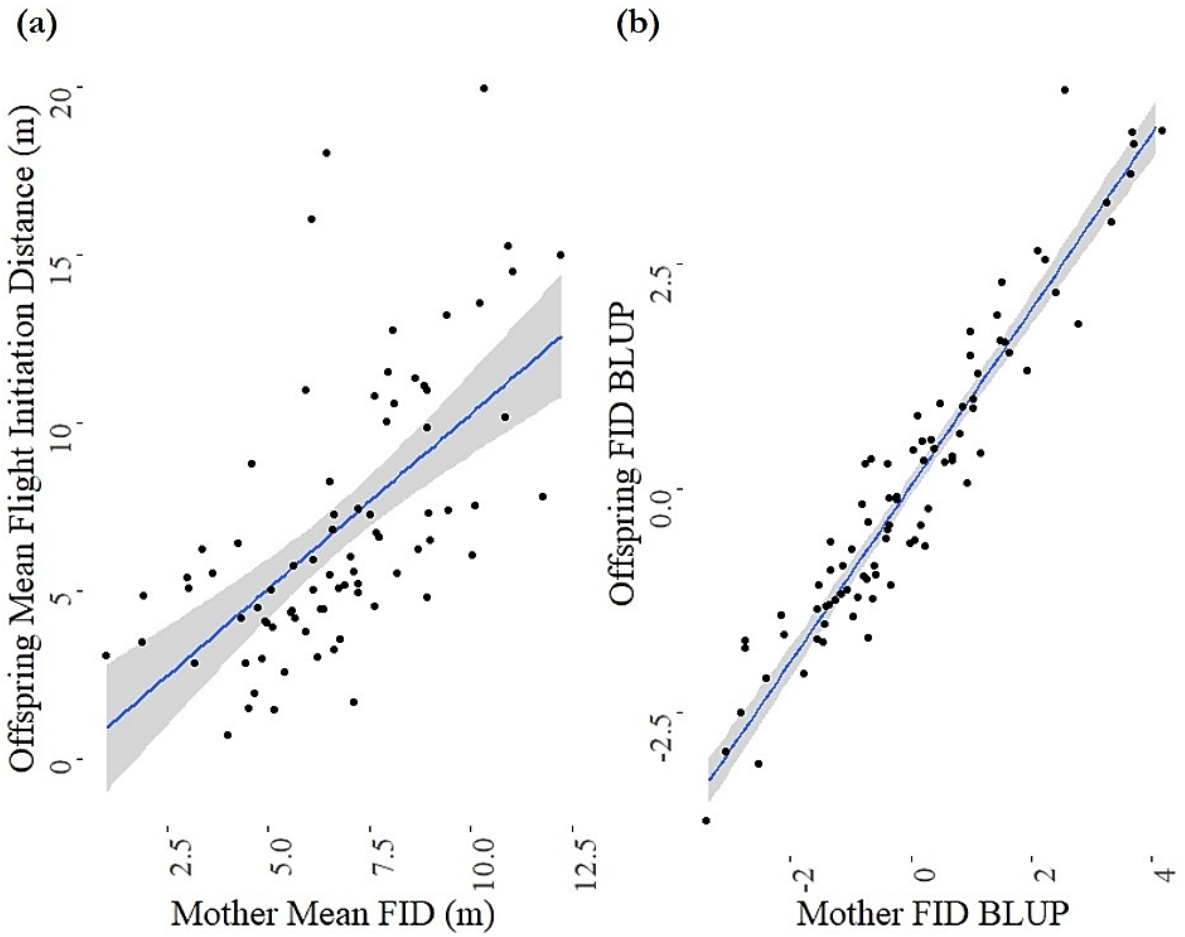


Figure 2. Relationship between mother and subadult offspring mean flight initiation distances (FIDs; n=85 mothers) in eastern grey kangaroos.

(a) *Regression of raw data points and* ***(b)*** *associations of the best linear unbiased predictor (BLUP) of the random effects from Model II, after accounting for other variables in the model (see Table S2). The blue lines indicate the linear regression, and the grey shades the 95% confidence intervals. The heritability values derived from the analyses were greater than 1, suggesting that maternal environmental effects inflated the covariance (see main text).*

Discussion

Our analyses indicate a substantial variation between offspring of different mothers in early-life behavior and survival. There was no indication of covariance between juvenile or maternal behavior and subadult survival. Mothers and offspring, however, had similar FID. Below, we discuss the implications of these results for our understanding of individual behavioral differences.

Maternal effects can affect a range of offspring phenotypes. Strictly speaking, the term 'maternal effects' refers to the impact a mother has on her offspring over and above the direct effects of the genes they inherit from her. Maternal effects can be driven by both environmental or genetic differences between mothers (Rossiter 1996; Reinhold 2002; Räsänen and Kruuk 2007). Maternal effects during early development may be especially important in mammals, as offspring are dependent on maternal provisioning through gestation and lactation (Reinhold 2002), and mothers may thus be expected to have the most influence on offspring phenotype at these early stages before other external influences increase in relevance (East et al. 2009; Maestripieri 2018; Samuni et al. 2020). These effects also include influences on early-life offspring survival (Nafus et al. 2015; Quesnel et al. 2017; Westrick et al. 2020). Maternal effects can be quantified from the differences between offspring of different mothers (Kruuk and Hadfield 2007), and maternal effect variance in mammals generally declines with offspring age (Wilson and Réale 2006; Gauzere et al. 2020). Marsupials offer an especially interesting opportunity to investigate phenotypic variation at very early developmental stages, but to our knowledge no study has considered the amount of variance due to maternal effects in pouch young. Some marsupial studies have addressed maternal traits' influence on offspring survival in the brushtail possum (Bannister et al. 2019) and eastern grey kangaroos (Quesnel et al. 2017, 2018), but none explicitly quantified maternal-level variance. Our results here show support for maternal repeatability in both behavior and survival of offspring. Variance between mothers in these offspring traits suggests genetically and/or environmentally induced effects on juvenile

offspring phenotype in a wild social marsupial. We note that since we did not estimate a genetic component of differences between individuals directly, these effects were not 'over and above the direct effects of genes inherited from the mother', in the strictest sense of maternal effects variance. Evidence of maternal repeatability in pouch young movement is particularly interesting because it reveals differences between offspring from different mothers at a very early life stage, when the offspring is still in the pouch – at a comparable stage to a fetus in late gestation in eutherian mammals (Tyndale-Biscoe and Renfree 1987).

In the other two traits (PY Movement and Subadult Survival), we found support for maternal repeatability, but no evidence of covariance with maternal behavioral trait. From a life-history evolution perspective, subadult survival can be considered a component of maternal total fitness (Wolf and Wade 2001), and the absence of covariance between adult female FID and offspring survival suggests that FID is not under selection via offspring survival in this population. This conclusion is in contrast to studies that found an association between maternal behavior and offspring behavior or survival: for example, in North American red squirrels, maternal aggressiveness was correlated with offspring survival (Boon et al. 2007), and there was also a maternal effect correlation between offspring activity and aggression (Taylor et al. 2012). Maternal dominance was correlated with offspring survival to 24 months in spotted hyenas (Hofer & East, 2003; Höner et al, 2010) and with stress hormone levels in baboons subadult males (Onyango et al. 2008). Maternal sociality was linked with offspring mortality via infanticide in white-faced capuchin monkeys (*Cebus capucinus imitator*) (Kalbitzer et al. 2017).

Among-mother variance may contribute to among-individual offspring behavioral variation, but few studies have tried to isolate such effects during repeatability estimation (Taylor et al. 2012; White and Wilson 2019; Zablocki-Thomas et al. 2019). This separation is important because genetically and/or environmentally induced effects can contribute to differences among individuals. Some researchers recommend not to 'correct' for intrinsic or extrinsic sources of individual heterogeneity when analyzing personality, as these may be relevant to the individual

differences observed in the population (Wilson 2018). However, it is still valid to ask how much of the total repeatability is actually due to maternal influences, in the same way as analyses may consider how much of repeatability is due to sex, age, or environmental effects (Bell et al. 2009). Our results showed that 70% of the total repeatability in offspring FID could be explained by among-mother differences, which suggests that less than a third of the total repeatability was due to individual differences in behavior between offspring of the same mother.

The field of animal personality has expanded rapidly in recent years (Beekman and Jordan 2017), but the investigation of how maternal personality contributes to offspring trait variation has received less attention (Reddon 2012; Schuett et al. 2013). We estimated covariances between the repeatable behavioral trait of FID in mothers (arguably a measure of their personality), and offspring behavioral traits. We found that maternal FID covaried with offspring FID (Figure 2). Of the increasing number of investigations in heritability of personality, most are restricted to primates and ungulates in mammals (Dochtermann et al. 2019). Our analyses estimate an upper limit on the heritability of FID, based on the regression of offspring values on maternal values. We only had information on the resemblance between mothers and offspring, though ideally, we would have fitted a model estimating both additive genetic variance and maternal effect variance separately, to separate the shared-genes effects from maternal environmental effects (Kruuk and Hadfield 2007). Some studies of personality have used this approach and found support for heritability of behavioral traits (Taylor et al. 2012; Dochtermann et al. 2015; Zablocki-Thomas et al. 2019). Given the definition of heritability as the proportion of total phenotypic variance due to additive genetic variance (Falconer and Mackay 1996), the value of heritability should always lie between 0 and 1. Our results show an unrealistically large upper-limit heritability estimate (>1), which was presumably inflated by environmental effects increasing similarities between mothers and offspring. A study on zebra finches (*Taeniopygia guttata*) cross-fostered offspring and found that personality variation arose primarily as a maternal effect relative to the foster mother, indicating that shared environmental

effects may be more important than genes in shaping personality variation (Schuett et al. 2013). For the few cases of adoption in our population, the covariance of offspring with the foster mother was nearly twice as large as that with the biological mother, indicating a similar trend to that found in zebra finches, but high uncertainty made meaningful conclusions difficult. If we ignore the uncertainty due to small sample sizes and consider the biological mother association as the better estimate of heritability, the upper-limit heritability is around 70% for FID; however, this estimate may still be confounded by environmental and maternal effects prior to the adoption. Future work on genomic data for the study population that will allow us to separate more elegantly the different environmental and genetic factors contributing to phenotypic variance in eastern grey kangaroos.

Several aspects of the environment could cause mothers and offspring to have similar personalities. Mother and offspring share physical and social environments, particularly when the juvenile is 10-19 months old, as they are closely associated before weaning (King and Goldizen 2016)). When we corrected for one spatial structure effect (Table S2), the correlation between maternal and offspring behavior increased (Figure 2). An alternative explanation for why mothers and offspring had similar personalities could be that offspring learn from the mother a responsiveness level to threat during the first days after pouch exit, which is likely a critical life stage (Breed and Sanchez 2010). We tested for mother and offspring FID responses at separate times, but FID response was independent of whether or not the mother and offspring were in the same group (Table S2). Therefore, offspring do not appear to copy their mother's behavior (or vice versa) at the time of the test but may have learned it early in life. Another indication of learning is that juveniles did not change their FID significantly from age 1 to 2 (Table S2), even though they were more independent of their mother at age 2. A laboratory study on a lizard, *Liopholis whitii,* (Munch et al. 2018) housed offspring either with their mother or alone during the first eight weeks of life and found that maternal presence affected juvenile expression of three key behaviors (latency in time until offspring emerged after a threat, length of time spent moving

in the familiar environment, and length of time spent moving in a novel environment) compared to individuals housed alone. The lizard study did not quantify the similarities in behavior between mother and offspring in either treatments, but provided evidence that even relatively simple forms of mother-offspring association early in life can significantly impact offspring behavioral phenotypes.

We used a Bayesian MCMC framework to fit multivariate GLMMs to behavioral traits (Hadfield 2010). This procedure allowed us to estimate variance components and their associated proportions of variance and covariance, along with estimates of uncertainty on all parameters. It also allowed us to estimate covariances between different traits and to partition these covariances into contributions from different sources (maternal-level vs. individual-level). Our models also illustrate the use of a generalized linear model to estimate components of variance in the binary trait of survival, and the necessary back transformation from the latent to data scale – with a corresponding reduction in variance explained. The Bayesian MCMC framework also allowed us to derive additional statistics such as the proportion of total repeatability attributable to mothers, Table 3, heritability estimates and their associated uncertainty. Translation of estimates from the GLMM of Subadult Survival from latent-scale values to data-scale values illustrates some of the complexities associated with GLMMs (Nakagawa and Schielzeth 2010; Villemereuil et al. 2016). We provide code for the statistical models in the Supplementary Information to encourage this approach in other analyses of behavioral trait variation.

In conclusion, our study revealed that differences between offspring of different mothers (equivalent to similarities among offspring of the same mother) are an important source of variation in offspring behavioral traits and early survival in a wild marsupial. Despite the support for maternal repeatability, Adult Female FID only covaried with Subadult FID. More than two thirds of offspring FID total repeatability was explained by maternal-level repeatability, which means that only 30% of total repeatability was due to consistent individual differences in offspring behavior over and above differences between offspring of different mothers. We

found a high resemblance between maternal and offspring FIDs, but our estimates of this covariance indicated that it was heavily inflated by shared environmental effects. We hope that more studies using multivariate approaches will provide valuable contributions to our understanding of the causes and consequences of variance and covariance between individuals in their behavior at different stages of life.

References

Archard GA, Braithwaite VA (2010) The importance of wild populations in studies of animal temperament. J Zool 281:149–160. https://doi.org/10.1111/j.1469-7998.2010.00714.x

Bannister HL, Hodgens P, Moseby KE (2019) Offspring sex and maternal effects influence the development and natal dispersal of an arboreal marsupial. J Mammal 100:423–434. https://doi.org/10.1093/jmammal/gyz021

Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? Behav Ecol 28:617–623. https://doi.org/10.1093/beheco/arx022

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Boon AK, Réale D, Boutin S (2007) The interaction between personality, offspring fitness and food abundance in North American red squirrels. Ecol Lett 10:1094–1104. https://doi.org/10.1111/j.1461-0248.2007.01106.x

Breed M, Sanchez L (2010) Both environment and genetic makeup influence behavior. Nat. Educ. Knowl. 3:68

Dingemanse NJ, Both C, Drent PJ, et al (2002) Repeatability and heritability of exploratory behaviour in great tits from the wild. Anim Behav 64:929–938. https://doi.org/10.1006/anbe.2002.2006

Dingemanse NJ, Dochtermann NA (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. J Anim Ecol 82:39–54. https://doi.org/10.1111/1365-2656.12013

Dochtermann NA, Schwab T, Anderson Berdal M, et al (2019) The heritability of behavior: a meta-analysis. J Hered 110:403–410. https://doi.org/10.1093/jhered/esz023

Dochtermann NA, Schwab T, Sih A (2015) The contribution of additive genetic variation to personality variation: heritability of personality. Proc Biol Sci 282:20142201. https://doi.org/10.1098/rspb.2014.2201

East ML, Höner OP, Wachter B, et al (2009) Maternal effects on offspring social status in spotted hyenas. Behav Ecol 20:478–483. https://doi.org/10.1093/beheco/arp020

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th ed. Prentice Hall, Harlow England

Gauzere J, Pemberton JM, Morris S, et al (2020) The genetic architecture of maternal effects across ontogeny in the red deer. Evolution (N Y) 74:1378–1391. https://doi.org/10.1111/evo.14000

Hadfield JD (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. J Stat Softw 33:1–22

Hall AJ, McConnell BJ, Barker RJ (2001) Factors affecting first-year survival in grey seals and their implications for life history strategy. J Anim Ecol 70:138–149. https://doi.org/10.1111/j.1365-2656.2001.00468.x

Höner OP, Wachter B, Hofer H, et al (2010) The fitness of dispersing spotted hyaena sons is influenced by maternal social status. Nat Commun 1:60. https://doi.org/10.1038/ncomms1059

Jarman PJ (1987) Group size and activity in eastern grey kangaroos. Anim Behav 35:1044–1050. https://doi.org/10.1016/S0003-3472(87)80161-6

Kalbitzer U, Bergstrom ML, Carnegie SD, et al (2017) Female sociality and sexual conflict shape offspring survival in a Neotropical primate. Proc Natl Acad Sci U S A 114:1892–1897. https://doi.org/10.1073/pnas.1608625114

King WJ, Festa-Bianchet M, Coulson G, Goldizen AW (2017) Long-term consequences of mother-offspring associations in eastern grey kangaroos. Behav Ecol Sociobiol 71:77. https://doi.org/10.1007/s00265-017-2297-1

King WJ, Forsyth DM, Coulson G, Festa-Bianchet M (2015a) Adoption in eastern grey kangaroos: a consequence of misdirected care? PLoS One 10:e0125182. https://doi.org/10.1371/journal.pone.0125182

King WJ, Garant D, Festa-Bianchet M (2015b) Mother-offspring distances reflect sex differences in fine-scale genetic structure of eastern grey kangaroos. Ecol Evol 5:2084–2094. https://doi.org/10.1002/ece3.1498

King WJ, Goldizen AW (2016) Few sex effects in the ontogeny of mother-offspring relationships in eastern grey kangaroos. Anim Behav 113:59–67. https://doi.org/10.1016/j.anbehav.2015.12.020

King WJ, Wilson ME, Allen T, et al (2011) A capture technique for free-ranging eastern grey kangaroos (*Macropus giganteus*) habituated to humans. Aust Mammal 33:47–51. https://doi.org/10.1071/AM10029

Kruuk LEB, Hadfield JD (2007) How to separate genetic and environmental causes of similarity between relatives. J Evol Biol 20:1890–1903. https://doi.org/10.1111/j.1420-9101.2007.01377.x

Mackay AE, Forsyth DM, Coulson G, Festa-Bianchet M (2018) Maternal resource allocation adjusts to timing of parturition in an asynchronous breeder. Behav Ecol Sociobiol 72:7. https://doi.org/10.1007/s00265-017-2419-9

Maestripieri D (2018) Maternal influences on primate social development. Behav. Ecol. Sociobiol. 72:1–12

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Mousseau TA, Fox CW (1998) Maternal effects as adaptations. Oxford Univ. Press.

Munch KL, Noble DWA, Budd L, et al (2018) Maternal presence facilitates plasticity in offspring behavior: insights into the evolution of parental care. Behav Ecol 29:1298–1306. https://doi.org/10.1093/beheco/ary122

Murray CM, Stanton MA, Wellens KR, et al (2018) Maternal effects on offspring stress physiology in wild chimpanzees. Am J Primatol 80:e22525. https://doi.org/10.1002/ajp.22525

Nafus MG, Todd BD, Buhlmann KA, Tuberville TD (2015) Consequences of maternal effects on offspring size, growth and survival in the desert tortoise. J Zool 297:108–114. https://doi.org/10.1111/jzo.12250

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:no-no. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Onyango PO, Gesquiere LR, Wango EO, et al (2008) Persistence of maternal effects in baboons: Mother's dominance rank at son's conception predicts stress hormone levels in subadult males. Horm Behav 54:319–324. https://doi.org/10.1016/j.yhbeh.2008.03.002

Poole W (1975) Reproduction in the two species of grey kangaroos, *Macropus giganteus* Shaw and *M. fuliginosus* (Desmarest). II. Gestation, parturition and pouch life. Aust J Zool 23:333. https://doi.org/10.1071/ZO9750333

Poole WE, Carpenter SM, Wood JT (1982) Growth of grey kangaroos and the reliability of age determination from body measurements I. the eastern grey kangaroo, macropus giganteus. Wildl Res 9:9–20. https://doi.org/10.1071/WR9820009

Quesnel L, King WJ, Coulson G, Festa-bianchet M (2018) Tall young females get ahead: size-specific fecundity in wild kangaroos suggests a steep trade-off with growth. Oecologia 186:59–71. https://doi.org/10.1007/s00442-017-4003-4

Quesnel L, MacKay A, Forsyth DM, et al (2017) Size, season and offspring sex affect milk composition and juvenile survival in wild kangaroos. J Zool 302:252–262. https://doi.org/10.1111/jzo.12453

Räsänen K, Kruuk LEB (2007) Maternal effects and evolution at ecological time-scales. Funct Ecol 21:408–421. https://doi.org/10.1111/j.1365-2435.2007.01246.x

Reddon AR (2012) Parental effects on animal personality. Behav Ecol 23:242–245. https://doi.org/10.1093/beheco/arr210

Reinhold K (2002) Maternal effects and the evolution of behavioral and morphological characters: a literature review indicates the importance of extended maternal care. J Hered 200293(6) 93:400–405 Using

Rossiter MC (1996) Incidence and consequences of inherited environmental effects. Annu Rev Ecol Syst 27:451–476. https://doi.org/10.1146/annurev.ecolsys.27.1.451

Saastamoinen M, Bocedi G, Cote J, et al (2018) Genetics of dispersal. Biol Rev 93:574–599. https://doi.org/10.1111/brv.12356

Samuni L, Tkaczynski P, Deschner T, et al (2020) Maternal effects on offspring growth indicate post-weaning juvenile dependence in chimpanzees (Pan troglodytes verus). Front Zool 17:1–12. https://doi.org/10.1186/s12983-019-0343-8

Schuett W, Dall SRX, Wilson AJ, Royle NJ (2013) Environmental transmission of a personality trait: Foster parent exploration behaviour predicts offspring exploration behaviour in zebra finches. Biol Lett 9:. https://doi.org/10.1098/rsbl.2013.0120

Schülke O, Ostner J, Berghänel A (2019) Prenatal maternal stress effects on the development of primate social behavior. Behav Ecol Sociobiol 73:1–22. https://doi.org/10.1007/s00265-019-2729-1

Smith BR, Blumstein DT (2008) Fitness consequences of personality: a meta-analysis. Behav Ecol 19:448–455. https://doi.org/10.1093/beheco/arm144

Strong JS, Weladji RB, Holand Ø, et al (2017) Personality and fitness consequences of flight initiation distance and mating behavior in subdominant male reindeer (Rangifer tarandus). Ethology 123:484–492. https://doi.org/10.1111/eth.12619

Taylor RW, Boon AK, Dantzer B, et al (2012) Low heritabilities, but genetic and maternal correlations between red squirrel behaviours. J Evol Biol 25:614–624. https://doi.org/10.1111/j.1420-9101.2012.02456.x

Thomson CE, Winney IS, Salles OC, Pujol B (2018) A guide to using a Multiple-Matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance. PLoS One 13:e0197720. https://doi.org/10.1371/journal.pone.0197720

Travis J, Reznick DN (2018) Natural selection: how selection on behavior interacts with selection on morphology. Curr Biol 28:R882–R884. https://doi.org/10.1016/j.cub.2018.07.006

Tyndale-Biscoe H, Renfree M (1987) Reproductive physiology of marsupials (Monographs on Marsupial Biology). Cambridge University Press

Villemereuil P de (2018) Package "QGglmm": Estimate Quantitative Genetics Parameters from Generalised Linear Mixed Models

Villemereuil P de, Schielzeth H, Nakagawa S, Morrissey M (2016) General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics 204:1281–1294. https://doi.org/10.1534/genetics.115.186536

Weiss A, King JE, Figueredo AJ (2000) The heritability of personality factors in chimpanzees (Pan troglodytes). Behav Genet 30:213–221. https://doi.org/10.1023/A:1001966224914

Westrick SE, Taylor RW, Boutin S, et al (2020) Attentive red squirrel mothers have faster-growing pups and higher lifetime reproductive success. Behav Ecol Sociobiol 74:1–13. https://doi.org/10.1101/594093

White SJ, Wilson AJ (2019) Evolutionary genetics of personality in the Trinidadian guppy I: maternal and additive genetic effects across ontogeny. Heredity (Edinb) 122:1–14. https://doi.org/10.1038/s41437-018-0082-1

Williamson DE, Coleman K, Bacanu SA, et al (2003) Heritability of fearful-anxious endophenotypes in infant rhesus macaques: A preliminary report. Biol Psychiatry 53:284–291. https://doi.org/10.1016/S0006-3223(02)01601-3

Wilson AJ (2018) How should we interpret estimates of individual repeatability? Evol Lett 2:4–8. https://doi.org/10.1002/evl3.40

Wilson AJ, Réale D (2006) Ontogeny of additive and maternal genetic effects: Lessons from domestic mammals. Am Nat 167:23–38. https://doi.org/10.1086/498138

Wolf JB, Wade MJ (2001) On the assignment of fitness to parents and offspring: Whose fitness is it and when does it matter? J Evol Biol 14:347–356. https://doi.org/10.1046/j.1420-9101.2001.00277.x

Zablocki-Thomas PB, Herrel A, Karanewsky CJ, et al (2019) Heritability and genetic correlations of personality, life history and morphology in the grey mouse lemur (Microcebus murinus). R Soc Open Sci 6:1–11. https://doi.org/10.1098/rsos.190632

Supplementary Materials

Table S1. Results from Model I: multivariate with a random effect of mother.

	PY Movement		Subadult FID		Juvenile Survival		Maternal FID	
Fixed effects	Estimate	*p-MCMC*	Estimate	*p-MCMC*	Estimate	*p-MCMC*	Estimate	*p-MCMC*
Intercept	0.3 (-0.4, 1.1)	0.508	**8.2** **(6.7, 9.6)**	<0.001	**0.9** **(0.1, 1.8)**	0.040	1.5 (-5.6, 9.2)	0.660
Age at Capture (in months)	0.0 (-0.1, 0.1)	0.606	-	-	-	-	-	-
Sex (Female)	0.1 (0.0, 0.2)	0.188	0.3 (-1.1, 1.5)	0.666	-0.6 (-1.6, 0.5)	0.240	-	-
East-west Coordinate (in metres)	-	-	-	-	-	-	0.001 (0.000, 0.002)	0.090
Group Size	-	-	-	-	-	-	-0.2 (-0.5, 0.0)	0.092
Associate Type:	-	-	-	-	-	-		
- mother with Small Pouch Young (≤6 months-old)	-	-	-	-	-	-	0.7 (-0.3, 1.6)	0.150
- mother with Large Pouch Young (7+ months-old)	-	-	-		-	-	**0.8** **(0.3, 1.5)**	0.010
- mother with 1-year-old offspring	-	-	-	-	-	-	0.4 (-0.3, 1.2)	0.316
Test Number	-	-	-	-	-	-	-0.2 (-0.4, 0.0)	0.064
Difference in Days from the Previous Test	-	-	-	-	-	-	0.001 (0.000, 0.003)	0.092
Year	-0.1	0.212	-2.4	0.002	-0.5	0.388	-0.7	0.278
-2017	(-0.2, 0.1)		(-3.6, -0.8)		(-1.7, 0.5)		(-2.0, 0.6)	
-2018	-0.2 (-0.4, -0.1)	0.010	-1.1 (-2.7, 0.5)	0.176	0.6 (-0.6, 2.0)	0.426	-0.6 (-1.6, 0.3)	0.180
-2019	(base)		(base)		-34.6	<0.001	(base)	

					(-64.8, -3.1)			
Observer (Observer 2)	-	-	-	-	-	-	0.7 (-0.4, 1.6)	0.166

Complete fixed effect results from the four-trait Model I, analyzing repeatability in behavior of eastern grey kangaroos, with response variables: Pouch Young (PY) Movement, Subadult Flight Initiation Distance (FID, averaged), Juvenile Survival, and Maternal FID. The 'base' levels for categorical variables were: sex male, associate type no associate, year 2019 in the behavioral traits and 2016 in Juvenile Survival, and observer Observer 1. As discussed in the main text, there were biases on the 2019 year effect in Juvenile Survival; hence, that particular effect is irrelevant.

Table S2. Results from Model II: bivariate with random effect of offspring and mother.

	Subadult FID		Maternal FID	
Fixed effects	Estimate	*p-MCMC*	Estimate	*p-MCMC*
Intercept	**-13.0 (-23.4, -1.8)**	0.040	-1.5 (-8.7, 6.8)	0.678
Age at Observation (in years)	-0.6 (-1.7, 0.6)	0.320	-	-
Sex (Female)	0.2(-0.8, 1.3)	0.746	-	-
East-west coordinate (in metres)	**0.003 (0.002, 0.004)**	<0.001	**0.001 (0.000, 0.002)**	0.018
Group Size	-0.3 (-0.6, 0.1)	0.162	-0.2 (-0.5, 0.0)	0.082
Associate Type:				
- 1-year-old with mother	-0.6 (-1.5, 0.2)	0.152	-	-
- mother with Small Pouch Young (≤6 months old)	-	-	0.7 (-0.2, 1.6)	0.158
- mother with Large Pouch Young (7+ months old)	-	-	0.8 (0.1, 1.4)	0.014
- mother with 1-year-old offspring	-	-	0.4 (-0.4, 1.1)	0.342
Test Number	**-0.4 (-0.6, -0.1)**	0.010	-0.2 (-0.4, 0.0)	0.052
Difference in Days from the Previous Test	**0.003 (0.001, 0.005)**	0.002	0.001 (0.000, 0.003)	0.128
Year				
-2017	**-2.7 (-4.0, -1.3)**	<0.001	-0.8 (-2.1, 0.4)	0.220
-2018	-0.6 (-1.5, 0.2)	0.170	-0.7 (-1.6, 0.2)	0.154
Observer (Observer 2)	-	-	0.7 (-0.2, 1.7)	0.134

Complete fixed effect results from bivariate Model II, using repeated measures of flight initiation distance (FID) on each subadult and adult female individual eastern grey kangaroo. The 'base' levels for categorical variables were: sex male, associate type no associate, year 2019, and observer Observer 1.

Table S3. Results from the supplementary test on adoptive young FID

	Biological Mother			**Adoptive Mother**		
	Estimate	Std Error	Pr(>|t|)	Estimate	Std Error	Pr(>|t|)
Intercept	7.1	5.324, 1.338	0.252	2.4	6.881, 0.345	0.753
Maternal FID	0.4	0.634, 0.588	0.588	0.6	0.694, 0.910	0.430

Results from linear regressions of maternal (biological or adoptive) flight initiation distance (FID) effects on adopted subadult FID in eastern grey kangaroo (n = 6).

MCMCglmm Codes for Model I

 # Data preparation

- Make sure R reads your factors as characters (i.e., special attention to Animal ID or Year whose levels tend to be numbers, and R automatically reads it as numeric variables)
- Before merging files, make sure to remove NAs from the variables – also, a good practice is to have a variable called 'Data File' before merging, so after the different files are merged, it is easier to do data cleaning if required
- After merging the files, variables that are not common across different responses will receive NAs – transform those NAs to '0', and we will set the models in the way those 0 values will not be used. ATTENTION! Do this only for explanatory variables.

 # Set a prior####

In this case, I have four response variables. I want covariances across all variables for the random effect G1, but residual covariance only for the three last responses listed in the model. I am also fitting the residual variance for the last variable listed to 1 (Juvenile Survival, as it is binary)

```
library(MCMCglmm)

PriorModel1 <- list(G = list(G1 = list(V = diag(4), n = 4)),

                    R = list(R1=list(V=diag(1), n=1),

                             R2 = list(V = diag(3), n = 3,fix=3)))
```

 # Fitting the model####

```
times<-10000 # start with lower values here, and increase depending on autocorrelation

Model_I<-MCMCglmm(cbind(FIDAdFem,AvFIDJuv,AvPYMov, JuvSurv) ~ trait - 1 +

                    at.level(x=trait,level=1):(East+ GroupSize+ TestNumber+
                    RelFemAlong+ Observer+ DifferenceDays)+

                    at.level(x=trait,level=1:2):(Year)+

                    at.level(x=trait,level=2):(Sex)+

                    at.level(x=trait,level=3):(AgeCapture+Sex+Year)+

                    at.level(x=trait,level=4):(Sex+Year),

                  random = ~ us(trait):MotherID,

                  rcov = ~idh(at.level(x=trait, level=1)):units+us(at.level(x=trait,
level=2:4)):units, family = c("gaussian","gaussian", "gaussian","categorical"),

                  data = data.merged, prior = PriorModel1, verbose = T,

                  nitt=1300* times,thin=1* times, burnin=300* times)

plot(Model_I $VCV)

autocorr(Model_I$Sol)

autocorr(Model_I$VCV) # columns that indicate autocorrelation have first lag as '1'. columns that
indicate autocorrelation have first lag as '1'. We want the all values across the consecutive lags to be
below 0. Look which lag is <0.1, and use the lag value as the 'times' object – that way, we will
multiply nitt, thin, and burnin for that value.

summary(Model_I)

save(Model_I,file="Model_I.Rdata")
```

 # Maternal Repeatability PY Mov#### (Gaussian trait)

```
# This is an example of calculating repeatability on pouch young movement – a Gaussian trait

summary(Model_I)

PYMAposterior.repeatability_M <- Model_I$VCV[, "traitAvPYMov:traitAvPYMov.MotherID"]/

  (Model_I$VCV[,"traitAvPYMov:traitAvPYMov.MotherID"] +

     Model_I$VCV[, "at.level(x = trait, level = 2:4)2:at.level(x = trait, level = 2:4)2.units"])

hist(PYMAposterior.repeatability_M)

posterior.mode(PYMAposterior.repeatability_M)

HPDinterval(PYMAposterior.repeatability_M)

save(PYMAposterior.repeatability_M,file="PYMAposterior.repeatability_M_RepeatabilitiesMCMC_
PYM_MotherID.Rdata")


# Within-mother Offspring Repeatability PYM####

summary(Model_I)

PYMAposterior.repeatability_O <- Model_I$VCV[, "at.level(x = trait, level = 2:4)2:at.level(x = trait,
level = 2:4)2.units"]/

  (Model_I$VCV[,"traitAvPYMov:traitAvPYMov.MotherID"] +

     Model_I$VCV[, "at.level(x = trait, level = 2:4)2:at.level(x = trait, level = 2:4)2.units"])

hist(PYMAposterior.repeatability_O)

posterior.mode(PYMAposterior.repeatability_O)

HPDinterval(PYMAposterior.repeatability_O)


save(PYMAposterior.repeatability_O,file="
PYMAposterior.repeatability_O_RepeatabilitiesMCMC.Rdata")
```

 # Maternal Repeatability Juv Surv#### (Binary trait)

```
# This is an example of calculating repeatability on juvenile survival – a non-Gaussian trait
# Data scale back-transformation to original data (0/1).
# Example here on maternal repeatability estimation
library(QGglmm)
RepsSurv_M2<-vector(length=1000)
for (i in 1:1000) {
 prevalues <- predict.MCMCglmm(object = Model_I, it=i, type = "terms")
 qgrSurv_M2<-QGicc(predict =prevalues,
                var.comp  = Model_I$VCV[i,"traitJuvSurv.1:traitJuvSurv.1.MotherID"],
                var.p = Model_I$VCV[i,"traitJuvSurv.1:traitJuvSurv.1.MotherID"]+
                  Model_I$VCV[i,"at.level(x = trait, level = 2:4)3:at.level(x = trait, level = 2:4)3.units"],
                model = "binom1.logit")
 RepsSurv_M2[i]<-qgrSurv_M2$icc.obs
 }
save(RepsSurv_M2,file="RepsSurv_M2_RepeatabilitiesMCMC_YoungSurvival_MotherID.Rdata")
plot(as.mcmc(RepsSurv_M2))
posterior.mode(as.mcmc(RepsSurv_M2), adjust = 1)
mean(RepsSurv_M2)
HPDinterval(as.mcmc(RepsSurv_M2))

# Proportions on the data scale
JuvSurvposterior.repeatability_M <- Model_I$VCV[, "traitJuvSurv.1:traitJuvSurv.1.MotherID"]/
  (Model_I$VCV[,"traitJuvSurv.1:traitJuvSurv.1.MotherID"] +
    Model_I$VCV[, "at.level(x = trait, level = 2:4)3:at.level(x = trait, level = 2:4)3.units"])
hist(JuvSurvposterior.repeatability_M)
posterior.mode(JuvSurvposterior.repeatability_M2)
mean(JuvSurvposterior.repeatability_M)
HPDinterval(JuvSurvposterior.repeatability_M)
```

MCMCglmm Codes for Model II

 # Fitting the model####

```
library(MCMCglmm)

PriorModelII <- list(G = list(G1 = list(V = diag(2), n = 2),
                              G2 = list(V=diag(1), n=1)),
                     R = list(V = diag(2), n = 2))

times<-10000

Model_II<-MCMCglmm(cbind(FIDAdFem,FIDJuv) ~ trait - 1 +
                     at.level(x=trait,level=1:2):(East+ GroupSize+ TestNumber+ RelFemAlong+
                     Year+ DifferenceDays)+
                     at.level(x=trait,level=1):(Observer)+
                     at.level(x=trait,level=2):(AgeatObs+ Sex),
                   random = ~ us(trait):MotherID+idh(at.level(x=trait, level=2)):YoungID,
                   rcov = ~idh(trait):units, family = c("gaussian","gaussian"),
                   data = data.merged2, prior = PriorModelII, verbose = T,
                   nitt=1300*times,thin=1*times,burnin=300*times)

plot(Model_II$VCV)

autocorr(Model_II$Sol)

autocorr(Model_II$VCV)

save(Model_II,file="Model_II_CovarianceForQuestion31000It.rdata")
```

 # Getting repeatability proportions####

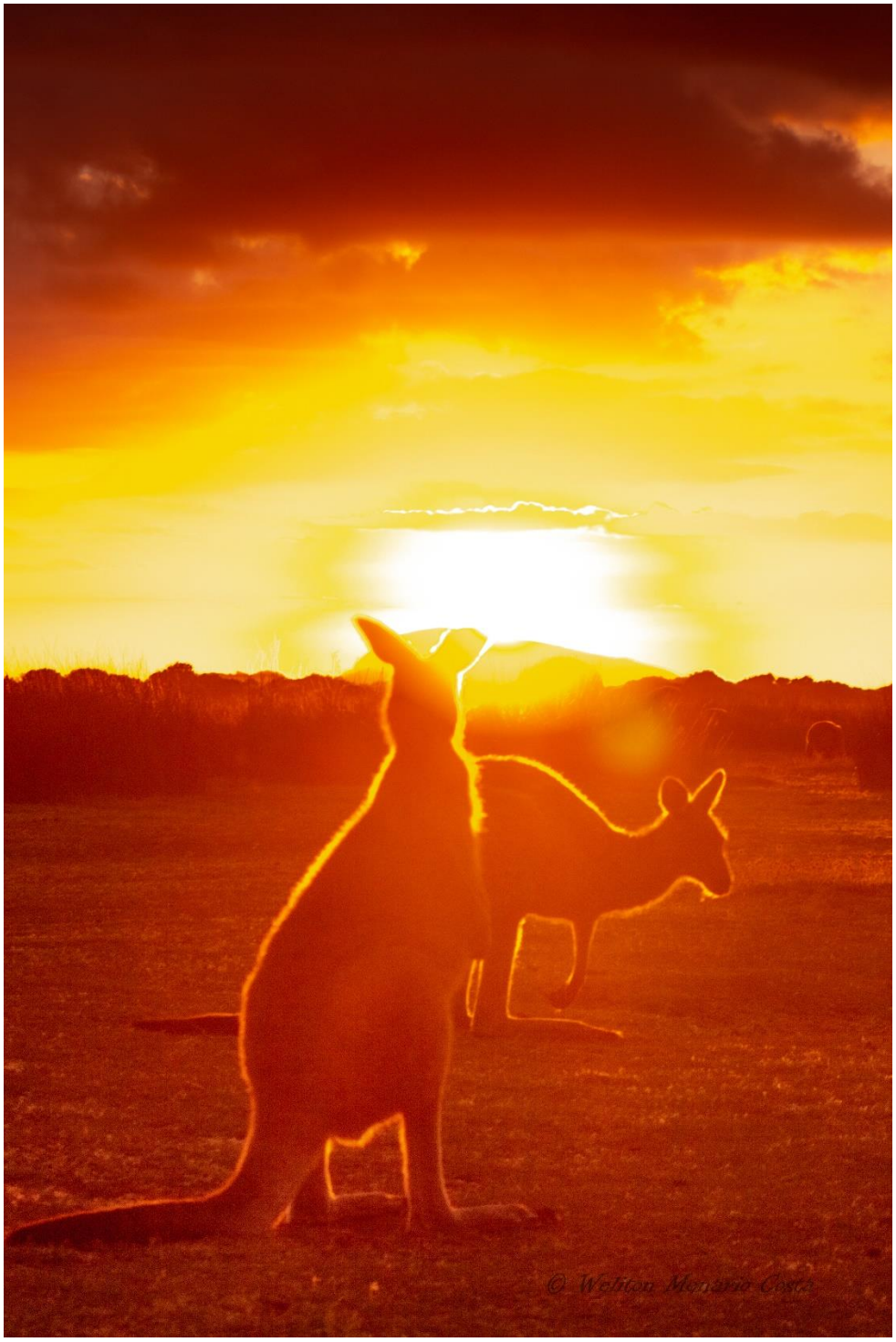
```
# Maternal-level repeatability proportion

rJuvProportion_M<-((Model_II3$VCV[, "traitFIDJuv:traitFIDJuv.MotherID"])/

    ((Model_II3$VCV[,"traitFIDJuv:traitFIDJuv.MotherID"])+(Model_II3$VCV[,"at.level(x = trait,
level = 2).YoungID"])))

mean(rJuvProportion_M)

HPDinterval(rJuvProportion_M)


# Individual-level repeatability proportion

rJuvProportion_O<-((Model_II$VCV[, "at.level(x = trait, level = 2).YoungID"])/

    ((Model_II$VCV[,"traitFIDJuv:traitFIDJuv.MotherID"])+(Model_II$VCV[,"at.level(x = trait,
level = 2).YoungID"])))

mean(rJuvProportion_O)

HPDinterval(rJuvProportion_O)


#Total repeatability

rJuvProportion_Total<-((Model_II$VCV[, "at.level(x = trait, level =
2).YoungID"])+(Model_II$VCV[,"traitFIDJuv:traitFIDJuv.MotherID"]))/

                ((Model_II$VCV[,"traitFIDJuv:traitFIDJuv.MotherID"])+(Model_II$VCV[,"at.level(x
= trait, level = 2).YoungID"])+

                   (Model_II$VCV[,"traitFIDAdFem.units"])+

                   (Model_II$VCV[,"traitFIDJuv.units"]))

mean(rJuvProportion_Total)

HPDinterval(rJuvProportion_Total)
```

 # Estimating an upper limite heritability####

```
h2<-((Model_II$VCV[, "traitFIDAdFem:traitFIDJuv.MotherID"])/
    (Model_II$VCV[,"traitFIDAdFem:traitFIDAdFem.MotherID"]))*2
mean(h2)
HPDinterval(h2)
```

Chapter V.

General Discussion



Overview

Studies of animal personality integrate questions from behavioural and evolutionary ecology, with a primary interest in understanding the causes and consequences of consistent behavioural differences among individuals (or personality). This research field has arguably fashionably (re)branded terms (Beekman and Jordan 2017), attracting extensive public attention and largely influencing current animal behaviour experiments. However, individual behaviour measured through 'personality tests' has often been inappropriately considered a direct measure of 'individual personality' – generating many studies using single-time tested responses (Niemelä and Dingemanse 2017). This practice tends to be supported by a dangerous assumption that evidence for a repeatable behavioural trait in a population means that individual behaviour may always be consistent across all contexts. Indeed, support for repeatability indicates simply the existence of personalities. In the natural world – for any social species, like humans - personality could affect an individual's observed behaviour at a given time point, but so could the social environment. Suppose a behavioural trait reflects personality (i.e., there is support for individual repeatability). In that case, the biological assumption should be that personality (or individual identity) was an important driver of that behaviour, given the factors into play during the experimental tests, such as the social or environmental contexts. The same logic applies for when we quantify the variance explained by different social groups. We can assume that the group social environment is an important driver of behaviour if we find support for among-group variance (or group repeatability). However, we can only quantify a social group variance component if we have properly incorporated the social context in our experiments. Dissecting the importance of both individual and group components allows us to compare which factors (personality or social environment) are the most important drivers of behavioural expression in the wild.

In my thesis, I used a broad range of techniques to characterise behavioural variation from various important angles, using a wild population of a social species. I considered the social context by sampling all individuals' responses in a group and allowing time for their responses to develop under social influences. I dissected the importance of different sources of behavioural variation, such as personality and social group effects (Chapter II). Then, using a more nuanced approach, I tested which factors contributed specifically to group-level differences (Chapter III) and to individual-level differences in behaviour early in life (Chapter IV). I used a wild population of marked eastern grey kangaroos at Wilsons Promontory National Park that has been the subject of long-term monitoring for over a decade. I had three primary questions, which were addressed through a series of experimental tests and behavioural observations. These questions are set out below (Section I), following which I consider (1) the relative importance of individual repeatability versus group effects, (2) the factors driving group-level responses, and (3) causes of among-individual differences in behaviour early in life. In sequence, in Section (II), I discuss the relationship among the main results in the different chapters (personality effects on social structure, the importance of spatial variation, and sex and age effects), the implications of my results to some main topics within my field of research (selection, and heritability), and the advances (Section III), comparison with other studies (Section IV), future directions (Section V), and concluding remarks (Section VI) of my research. In Figure 1, I summarise the main take-home results from my thesis.

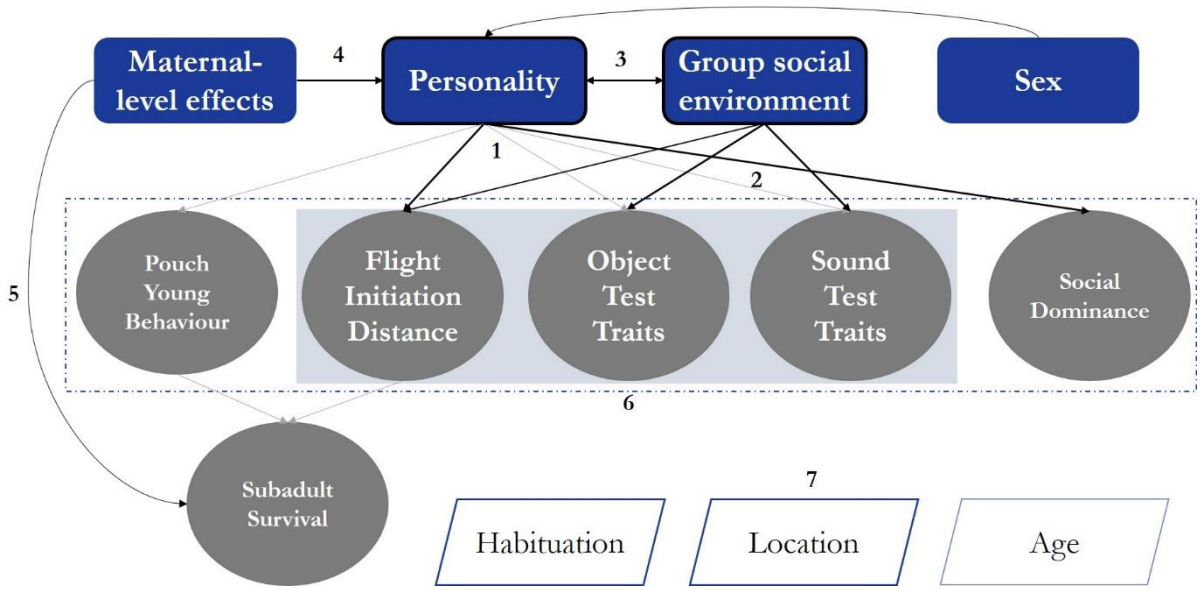


Figure 1. Summary of main results.

Grey circles represent the traits in which I was interested in understanding the causes of variation. The blue boxes represent the overarching concepts that I investigated the implications of for the traits. The large dashed box shows the range of behavioural data across which I have tested for covariances among traits. The central shaded box within the dashed box highlights the experimental traits that I measured on individuals that were at least 13 months of age. The blue-outlined parallelograms on the bottom row show the main aspects that I tested for behavioural plasticity. Bold arrows indicate statistical support, and transparent arrows show a lack of statistical support.

In summary, the main take-home results were:

1. There was strong support for personality differences in FID and Social Dominance, and there was some evidence for sex differences in personality (FID) (Chapter II).
2. The magnitude of variance explained by social group differences was more substantial than individual differences (i.e., personality) in the Object and Sound traits (Chapter II).
3. Personality is one of the determinants of the group-social-environment: every group member's personality contributes to the group effects, and individuals are grouped based on their personalities (Chapter III).

4. Two-thirds of the variability in personality among subadults were attributable to maternal differences (or maternal-level effects) (Chapter IV).
5. Differences between mothers also resulted in differences in offspring survival probability between 6 and 21 months of age (Chapter IV).
6. FID covaried with Object and Sound traits, but not with dominance (Chapter II). In juveniles, pouch young (6-to-10-month-olds) behaviour and subadult (13-to-35-month-olds) FID did not covary together (Chapter IV). Maternal FID was highly correlated with subadult FID (Chapter IV).
7. There was strong support for habituation in FID, Object, and Sound traits (central rectangle in Figure 1), where individuals differed in their habituation rates (Chapter II). There was also strong support for spatial effects on FID, but I did not find age-class effects in any of the traits (Chapter II).

(I) Summary of research results

(1) What is the relative importance of individual repeatability vs group effects for behavioural responses?

In Chapter II, I showed that estimates of among-individual variance (or repeatability) were smaller than of the among-group variance in a series of experimental traits. Among-group variance will be shaped by social plasticity, so my results highlight the importance of the group social environment in shaping behavioural phenotypes in the wild. The estimates of individual repeatability were the highest in the flight initiation distance trait, in which I recorded an individual's immediate fleeing response to an approaching human (Weston et al. 2012). For this trait, I could not quantify the magnitude of variance explained by the social environment effects as the experimental design allowed me to measure only a single individual's response per group. I believe experiments that test for immediate responses of a target individual may not entirely capture how individual behaviour progresses under social influences, leading to less realistic repeatability estimates. I cannot say conclusively from my results that immediate responses will always give higher individual repeatability estimates as I only had one behavioural experiment where I recorded immediate responses (FID). Still, if we measure only a target individual per group, we will not be able to partition the variance due to group social effects. Nevertheless, experiments like the flight initiation distance test are still useful as they allow us to quantify individual behaviour with less group influence – so these traits may be a closer measure of personality and this will be a useful measure to address a range of ecological questions. Considerations on how to incorporate the social context in behavioural experiments and more studies comparing the magnitude of personality and group social environment effects deserve more attention.

(2) Are group-level responses driven by the collective behaviour of all individuals in the group, or by particular individuals' behaviour?

In Chapter III, I found that the average personality of all group members was a better predictor of group-level responses than extreme personalities or elders' personalities, which implies shared or 'democratic' group decisions (Conradt and Roper 2007). In other words, key individuals such as those with extreme personalities or older individuals are not more important than all group members for group-level outcomes, which suggests a lack of leadership. In eastern grey kangaroos, dominance is weak in females (Kaufmann 1975; Maguire et al. 2006) and, despite being strong in males (Miller et al. 2010), dominance does not seem to have much importance for reproductive success (Rioux-Paquette et al. 2015). As males do not form territories or harems (Poole 1982), less dominant individuals can still overlap and mate with females under less competition (Montana et al. 2020). The lack of importance of dominance in this species may be associated with a lack in leadership or unshared decisions (Conradt and Roper 2003). The realisation that shared decisions could occur beyond eusocial insects is relatively recent, with relatively few empirical studies (Conradt and Roper 2007).

(3) How early in life is among-individual variance in behaviour apparent, and to what extent is it shaped by maternal-level effects causing differences among offspring of different mothers?

In Chapter IV, I demonstrated the existence of among-mother variance in pouch young behaviour, a life stage arguably comparable to a late-term fetus in eutherian mammals (Tyndale-Biscoe and Renfree 1987). There was also support for among-mother variance in the behaviour of 1- to 2- year-old subadults, and this variance was more important for total

repeatability than among-subadult variance. Individual repeatability can be shaped by different intrinsic factors (Stoffel et al. 2017; Wilson 2018), and my results indicate that maternal-level variance can drive a large proportion of the estimate of the importance of personality, being possibly more important than age or sex in driving individual differences in behaviour. These results suggest that more attention needs to be given to maternal-level effects in animal personality studies. I could not separate how much maternal-level variance was due to heritable genetics versus maternal effects (Mousseau and Fox 1998; Räsänen and Kruuk 2007), but my upper-limit estimate of heritability was so high that it had to have been inflated by maternal shared environment effects.

(II) Contributions

Results of Chapters II and III

At face value, the results of Chapters II and III may seem slightly contrary: II indicates greater importance of group differences than of individual personality, whereas III indicates that group differences are shaped by individual personalities. Eastern grey kangaroos have personalities, but the behavioural phenotype may be more socially plastic than intrinsically consistent in the wild, dependent on the social context. Personality is one of the determinants of the group-social-environment (but there are also other determinants), and overall when you look at the relative importance of individual personality versus group-social environment, the latter comes out more important. In particular, multiple personalities and the interactions between them make up the group-social-environment, so what Chapter II shows is it is the integrated effect of all the different group environment determinants (into one overall group effect) that is most important. Chapter III shows that there are contributions of every individual to the group effects, and Chapter II that these are best integrated into the overall group effects. These results suggest that personality may be the baseline for the social effects to act on – as if an individual initial response tended to be more influenced by its personality (i.e., a predisposition to behave in a particular way), but as group members experience the behaviour of each other, they adjust collectively. Thus, all group members will influence and adjust to each other's behaviour, leading to stronger behavioural similarities within groups than within individuals. As kangaroos range across different groups within minutes (King and Goldizen 2016), the behavioural phenotype will be less intrinsically consistent across time. Average personality (in this case, the average of the group members' FID scores) was the best predictor of group-level responses and

individuals were grouped based on their personalities, so the large among-group variance found in this thesis may be due mainly to having different personalities composing different groups. Personality is, therefore, largely important for social interactions in eastern grey kangaroos, and social interactions seem to be the most important factor contributing to an individual's behavioural expression.

Social structure

My study shows evidence for the contributions of personality to grouping patterns in the wild. These results are relevant to the field of 'social evolution' as they show that an individual's personality affects how it interacts with others, determining group composition and distribution variation (Croft et al. 2009). Knowledge of how individual variability in behaviour affects sociality helps us to understand nuances of population structure (Aplin et al. 2013). In my study, individuals of different personalities were consistently grouped separately (Chapter III), which indicates an effect of personality on social organisation. Previous studies of another eastern grey kangaroo population used adult females and also found that personality underlines social behaviour (e.g., mean group sizes and the number of preferred associates) (Best et al. 2015; Menz et al. 2017). Other studies on eastern grey kangaroos have also highlighted the importance of spatial structure and density (Best et al. 2014; King et al. 2015) and habitat structure (Neaves et al. 2017) to social organisation. Social behaviour patterns can partially determine the social organisation's nature (Scott 1956), so those previous studies provided some evidence for the importance of personality for social structure. In my thesis, I explicitly tested the contributions of personality to grouping patterns (i.e., group composition), finding that the personalities of individuals grouped together were consistently similar.

Spatial structure

Female kangaroos are philopatric, with individuals tending to spend most of their lives in roughly the same location (King and Goldizen 2016). Spatial structure can then be an important factor contributing to behavioural variation as individuals from different habitats experience different physical and social environments (Dubuc-Messier et al. 2016). I found a very strong effect of location on flight initiation distance – with individuals in the west of the study area being more approachable. I believe that the difference in flight initiation distance across the west-east axis mainly reflects long-term habituation. There is a road crossing the far west of our study site, and tourists regularly visit this area. In contrast, the far-east individuals will rarely see a tourist. Easting did not affect the other behavioural traits (duration of feeding and hopped away) at the individual level, but did affect group-level duration of feeding. A kangaroo could move from the west to the east in only a couple of minutes if it 'decided' to, and vice versa, but that was very rarely observed in the field. Therefore, the reasons why individuals in different areas behave differently are not entirely conclusive. Individuals living in different areas likely have different long-term habituation rates, which means their behaviour was shaped by the environment where they settled, but individuals could also have settled in locations that 'matched their personality' (Holtmann et al. 2017).

Sex and age effects

My study considered individuals of both sexes and all ages, ranging from pouch young to adults. I found very limited support for the contributions of sex and age class to the behavioural phenotype. Demographic attributes such as adult sex ratio and proportion of 2+ year-olds also did not greatly affect behaviour at the group-level. Another set of analyses

was conducted by a Master's student (Carolina Dolan) on the ageing of personality, using my FID data of individuals with precise age only. Male and female FID had different ageing trajectories: females decreased FID in late life, whereas male behaviour became more variable with age (Dolan 2019). Therefore, age seems to be more important for behavioural phenotypes late in life than early in life, and sex effects were not predominant. A recent large meta-analysis using data across several species, from invertebrates to mammals, including our data here, found no evidence of sex differences in personality mean or variability (Harrison, Noble, Jennions, manuscript in preparation).

Selection

Natural selection occurs when differences among individuals' fitness emerge from differences in their phenotypes (e.g., behavioural types), so there is an association among fitness and a phenotypic trait (Kingsolver and Pfennig 2007; Ellegren and Sheldon 2008; Linnen and Hoekstra 2009). Phenotypic variation among individuals is an important factor affecting a population's ecology (e.g., outcomes in competition and dynamics of ecological networks) (Wolf and Weissing 2012), and this variation may also be fundamental for the occurrence of evolutionary processes (Linnen and Hoekstra 2009). Fitness refers to an individual's contributions to future generations, and is usually assessed by its survival and reproductive rates, encompassing, when possible, the success of the successive generations of progeny (Ellegren and Sheldon 2008; Orr 2009). Correlations among fitness and personality traits have been found in some species, indicating that personality is under selection (Dingemanse and Réale 2005; Smith and Blumstein 2008). Some behavioural types may have an increased risk of mortality or reproductive failure than others. For example, willingness to forage under predation risk, tendency to initiate fights against conspecifics,

dispersal tendency, and differences in habitat preferences are common aspects of personality where individuals vary in the level of direct risk they take, but which will all also affect the probability of survival (Wolf and Weissing 2012). Alternatively, depending on the conditions, taking these risks may be advantageous as an individual may benefit from gaining access to more resources (Smith and Blumstein 2008). This logic works clearly in theory, but to date relatively few studies have investigated the implications for fitness of behavioural variation in the wild (Cameron et al. 2009).

I found no support for selection on FID via offspring survival. A Master's student who used the same animals I used here, tested for the effects of age and personality on fitness variation (measures of female or male reproductive success and survival) (Dolan 2019). She extracted individual intercepts from FID linear mixed models to categorise individuals as 'more' or 'less' responsive, and this allowed her to investigate how eastern grey kangaroo ageing in life history traits varied with personality. She found no evidence for differences between the age trajectories of life history traits in more *versus* less responsive individuals. She also found no support for the effect of personality class itself on the fitness measures she used. As individual behaviour tends to vary largely across social groups, we should also find ways of incorporating group effects into our analyses of selection on behaviour as selection will act on the observed phenotype (i.e., on 'behaviour' instead of 'personality'). Furthermore, studies that have found an association between fitness and personality tend to use highly repeatable traits when testing for selection on behaviour (Smith and Blumstein 2008), which may not be frequent forms of behavioural expression in a social species – for example, as in my study system here. Finally, as personality and other morphological or physiological traits may covary (Dammhahn et al. 2018), selection on personality could be indirect – and the associations found may be false positives. Poor

incorporation of the social influences in behavioural measures could even mask some significant behavioural contributions to fitness variation (false negatives).

The habitat and the social environment are variable, and these environmental differences will contribute to an individual differences in behaviour as individuals consistently live in separate locations and experience consistently different social groups (Dochtermann et al. 2012; Dubuc-Messier et al. 2016). Still, individuals may behave consistently differently over and above spatial structure and be grouped with similar personalities, as shown in this thesis. So the immediate behaviour will tend to follow what is happening in the environment, and the social environment is a major driver of the behavioural phenotype. Nevertheless, individual behaviour may show some intrinsic consistency, reflecting individual personality (Stamps and Biro 2016). As the group social environment plays such a substantial effect on behavioural expression, the question of fitness consequences of personality may be less realistic than fitness consequences of behaviour. Pressures act on the individual behavioural phenotype, and social individuals do not act alone. If individuals prioritise adjusting their behaviour to the group, how would personality be under viability selection in wild social animals? I believe that for personality to underline survival variation in the wild, there must be an extreme and consistent selective pressure – for example, constant predator attacks, which is not the case of our population. However, fitness will also be greatly shaped by reproductive success, which could be determined by variation in personality and behaviour, especially for males. Incorporating the social context in selection analysis remains a challenge.

Heritability

Resemblance among relatives can indicate a genetic basis for a trait. However, we also need to distinguish the fractions of the similarity due to additive genetic vs shared environment effects (e.g., maternal effects) (Kruuk and Hadfield 2007). In my study, I have shown that mothers and offspring have a high resemblance in behaviour and that the same mother's offspring also share behavioural similarities. I could not separate additive genetics from maternal effects here, but the very high estimate of the upper limit of heritability indicates significant inflation by maternal environmental effects. Heritability of personality has been found in some species (Dingemanse et al. 2002; Winney et al. 2018), but this has not been largely shown in the wild. Studies suggest that a genetic component partially underlines personality variation and partly attributed to ontogeny and learning (Van Oers and Mueller 2010). Describing personality quantitatively in terms of repeatability (Bell et al. 2009) helps us to address these important evolutionary questions (Dochtermann et al. 2015). Evolutionary change occurs when the distribution of genotypes is altered across generations. It requires natural selection to act on a phenotypic distribution underlined by heritable genetic variation (Kruuk and Hill 2008). Quantifying the genetic basis of a behavioural trait is possibly one of the most challenging parts of any study in the wild. It requires long-term phenotypic data across several generations and a well-developed pedigree (Dingemanse and Dochtermann 2014). In species with slow reproductive cycles and relatively high juvenile mortality, such as large mammals, it is challenging to build a strong pedigree in less than a few decades or without genomic data. This area is one that will be the subject of future research in this study population.

(III) Advances

I believe my project constitutes several useful advances for the study of animal behaviour. I applied a state-of-the-art statistical (co)variance-partitioning Bayesian approach to compare the magnitude of behavioural variance explained by different biological levels, such as individual, social group, and maternal identity. The analyses illustrate how to incorporate the social context in behavioural experiments and how to dissect the variance due to social effects statistically – which we discussed as an angle into the quantification of social plasticity. I also showed how to choose responses from several behavioural traits and illustrated why we should prefer raw trait values over principal component summaries. I also dealt with different data distributions while addressing complex statistical questions, providing guidance and codes.

(IV) Comparison with other studies

Conceptually, I have addressed a range of important topics in the fields of behavioural and evolutionary ecology. Previous studies of another eastern grey kangaroo population found evidence of personality via repeatability in a range of traits (e.g., FID (Best et al. 2015) and vigilance duration (Edwards et al. 2013)). Despite some methodological differences (most notably, use of adult females only, photo-identification of natural marks, smaller sample sizes), their overall conclusions were similar to mine. FID in (Best et al. 2015) had significant individual repeatability estimates (R^2: 84.4%, credible intervals (CI) 83.9%-85.6%)), using adult females across two years, but their estimates were much higher than the ones in my study (R^2: 48.7%, CI 43.6, 56.0%), using individuals of all ages and sexes across three years. Vigilance duration in (Edwards et al. 2013) had some repeatability (R^2: 7.2%), but the credible intervals were not shown. I did not estimate repeatability for vigilance duration in my study, but I did for duration of feeding, which was highly correlated with vigilance duration (Pearson correlation: -0.845 (confidence interval -0.863,-0.826) in the Object test, and -0.855 (-0.874,-0.833) in the Sound test). In my study, duration of feeding had R^2 of 7.9% (CI 1.9%, 15.3%) in the Object test, but it did not differ from zero in the Sound test (Chapter II). In the (Best et al. 2015) study, there was a positive correlation between individual average FID and average vigilance duration, which is similar to the negative covariance between FID and feeding duration in my study (Chapter II). Results for plasticity were also similar between my study and others on eastern grey kangaroos. Individuals habituated to FID tests (response decreased) (Edwards et al. 2013; Austin and Ramp 2019), and an increase in group size decreased FID response (Best et al. 2015) and anti-predator vigilance (Favreau et al. 2010), and increased bite rate (Favreau et al. 2018). Finally, EGK studies have found personality to underline social behaviour. For example, FID was

associated with the number of preferred associates, and more reactive females were found in bigger groups (Best et al. 2015), and individuals differed in their sociability level (mean group size and mean nearest-neighbour distances) (Menz et al. 2017). Like my results, these studies highlight ways by which personality can contribute to the social organisation's nature. My study's main advantage compared to these studies is that I partitioned variance-covariance due to individual versus social group effects, using large sample sizes of marked individuals of all ages and sexes, across three years of observations.

Many other aspects of my thesis have not been addressed in marsupials before, for example, individual differences in habituation, among-group variance, and contributions of personality to group-level decisions and grouping patterns. However, they have been considered in studies of other taxa. For example, evidence for individual differences in habituation has been shown in threespine sticklebacks (*Gasterosteus aculeatus*) (Bell and Peeke 2012) and great tits (*Parus major)* (Dingemanse et al. 2012), but these studies were conducted in the laboratory or using captive wild animals, retesting individuals only twice – different from mine, which tested individuals in the wild and multiple times. A study of threespine sticklebacks (Jolles et al. 2018) also found among-group variance in behaviour, but they used individuals in the laboratory and did not quantify individual repeatability. Studies of feral guppies (*Poecilia reticulata*) (Brown and Irving 2014) and of cockroaches (*Periplaneta americana*) (Planas-Sitjà 2020) tested the effect of different personalities to group-level responses, and in contrast to my study, found strong support for a disproportionate influence of particular personalities on group-level responses. These studies were carried out in the laboratory or used simulated data, and the researchers artificially allocated individuals in each group. Personality has also been shown to underpin multiple aspects of network structure in bottlenose dolphins (*Tursiops truncates*) (Díaz López 2020), great tits (Aplin et al. 2013),

Trinidadian guppies (*Poecilia reticulata*) (Croft et al. 2009), and sticklebacks (Pike et al. 2008). These results are similar to my evidence of personality contributing to grouping patterns, but only the first two studies were conducted in the wild.

Furthermore, some aspects related to causes and consequences of personality differences have only been addressed in other taxa, such as analysis of behavioural variance among non-siblings, maternal-offspring behavioural covariance, and contributions of personality to fitness variation. Heritable genetic or maternal effects were found in North American red squirrels (*Tamiasciurus hudsonicus*) (Taylor et al. 2012), grey mouse lemurs (*Microcebus murinus*) (Zablocki-Thomas et al. 2019), and Trinidadian guppies (*Poecilia reticulata*) (White and Wilson 2019). In these studies, however, behaviour was measured from captured or laboratory animals and not all studies have shown the impact of maternal-level variance on repeatable individual variation. As advantages, some of these studies have separated additive genetic effects from maternal effects – which is a limitation of my study, as I was not able to do so. Mother-offspring behavioural covariance has been reported in red squirrels (Taylor et al. 2012) and grey mouse lemurs (Zablocki-Thomas et al. 2019) (i.e., similarities in personality traits), and rhesus macaques (*Macaca mulatta*), baboons (*Papio* sp.), vervet monkeys (*Cercopithecus aethiops sabaeus*) (see review: (Maestripieri 2018)), and spotted hyenas (East et al. 2009) (i.e., similarities in dominance rank). Some of these studies have conducted an overall association among maternal and offspring traits, not using multivariate models and a variance-partitioning approach to dissect covariance across mother and residuals as in this thesis. Several studies have tested for personality and fitness covariance (see these reviews: (Smith and Blumstein 2008), (Travis and Reznick 2018), (Wolf and Wade 2001)). However, most of these studies have not been carried out in the wild or have

considered secondary terms to refer to personality traits (such as boldness, activity, exploration), making a comparison across species difficult.

(V) Future directions

My work suggests several avenues for further research, exploring aspects of the system here that would follow from work in this thesis. In particular, social network analysis offers a means of studying social interaction patterns to characterise the social structure as it allows to incorporate the (non-) associations among individuals quantitatively. This framework allows us to test for the importance of individual attributes (such as sex) and traits (such as personality) for the network structure (Aplin et al. 2013; Belton et al. 2018). It also allows us to estimate metrics of 'sociability' for each individual, such as their number of frequent preferences or avoidances (Strickland et al. 2017) or how important (or 'central') an individual is within its network cluster (Menz et al. 2017). Analyses of networks in this population have focused on maternal-offspring associations, showing that time spent with the mother decreased as young aged, and also that the association duration differed among juveniles (King and Goldizen 2016). Juveniles with strong associations with their mothers before weaning (<21-month-olds) had a beneficial rates of growth and survival (King et al. 2017). This research could be further extended to other types of associations, testing for covariance among individual sociability metrics and the behavioural traits discussed here, and hence generating greater understanding of how personality contributes to social organisation.

Further exploration of the associations among behaviour and life history would elucidate natural selection patterns acting on behaviour. In particular, it would be interesting to explore any associations among behaviour and life-history traits at many life stages and across sexes, exploiting to the maximum our long-term data basis on survival and reproductive success of females and males. Our results show the importance of the social context for the individual behavioural phenotype, and also how the social environment may play an important role in selection outcomes (Cameron et al. 2009). Further challenges

include the quantification of the social environmental effects for fitness variation of behaviour, which will be particularly interesting in species that live in fission-fusion dynamics. Determination of the relatedness among individuals using genomic data (Gienapp et al. 2017) will allow the construction of a pedigree that is required to dissect additive genetic from maternal effects on behavioural variation in a population (Kruuk and Hill 2008).

I would also like to use this part of the thesis to suggest guidance for future research on the topic, based on my experience with the study and the data analysis here. In particular, I encourage more studies using experimental designs that allow measurement of the response of all group members at a time, allowing the responses to develop under natural social influences. This way, more studies can quantify group environmental effects in other species and compare them with personality effects, as I have done in this thesis. It is important to be careful about knowing well what we measure through our experiments to be more precise on the types of ecological interpretations we can make based on our results. When designing an experiment, be sure to pay attention to whether the setting is the most natural possible. If studying a social species' behaviour, be clear if you want to measure a response that progressed under the social influence of group members or wants an instant response with fewer social influences. With this distinction clearer, you will be less likely to engage blindly in the search for high individual repeatability estimates. I also encourage future research to opt to use raw trait values over principal component summaries. This choice helps you avoid secondary terms and be more explicit about the trait's meaning when making interpretations. My final advice is to avoid 'statistics on statistics' – rather than deriving one set of statistics from a first step of analyses and then using them in a second

step, use statistical tools to address all questions in a single go, as shown throughout this thesis.

(VI) Concluding remarks

What excites me the most about this study is the interplay among personality and group effects: the group social environment is more important than personality for the observed behaviour, but personality is still an essential factor driving group-level responses and social interactions. My results suggest that personality is the baseline for social influences to act on, and the personalities of all group members influence all group members' behaviour expression, so group-level 'decisions' are effectively shared. These results suggest a minimal role of 'leadership' in eastern grey kangaroos. The contributions of personality to behaviour, under the social context, are therefore mostly indirect. Therefore, a particular individual's behaviour is more likely to change as it ranges across social groups rather than be intrinsically consistent. My findings could impact the study of selection and other evolutionary questions in species living in fission-fusion dynamics. In summary, my thesis highlights the importance of incorporating the social environment created by groups of conspecifics in behavioural experiments and statistical analyses in social species.

References

Aplin LM, Farine DR, Morand-Ferron J, et al (2013) Individual personalities predict social behaviour in wild networks of great tits (Parus major). Ecol Lett 16:1365–1372. https://doi.org/10.1111/ele.12181

Austin CM, Ramp D (2019) Flight responses of eastern gray kangaroos to benign or harmful human behavior. Ecol Evol 9:13824–13834. https://doi.org/10.1002/ece3.5818

Croft DP, Krause J, Darden SK, Ramnarine IW, Faria JJ, James R (2009) Behavioural trait assortment in a social network: patterns and implications. Behav Ecol Sociobiol 63(10), 1495-1503. https://doi:10.1007/S00265-009-0802-X

Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? Behav Ecol 28:617–623. https://doi.org/10.1093/beheco/arx022

Bell AM, Hankison SJ, Laskowski KL (2009) The repeatability of behaviour: a meta-analysis. Anim Behav 77:771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Bell AM, Peeke HVS (2012) Individual variation in habituation: behaviour over time toward different stimuli in threespine sticklebacks (Gasterosteus aculeatus). Behaviour 149:1339–1365. https://doi.org/10.1163/1568539X-00003019

Belton LE, Cameron EZ, Dalerum F (2018) Social networks of spotted hyaenas in areas of contrasting human activity and infrastructure. Anim Behav 135:13–23. https://doi.org/10.1016/j.anbehav.2017.10.027

Best EC, Blomberg SP, Goldizen AW (2015) Shy female kangaroos seek safety in numbers and have fewer preferred friendships. Behav Ecol 26:639–646. https://doi.org/10.1093/beheco/arv003

Best EC, Dwyer RG, Seddon JM, Goldizen AW (2014) Associations are more strongly correlated with space use than kinship in female eastern grey kangaroos. Anim Behav 89:1–10. https://doi.org/10.1016/j.anbehav.2013.12.011

Brown C, Irving E (2014) Individual personality traits influence group exploration in a feral guppy population. Behav Ecol 25:95–101. https://doi.org/10.1093/beheco/art090

Cameron EZ, Setsaas TH, Linklater WL (2009) Social bonds between unrelated females increase reproductive success in feral horses. Proc Natl Acad Sci U S A 106:13850–13853. https://doi.org/10.1073/pnas.0900639106

Conradt L, Roper T. (2007) Democracy in animals: the evolution of shared group decisions. Proc R Soc B Biol Sci 274:2317–2326. https://doi.org/10.1098/rspb.2007.0186

Conradt L, Roper TJ (2003) Group decision-making in animals. Nature 421:155–158. https://doi.org/10.1038/nature01294

Croft DP, Krause J, Darden SK, et al (2009) Behavioural trait assortment in a social network: patterns and implications. Behav Ecol Sociobiol 63:1495–1503. https://doi.org/10.1007/s00265-009-0802-x

Dammhahn M, Dingemanse NJ, Niemelä PT, Réale D (2018) Pace-of-life syndromes: a framework for the adaptive integration of behaviour, physiology and life history. Behav Ecol Sociobiol 72:62. https://doi.org/10.1007/s00265-018-2473-y

Díaz López B (2020) When personality matters: personality and social structure in wild bottlenose dolphins, Tursiops truncatus. Anim Behav 163:73–84. https://doi.org/10.1016/j.anbehav.2020.03.001

Dingemanse NJ, Both C, Drent PJ, et al (2002) Repeatability and heritability of exploratory behaviour in great tits from the wild. Anim Behav 64:929–938. https://doi.org/10.1006/anbe.2002.2006

Dingemanse NJ, Bouwman KM, van de Pol M, et al (2012) Variation in personality and behavioural plasticity across four populations of the great tit Parus major. J Anim Ecol 81:116–126. https://doi.org/10.1111/j.1365-2656.2011.01877.x

Dingemanse NJ, Dochtermann NA (2014) Individual behaviour: behavioural ecology meets quantitative genetics. In: CharmantierA., Garant D, Kruuk LEB (eds) Quantitative Genetics in the Wild. Oxford University Press, pp 54–67

Dingemanse NJ, Réale D (2005) Natural selection and animal personality. Behaviour 142:1159–1184. https://doi.org/10.1163/156853905774539445

Dochtermann NA, Jenkins SH, Swartz MJ, Hargett AC (2012) The roles of competition and environmental heterogeneity in the maintenance of behavioral variation and covariation. Ecology 93:1330–1339. https://doi.org/10.1890/11-1025.1

Dochtermann NA, Schwab T, Sih A (2015) The contribution of additive genetic variation to personality variation: heritability of personality. Proc Biol Sci 282:20142201. https://doi.org/10.1098/rspb.2014.2201

Dolan C (2019) Ageing and senescence in a wild eastern grey kangaroo (Macropus giganteus) population. The Australian National University

Dubuc-Messier G, Réale D, Perret P, Charmantier A (2016) Environmental heterogeneity and population differences in blue tits personality traits. Behav Ecol 28:arw148. https://doi.org/10.1093/beheco/arw148

East ML, Höner OP, Wachter B, et al (2009) Maternal effects on offspring social status in spotted hyenas. Behav Ecol 20:478–483. https://doi.org/10.1093/beheco/arp020

Edwards AM, Best EC, Blomberg SP, Goldizen AW (2013) Individual traits influence vigilance in wild female eastern grey kangaroos. Aust J Zool 61:332–341. https://doi.org/10.1071/ZO13025

Ellegren H, Sheldon BC (2008) Genetic basis of fitness differences in natural populations. Nature 452:169–175. https://doi.org/10.1038/nature06737

Favreau F-R, Goldizen AW, Pays O (2010) Interactions among social monitoring, anti-predator vigilance and group size in eastern grey kangaroos. Proc R Soc B Biol Sci 2089–2095. https://doi.org/10.1098/rspb.2009.2337

Favreau FR, Goldizen AW, Fritz H, Pays O (2018) Food supply fluctuations constrain group sizes of kangaroos and in turn shape their vigilance and feeding strategies. Anim Behav 135:165–176. https://doi.org/10.1016/j.anbehav.2017.11.020

Gienapp P, Fior S, Guillaume F, et al (2017) Genomic quantitative genetics to study evolution in the wild. Trends Ecol Evol 32:897–908. https://doi.org/10.1016/j.tree.2017.09.004

Holtmann B, Santos ESA, Lara CE, Nakagawa S (2017) Personality-matching habitat choice, rather than behavioural plasticity, is a likely driver of a phenotype – environment covariance. Proc R Soc B 284:

Jolles JW, Laskowski KL, Boogert NJ, Manica A (2018) Repeatable group differences in the collective behaviour of stickleback shoals across ecological contexts. Proc R Soc B Biol Sci 285:20172629. https://doi.org/10.1098/rspb.2017.2629

Kaufmann JH (1975) Field observations of the social behaviour of the eastern grey kangaroo, *Macropus giganteus*. Anim Behav 23:214–221. https://doi.org/10.1016/0003-3472(75)90066-4

King WJ, Festa-Bianchet M, Coulson G, Goldizen AW (2017) Long-term consequences of mother-offspring associations in eastern grey kangaroos. Behav Ecol Sociobiol 71:77. https://doi.org/10.1007/s00265-017-2297-1

King WJ, Garant D, Festa-Bianchet M (2015) Mother-offspring distances reflect sex differences in fine-scale genetic structure of eastern grey kangaroos. Ecol Evol 5:2084–2094. https://doi.org/10.1002/ece3.1498

King WJ, Goldizen AW (2016) Few sex effects in the ontogeny of mother-offspring relationships in eastern grey kangaroos. Anim Behav 113:59–67. https://doi.org/10.1016/j.anbehav.2015.12.020

Kingsolver JG, Pfennig DW (2007) Patterns and power of phenotypic selection in nature. Bioscience 57:561–572. https://doi.org/10.1641/B570706

Kruuk LEB, Hadfield JD (2007) How to separate genetic and environmental causes of similarity between relatives. J Evol Biol 20:1890–1903. https://doi.org/10.1111/j.1420-9101.2007.01377.x

Kruuk LEB, Hill WG (2008) Introduction. Evolutionary dynamics of wild populations: The use of long-term pedigree data. Proc R Soc B Biol Sci 275:593–596. https://doi.org/10.1098/rspb.2007.1689

Linnen CR, Hoekstra HE (2009) Measuring natural selection on genotypes and phenotypes in the wild. Cold Spring Harb Symp Quant Biol 74:155–168. https://doi.org/10.1101/sqb.2009.74.045

Maestripieri D (2018) Maternal influences on primate social development. Behav. Ecol. Sociobiol. 72:1–12

Maguire G, Ramp D, Coulson G (2006) Foraging behaviour and dispersion of eastern grey kangaroos (*Macropus giganteus*) in an ideal free framework. J Zool 268:261–269. https://doi.org/10.1111/j.1469-7998.2005.00026.x

Menz CS, Goldizen AW, Blomberg SP, et al (2017) Understanding repeatability and plasticity in multiple dimensions of the sociability of wild female kangaroos. https://doi.org/10.1016/j.anbehav.2017.01.012

Miller EJ, Eldridge MDB, Cooper DW, Herbert CA (2010) Dominance, body size and internal relatedness influence male reproductive success in eastern grey kangaroos (*Macropus giganteus*). Reprod Fertil Dev 22:539–549. https://doi.org/10.1071/RD09061

Montana L, Rousseu F, Garant D, Festa-Bianchet M (2020) Siring success in kangaroos: size matters for those in the right place at the right time. Behav Ecol 31:750–760. https://doi.org/10.1093/beheco/araa020

Mousseau TA, Fox CW (1998) Maternal effects as adaptations. Oxford Univ. Press.

Neaves LE, Roberts MW, Herbert CA, Eldridge MDB (2017) Limited sex bias in the fine-scale spatial genetic structure of the eastern grey kangaroo and its relationship to habitat. Aust J Zool 65:33. https://doi.org/10.1071/ZO16051

Niemelä PT, Dingemanse NJ (2017) Individual versus pseudo-repeatability in behaviour: Lessons from translocation experiments in a wild insect. J Anim Ecol 86:1033–1043. https://doi.org/10.1111/1365-2656.12688

Orr HA (2009) Fitness and its role in evolutionary genetics. Nat Rev Genet 10:531–539. https://doi.org/10.1038/nrg2603

Pike TW, Samanta M, Lindström J, Royle NJ (2008) Behavioural phenotype affects social interactions in an animal network. Proc R Soc B Biol Sci 275:2515–2520. https://doi.org/10.1098/rspb.2008.0744

Planas-Sitjà I (2020) Personality variation improves collective decision-making in cockroaches. Behav Processes 177:. https://doi.org/10.1016/j.beproc.2020.104147

Poole WE (1982) *Macropus giganteus*. Mamm Species 1. https://doi.org/10.2307/3504005

Räsänen K, Kruuk LEB (2007) Maternal effects and evolution at ecological time-scales. Funct Ecol 21:408–421. https://doi.org/10.1111/j.1365-2435.2007.01246.x

Rioux-Paquette E, Garant D, Martin AM, et al (2015) Paternity in eastern grey kangaroos: moderate skew despite strong sexual dimorphism. Behav Ecol 26:1147–1155. https://doi.org/10.1093/beheco/arv052

Scott JP (1956) The analysis of social organization in animals. Ecology 37:213–221

Smith BR, Blumstein DT (2008) Fitness consequences of personality: a meta-analysis. Behav Ecol 19:448–455. https://doi.org/10.1093/beheco/arm144

Stamps JA, Biro PA (2016) Personality and individual differences in plasticity. Curr. Opin. Behav. Sci. 12:18–23

Stoffel MA, Nakagawa S, Schielzeth H (2017) rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol Evol 8:1639–1644. https://doi.org/10.1111/2041-210X.12797

Strickland K, Levengood A, Foroughirad V, et al (2017) A framework for the identification of long-term social avoidance in longitudinal datasets. R Soc open Sci 4:170641. https://doi.org/10.1098/rsos.170641

Taylor RW, Boon AK, Dantzer B, et al (2012) Low heritabilities, but genetic and maternal correlations between red squirrel behaviours. J Evol Biol 25:614–624. https://doi.org/10.1111/j.1420-9101.2012.02456.x

Travis J, Reznick DN (2018) Natural selection: how selection on behavior interacts with selection on morphology. Curr Biol 28:R882–R884. https://doi.org/10.1016/j.cub.2018.07.006

Tyndale-Biscoe H, Renfree M (1987) Reproductive physiology of marsupials (Monographs on Marsupial Biology). Cambridge University Press

Van Oers K, Mueller JC (2010) Evolutionary genomics of animal personality. Philos Trans R Soc B Biol Sci 365:3991–4000. https://doi.org/10.1098/rstb.2010.0178

Weston MA, Mcleod EM, Blumstein DT, Guay PJ (2012) A review of flight-initiation distances and their application to managing disturbance to Australian birds. Emu - Austral Ornithol. 112:269–286

White SJ, Wilson AJ (2019) Evolutionary genetics of personality in the Trinidadian guppy I: maternal and additive genetic effects across ontogeny. Heredity (Edinb) 122:1–14. https://doi.org/10.1038/s41437-018-0082-1

Wilson AJ (2018) How should we interpret estimates of individual repeatability? Evol Lett 2:4–8. https://doi.org/10.1002/evl3.40

Winney IS, Schroeder J, Nakagawa S, et al (2018) Heritability and social brood effects on personality in juvenile and adult life-history stages in a wild passerine. J Evol Biol 31:75–87. https://doi.org/10.1111/jeb.13197

Wolf JB, Wade MJ (2001) On the assignment of fitness to parents and offspring: Whose fitness is it and when does it matter? J Evol Biol 14:347–356. https://doi.org/10.1046/j.1420-9101.2001.00277.x

Wolf M, Weissing FJ (2012) Animal personalities: consequences for ecology and evolution. Trends Ecol Evol 27:452–461. https://doi.org/10.1016/j.tree.2012.05.001

Zablocki-Thomas PB, Herrel A, Karanewsky CJ, et al (2019) Heritability and genetic correlations of personality, life history and morphology in the grey mouse lemur (Microcebus murinus). R Soc Open Sci 6:1–11. https://doi.org/10.1098/rsos.190632