THE UNIVERSITY *of* EDINBURGH

Coping with stress: personality, life history and social dominance in swordtail fishes, *Xiphophorus sp.*

Kay Boulton

"Everyone knows what stress is and no-one knows what it is."

(Selye, 1973)

Submitted for the degree of

Doctor of Philosophy

The University of Edinburgh

2014

LAY SUMMARY

Competition for resources plays an important role in natural selection, creating winners and losers. Winners become socially dominant, obtain resources and so increase their fitness at the expense of losers. Provided they are heritable, phenotypic traits promoting competitive success will be inherited by subsequent generations. Thus, while resource dependent traits (e.g. growth) that rely on competitive outcomes are widely recognised as being under strong selection, this is also likely to be the case for those traits that determine competitive ability and social dominance. In addition, competition is expected to be an important source of stress, for example, harassment of subordinates by dominant individuals. Consequently individual fitness may depend not only on the ability to win resources, but also on the ability to cope with stress. This thesis proposes that social dominance is not just a simple consequence of body size or weaponry, but rather that the interplay between growth, repeatable behavioural characteristics (i.e., personality), and the ability to cope with social and environmental stressors are equally important factors. Thus the *dynamic of dominance* arises, a model that highlights the expectation of complex relationships between traits causal and consequent to social dominance. Here, empirical studies of *Xiphophorus sp.* are used to test each element in the model. First the concept of individual personality is explored, asking to what extent it is really stable over long periods of time (equivalent to life-spans). Next, the links between behaviour, physiological stress and contest outcome are considered and, using a repeated measures approach, the hypothesis that individuals differ in "stress coping style" is evaluated. Finally, using a quantitative genetic approach the genetic relationship is estimated between behavioural and life history traits under experimentally manipulated levels of competition. In this way the contribution of genetic and environmental effects to the patterns of trait (co)variation that make up the *dynamic of dominance* is assessed.

DECLARATION

The work described in this thesis has been carried out by myself or as otherwise acknowledged below. It is entirely of my own composition, the main text and bibliography do not exceed 70,000 words and it has not, in whole or part, been submitted for any other degree.

Chapters 2, 4 and 5 are based on experiments designed by K Boulton, AJ Wilson and CA Walling. All data were collected by K Boulton from 2010-2013 with assistance where necessary from AJ Grimmer, at the University of Edinburgh and the Tremough campus of Exeter University. AJ Wilson, CA Walling and JD Hadfield assisted with statistical models. K Boulton analysed the data and wrote the manuscripts.

Chapter 3 includes data provided by RL Earley, collated and scored by two students, RM Pearce and B Sinderman, University of Alabama, USA. AJ Wilson assisted with statistical models. K Boulton analysed the data and wrote the manuscript.

Chapter 4 includes data from samples collected and extracted by K Boulton with RIA performed by Elsa Couto in the laboratory of AVM Canario, CCMar, Faro, Portugal during 2013. K Watt, AVM Canario and RL Earley provided assistance with hormone collection design.

Kay Boulton

March 2014

DEDICATION

Patrem meum Bowden Dennis

(1935-1964)

ABSTRACT

Competition for resources plays an important role in natural selection, creating winners and losers. Winners become socially dominant, obtain resources and so increase their fitness at the expense of losers. Provided they are heritable, phenotypic traits promoting competitive success will be inherited by subsequent generations. Thus, while resource dependent traits (e.g. growth) that rely on competitive outcomes are widely recognised as being under strong selection, this is also likely to be the case for those traits that determine competitive ability and social dominance. In addition, competition is expected to be an important source of stress, for example, harassment of subordinates by dominant individuals. Consequently individual fitness may depend not only on the ability to win resources, but also on the ability to cope with stress. This thesis proposes that social dominance is not just a simple consequence of body size or weaponry, but rather that the interplay between growth, repeatable behavioural characteristics (i.e. personality), and the ability to cope with social and environmental stressors are equally important factors. Thus the *dynamic of dominance* arises, a model that highlights the expectation of complex relationships between traits causal and consequent to social dominance. Here, empirical studies of *Xiphophorus sp.* are used to test each element in the model. First the concept of individual personality is explored, asking to what extent it is really stable over long periods of time (equivalent to life-spans). Next, the links between behaviour, physiological stress and contest outcome are considered and, using a repeated measures approach, the hypothesis that individuals differ in stress coping style is evaluated. Finally, using a quantitative genetic approach the additive genetic variance-covariance matrix (**G**) is estimated between behavioural and life history traits under experimentally manipulated levels of competition. In this way the contribution of genetic and environmental effects to the patterns of trait (co)variation that make up the *dynamic of dominance* is assessed.

ACKNOWLEDGEMENTS

I am indebted to my main supervisors Alastair J Wilson and Craig A Walling for three and a half years of enthusiasm, help, humour and endless, infinite patience. Without them this thesis would not have been possible, and there are not large enough bottles of whisky that will thank them adequately. Additionally I must also thank Per T Smiseth and Sarah Reece for enthusiasm and encouragement, especially during the first year of this work.

Special thanks also to the following: Andrew J "Grim" Grimmer for fish expertise, help in the lab and again, endless patience; Jarod D Hadfield, Ian MS White and Neils J Dingemanse for (extra) stats advice and wit along the way; Gil G Rosenthal for sending the little fishy fishes from Mexico; Kathryn Watt, Ryan L Earley, Adelino VM Canario and Elsa Couto who helped, advised and offered yet more patience in the endless quest for cortisol.

I am also indebted to the Institute of Evolutionary Biology PhD committee for trusting me with the opportunity to complete this work, to the Engineering and Physical Research Council for PhD funding, the Edinburgh Alliance for Complex Traits Genetics for statistics course funding and the Genetics Society for a Hereditary Fieldwork Grant that funded the visit to Adelino's lab in Portugal.

To my friends, family and loved ones: thank you all for your support and encouragement. Special thanks to the exceptionally talented Class of 2014 (Sarah Mattey (Smattey), Lucy Carter, Elisa Schaum, Emma Hodcroft (honorary member), Hannah Froy, Reuben Nowell, Kevin Donnelly, Emily Moore and Cheryl Gibbons), Classes of 2013 (especially Steph Smith, Jess Hedge and Laura Appleby) and 2015 (especially Elisa Anastasia, Christine Tansey and Jess Flood), parents June and George, daughters Emma and Jennie, sisters JoJo and Denise, friends Anna Murray from Currie, Tango Bob, Jean Budge, Sarah from JFK, the three Jonathans (Moss, Cooke and Strassberg), Maggie and Geoff and last but not least Marmalade, for helping to keep body and soul together.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 SUMMARY

Throughout the living world competition exists for the energy resources, territory and mates that are essential for survival, growth, reproduction and ultimately, fitness (West-Eberhard, 1979). Competition plays an important role in natural selection because it creates winners and losers, with winners becoming socially dominant and thus increasing their fitness at the expense of those that lose (Brockelman, 1975). Provided they are heritable, phenotypic traits that promote success in particular environments will be selected for and preferentially passed to the next generation. In this way traits influencing competitive social dominance, defined simply as an individual's repeatable ability to win resources in competition (Wilson et al., 2011a) are expected to be under selection. The same is true for those resource-dependent life history traits (e.g. growth, maturation, fecundity and survival) that depend on the outcome of competition (following Wilson et al., 2011a; Wilson, 2014).

However, competition can also influence fitness by a second route. Specifically, behavioural interactions associated with competition are expected to be an important source of stress that can negatively influence fitness when exposure becomes chronic, i.e. is prolonged and beyond individual control, (e.g. Pickering and Pottinger, 1989; Blanchard et al., 1998; Wingfield et al., 1998; Gregory and Wood, 1999; Barton, 2002; Goymann and Wingfield, 2004). While social stress is not experienced equally by all members of a group (e.g. harassment of subordinates is usually by dominant animals) individuals also differ in their behavioural and physiological responses to stress. Consequently, under competition individual fitness may depend not only on the ability to win resources, but also on the ability to cope with stress. This thesis proposes that social dominance is not just a simple consequence of body size or weaponry; rather, the interplay between growth, repeatable behavioural characteristics (personality) and the ability to cope with social and environmental stressors are equally important factors.

1.2 THE ROLE OF ANIMAL PERSONALITY

Classically, studies of dominance have focussed on morphological traits especially those of body size and/or weaponry in determining the outcome of competition, and this is particularly so in dyadic studies of male-male competition (e.g. Beaugrand and Cotnoir, 1996; Réale and Festa-Bianchet, 2000; Réale et al., 2000; Prenter et al., 2008). These morphological traits have long been considered appropriate as suitable proxies for resource holding potential (RHP), i.e. an individual's absolute fighting ability (Parker, 1974). The necessity to win essential resources can thus explain the evolution of traits that enhance RHP even if these traits themselves are resource dependent (Kruuk et al., 2002). Recent competitive social experiences may not change morphological RHP but they may impact on physiological and psychological states, with notable differences in these effects on winners and losers (Price et al., 1994; Hsu and Wolf, 2001; Benson and Basolo, 2006; Briffa and Sneddon, 2007; Bernier et al., 2008; Hsu et al., 2009). Further, post-fight changes in physiological or psychological state may affect the ability of an individual to assess the RHP of an opponent (Arnott and Elwood, 2009b). Therefore, while classic morphological RHP traits such as size and weaponry are important for resource acquisition, other less obvious traits may be important too. Morphological traits may allow some individuals to outsize opponents; however, the role of animal *personality* is also important in the determination of social dominance (e.g. Dingemanse and de Goede, 2004; Earley, 2006; Ostner et al., 2008; Oliveira et al., 2011; Wilson et al., 2013). Thus agonistic behaviours such as aggression and boldness may also be viewed as part of an individual's RHP (Rudin and Briffa, 2012).

In behavioural ecology, growing use of the term *personality* reflects research parallels with human psychology (Budaev, 1997a; Moretz, 2003). In humans, personality is broadly used by psychologists to denote those characteristics that describe and account for consistent patterns of feeling, thinking and behaving (Pervin and John, 1997; Gosling, 2001). Animal personality is not so easily defined (Toms et al., 2010; Carter et al., 2013) and debate continues about if and how personality is distinct from closely related concepts such as temperament (Boissy, 1995; Réale et al., 2000), behavioural syndromes (Sih et al., 2004a; Sih et al., 2004b) and stress coping styles (Koolhaas et al., 1999). Nonetheless, a consensus definition of personality used throughout this thesis is among-individual differences in behaviour that are consistent across time and situation (Ariyomo et al., 2013a).

Animal personality has many important implications for ecological and evolutionary studies. For example, species dispersal may depend upon boldness and a willingness to explore in search of food or mates (Dingemanse et al., 2003; Cote et al., 2010; Sih et al., 2012; Brodin et al., 2013). Additionally, if personality differences evolve simultaneously with morphological changes (i.e. change together in a correlated manner) this may be important in allowing speciation to occur (Wcislo, 1989; Wilson, 1998; Dall et al., 2004). A core idea of this thesis is that personality is linked to an individual's ability to acquire resources and/or cope with stress under conditions of competition. As such personality is likely to be a major determinant of resource dependent life history traits and so of fitness itself.

In relation to competition and social dominance the most commonly studied personality traits are aggression and boldness (Conrad and Sih, 2009; Carter et al., 2013). An important point to recognise is that while aggression is not the same as social dominance, it is a behavioural strategy that is often used to assert dominance (Bernstein, 1976). As such, patterns of aggressive behavioural expression among individuals in a group are commonly a good predictor of dominance hierarchies (Francis, 1988; Jackson, 1991). While defining aggression is relatively straightforward (actual, threat or signal of attack, Hand, 1986; Francis, 1988), a universally agreed upon definition for boldness is less easy to find (Carter et al., 2012). However, boldness is loosely defined here as a willingness to take risks, e.g. approach a novel object or leave a refuge, especially in novel situations.

If defining personality traits is contentious then it is perhaps unsurprising that many different experimental designs have been used to measure them (Carter et al., 2012; Carter et al., 2013). However, the key to evaluating the consistency across time and/or context component of personality lies in the ability to obtain repeated measures of behavioural traits on individual animals. This is most readily done in captive animal populations, although in some cases appropriate data can be collected *in situ* on wild animals (e.g. Réale et al., 2000; Ferrari et al., 2013). Repeated measures allow statistical separation of observed variation into among-individual differences, i.e. how repeatable particular traits are for individuals within a population (potentially indicative of personality) and within-individual variation attributable to phenotypic plasticity and/or measurement error.

This basic strategy of obtaining data from repeated observations of individual behaviour has been used to investigate a range of personality traits including not only aggression and boldness, but also fearfulness, exploration, general activity and sociability (e.g. Dingemanse and de Goede, 2004; Svartberg et al., 2005; Dzieweczynski and Crovo, 2011). As may be expected, different behavioural trials have become standard for the assessment of different personality traits in different animal taxa. One widely used experimental test is the open field trial (OFT). Long used in rodent studies, the OFT is widely recognised as an appropriate method for testing boldness (Walsh and Cummins, 1976; Burns, 2008). The OFT comprises observation of an animal in a confined, empty space, over a specific length of time, often with a lead-in acclimation period. Simple modifications of the OFT often include placement of novel objects that individuals can choose to approach/investigate and/or a refuge that the animal can choose to emerge from (e.g. Cote et al., 2010). The introduction of a (simulated) predator to asses risk taking and behavioural response to acute stress is also common (Blanchard et al., 1998; Budaev, 1999; Budaev and Zworykin, 2002; Webster et al., 2007; Jones and Godin, 2010; Dammhahn and Almeling, 2012; Muller, 2012; Brodin et al., 2013).

It is currently unclear just how important personality traits such as aggression and boldness are for determining social dominance. Nonetheless, there is evidence that, at least in some cases, personality may be more important than classical RHP traits for determining competition outcome (e.g. aggressiveness was a better predictor of dominance in dyadic contests than size, Wilson et al., 2013). In addition, personality traits such as aggression and boldness are commonly found to (positively) correlate with each other and with fitness-related and life history traits such as reproductive success, growth, dispersal and response to predation risk in a wide range of taxa (e.g. Dingemanse et al., 2004; Brown et al., 2005a; Bell and Sih, 2007; Stamps, 2007; Biro and Stamps, 2008; Cote et al., 2010; Ariyomo and Watt, 2012; Rudin and Briffa, 2012; Mutzel et al., 2013). This reinforces the general point that resource acquisition under competition may depend on among-individual differences in behavioural characteristics (i.e. personality) as well as morphological traits. However, there is a need for more studies to investigate the behavioural mechanisms hypothesised to link resource acquisition and life history variation (e.g. Biro and Stamps, 2008; Dingemanse et al., 2012a).

1.3 COMPETITION AS A SOURCE OF STRESS

If personality may be a key determinant of social dominance and competition outcome, then the ability to cope with stress arising from competition may also be important in determining fitness. In biology, stress is another poorly defined - and much maligned - term. Selye (1973) described stress as "the nonspecific response of the body to any demand made upon it", while more recently, stress has been described as a mechanism of adaptation to threats on homeostasis involving a complex suite of responses to regain equilibrium (Chrousos, 1998). The latter definition emphasises the point that stress responses should primarily be seen as beneficial (or adaptive). By maintaining (or recovering) homeostasis, the behavioural and physiological changes that comprise the vertebrate stress response are critical for dealing with environmental challenges. However, it is also the case that if a stressor is prolonged, the physiological response mechanisms can become compromised and maladaptive, resulting in damage to the health of the individual (Barton and Iwama, 1991).

Stressors can arise from both biotic and abiotic features of the environment and may be described as acute (short term) or chronic (prolonged). Competition with conspecifics can induce stress directly through behavioural interactions (e.g. bullying or fighting), and indirectly through reducing availability of resources (e.g. food). Consequently, the ability to cope with stress may be crucial to becoming socially dominant or for maintaining or regaining social rank (e.g. Koyama, 1970; Lincoln, 1972; Cobb and Tamm, 1975; Dingemanse and de Goede, 2004) and managing the consequences of competitive outcomes (e.g. a loss of resource).

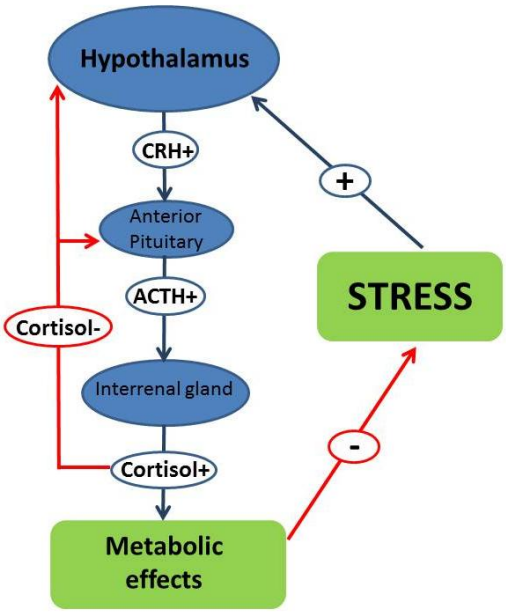
1.4 PHYSIOLOGICAL STRESS RESPONSE

While the vertebrate stress response includes both physiological and behavioural response mechanisms, it is the former that have been most extensively studied. The three stages of physiological stress response are well documented in the literature and this is especially true for fishes (Wedemeyer et al., 1990; Barton and Iwama, 1991; Wendelaar Bonga, 1997; Barton, 2002; Ashley, 2007; Pottinger, 2008). Briefly, the primary response to stressor exposure includes elevation of circulating levels of the corticosteroids catecholamine and cortisol that induce secondary changes in glucose and glycogen metabolism (Wendelaar Bonga, 1997). The secondary response comprises the diversion of metabolic resources from investment activities such as reproduction and growth toward the intensification of activities such as locomotion, respiration, tissue repair and hydromineral regulation, thus ensuring that homeostasis

is regained(Barton, 2002). These responses to acute stress are a normal and important part of daily life, controlled to a large degree by the hypothalamic-pituitary-adrenal axis (HPA) in vertebrates. Lacking an adrenal cortex, fish instead produce corticosteroids from the interrenal gland of the head kidney and thus control physiological stress response by the hypothalamic-pituitary interrenal axis (HPI axes) (Fig. 1.1).

However, if individuals are exposed to prolonged (chronic) stressors then damaging tertiary stress responses can occur. These result in changes to whole-animal performance that can include reductions in growth rate, loss of condition, reduced immune function and increased mortality risk (Wedemeyer et al., 1990; Balm, 1997; Fletcher, 1997; Wendelaar Bonga, 1997). The tertiary stress response is thought to be a result of prolonged elevated cortisol levels that do not return to normal due to exhaustion of the negative feedback mechanism (Fig.1.1) (Wedemeyer et al., 1990; Pickering, 1993). The central role of corticosteroids in mediating both acute and chronic stress responses in vertebrates has meant that circulating levels of cortisol (fish and mammals) and cortisone (birds and reptiles) are widely used as physiological measure of stress.

Fig. 1.1 The Hypothalamic-pituitary-interrenal axis of fish. Corticotropin-releasing–hormone (CRH) is produced by the hypothalamus in response to stress, stimulating release of adenocorticotropic hormone (ACTH) in the anterior pituitary. In turn, ACTH stimulates the secretion of cortisol by the interrenal gland. Homeostatic regulation is safeguarded by a negative feedback system (red arrows) acting upon the hypothalamus.



1.5 Stress Coping Style

Over-production of corticosteroid as the physiological response to stressors is sometimes seen as synonymous with stress. However, behavioural stress responses such as the fight or flight reactions to competitors and predators are also imperative to fitness and survival. An adaptive stress response is therefore likely to involve the integration of physiological and behavioural processes (Wingfield et al., 1998; Boonstra et al., 2001; Dufty et al., 2002) with expectations of correlation structure between them (Archard et al., 2012). As with any other aspect of the phenotype, it is also plausible that stress responses will vary among individuals within a population due to underlying differences in genetic factors and/or environmental conditions experienced. Indeed, among-individual variation in stress response traits has proven to be commonplace (Huntingford, 1976; Verbeek et al., 1996; DeVries, 2002), and this has led to the concept of the stress coping style (SCS) (Benus et al., 1991; Koolhaas et al., 1997; Koolhaas et al., 1999; Korte et al., 2005)

As originally proposed, the SCS model suggests that individuals within a population can be categorised as having one of two coping styles. *Proactive* individuals are those that actively challenge stressors and present behavioural profiles consistent with bold personalities (e.g. Brown et al., 2007; Thomson et al., 2011). Rapid development of rigid routines and the presence of low hypothalamic-pituitary-axis (HPA) activity are also features of a proactive style. In contrast, *reactive* individuals are shy and demonstrate low levels of aggression but have more flexible behavioural responses and tend toward raised HPA activity (e.g. Øverli et al., 2007; Carere et al., 2010).

Some links between SCS and social dominance have been found in a number of empirical studies. For instance, individuals with raised cortisol levels following socially stressful encounters (and thus likely to be defined as reactive) may become subordinate (Fox et al., 1997). It is also true that social and reproductive status along with stability of social situation can sometimes predict circulating cortisol levels although correlations are not always straightforward. For instance, increased circulating plasma corticosteroid levels have been reported in both socially dominant and subordinate animals following aggressive encounters; however those of subordinate individuals tend to be highest and of longer duration (Bronson, 1973; Sloman et al., 2001; Summers, 2002; Abbott et al., 2003; Carere et al., 2003; Earley et al., 2006). Although typically presented as dichotomous, proactive and reactive coping

styles may actually represent opposite ends of a continuously varying axis. If the SCS model is valid, then stress response traits should not only be repeatable, but physiological and behavioural responses should also change in an integrated manner along a major axis of among-individual variation (Wechsler, 1995). SCS have been likened to personality, temperament and behavioural syndromes, and in some circles, such descriptions of consistent multivariate among-individual behaviour are interchangeable (Øverli et al., 2007).

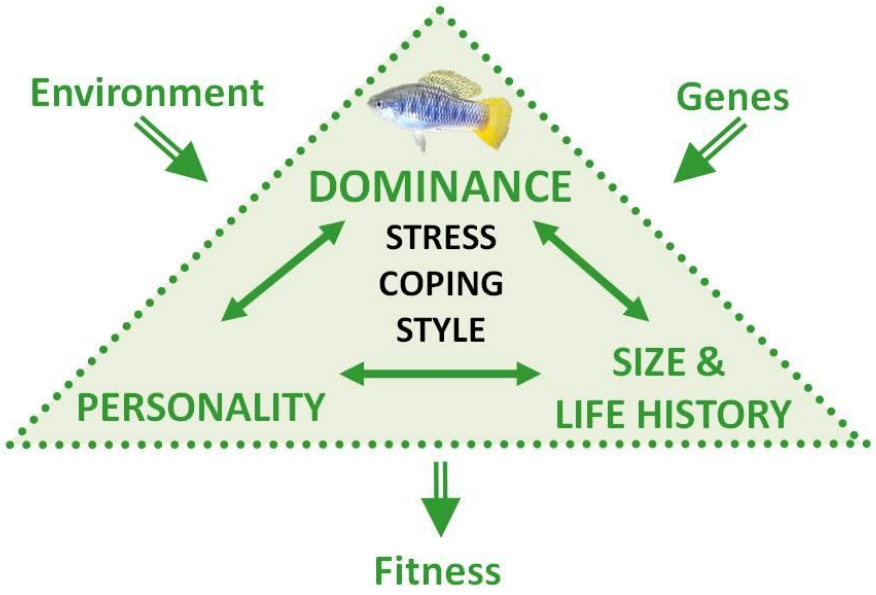
1.6 THE *DYNAMIC OF DOMINANCE*

When animals compete for resources, social dominance status (defined here as an individual's ability to acquire resources in competition) is expected to determine individual fitness through effects on resource-dependent life history traits. However, the traits that determine social dominance remain unclear. Classical studies emphasised the importance of size and other morphological measures of RHP (above) (Parker 1974, Dugatkin and Ohlsen 1990) and these are certainly important in contest outcomes. Being or becoming socially dominant may not be quite as straightforward as simply having high resource holding potential (RHP), however. Other factors such as personality are likely to be important (Cobb and Tamm, 1975; Hinde and Datta, 1981; Francis, 1988; Fox et al., 1997; Creel, 2001; Carlson et al., 2004; Ostner et al., 2008; Taves et al., 2009; Dahlbom et al., 2011), and causal relationships will sometimes be circular. This may be especially true where dominance has been measured based on the pattern of resource access (e.g. Appleby 1980, Wilson et al. 2013). For instance, if rapid growth is dependent on acquiring resources then large size will be a consequence and not just a cause of dominance (Wilson e al. 2013). Furthermore, the relationships between personality, life history and social dominance are expected to be mediated by responses to stress caused by competitive interactions within the social environment.

Throughout this thesis, the complex association of traits both causal and consequent to social dominance is referred to as the *dynamic of dominance* (Fig 1.2). A core theme of the following chapters is that to understand each component of the *dynamic of dominance* it is necessary to take a multivariate approach and to understand how variation in, and covariation between traits is distributed at the among-individual level. To enable this, statistical methods more widely used in quantitative genetics are adopted, applying them here to model repeated measures of data on behavioural, endocrine and life history data. This approach follows the recommendations of others

who have highlighted the great potential of linear mixed effect models for testing hypotheses about personality (Dingemanse et al., 2010; Dochtermann and Roff, 2010; Dingemanse et al., 2012a; Dingemanse and Dochtermann, 2013; Araya-Ajoy and Dingemanse, 2014).

Fig 1.2 The *dynamic of dominance* highlights the expectation of complex relationships between social dominance, stress coping style, personality, life history, genes and the environment, all ultimately affecting individual fitness.



1.7 STUDY SPECIES

This thesis is based on empirical studies of swordtail fishes, *Xiphophorus sp*. (Family: Poeciliidae, Order: Cyprinodontiformes). Chapters 2, 4 and 5 are based on studies of the sheepshead swordtail, *X. birchmanni* (Fig. 1.3a) using data collected by the author. Chapter 3 describes an analysis of data from a related species, the green swordtail *X. helleri* (Fig1.3b), collected by collaborators (RL Earley, B Sinderman and RM Pearce). Swordtails are small, sexually dimorphic live bearing tropical freshwater fish originating from different sites in Central America. In general males are highly ornamented, having long sword-like extensions to the caudal fin (although this is not present in *X. birchmanni*) making them a popular group among aquaria hobbyists. In the behavioural sciences, members of the *Xiphophorus* genus (especially the green swordtail *X. helleri*) have been extensively used as models in studies on sexual selection

(Basolo, 1988; Ryan and Keddyhector, 1992; Rosenthal et al., 1996; Rosenthal and Evans, 1998; Wong and Rosenthal, 2006), and on male-male aggression and dominance (e.g. Beaugrand and Zayan, 1985; Franck et al., 1985; Ribowski and Franck, 1993; Earley, 2006).

Fig. 1.3 Study species: a) Sheepshead swordtail (*Xiphophorus birchmanni*), male (above) and female (below) photographs from the Wilson Lab, University of Edinburgh, used in Chapters 2, 4 and 5; b) Green swordtail (*Xiphophorus helleri*) male, photograph from the Earley Lab, University of Alabama, used in Chapter 3.

a)

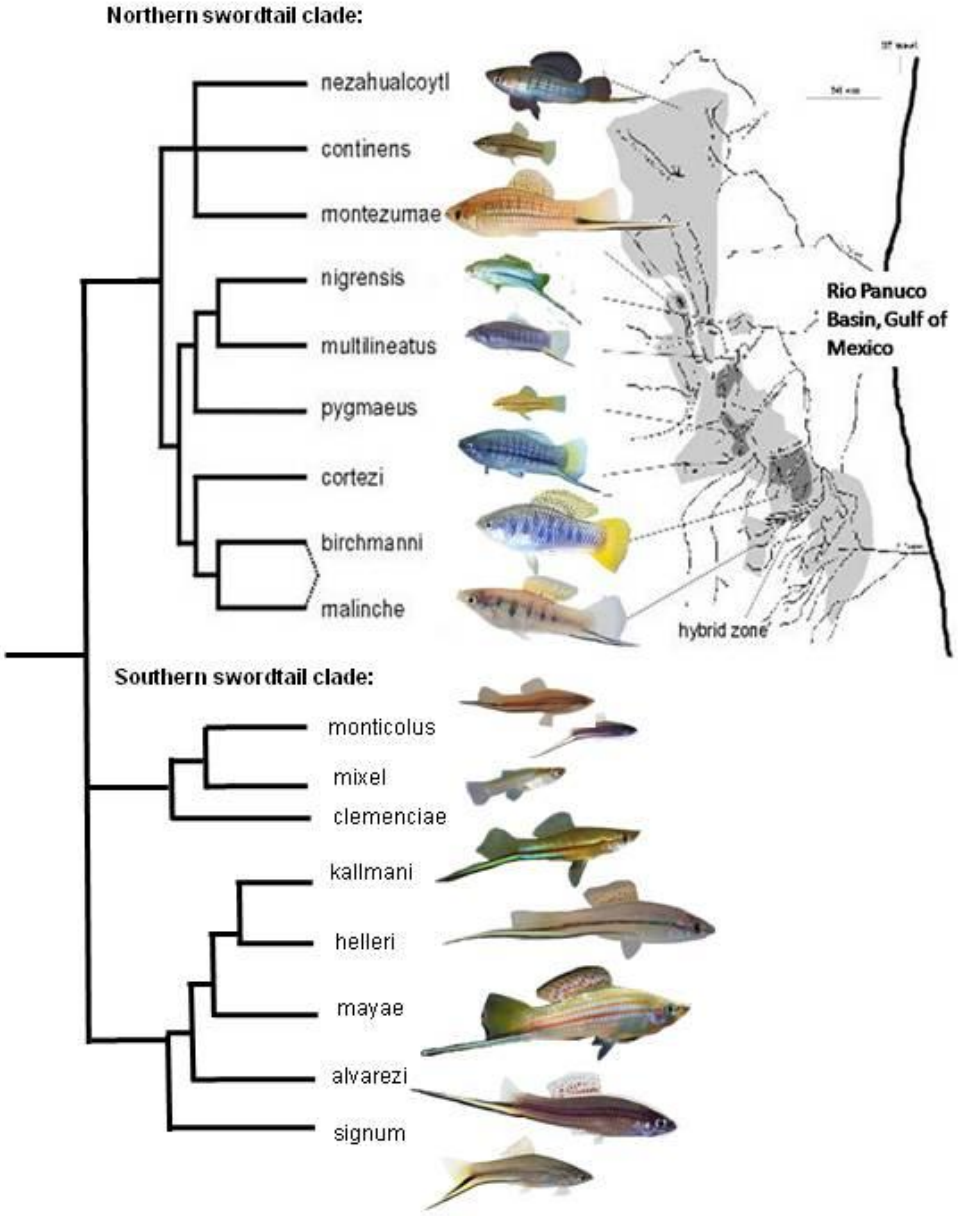
b)

1.7.1 PHYLOGENY

Native to areas of north and central America, extensive research has been focussed in diverse areas around three presumed clades of *Xiphophorus* - northern swordtails, southern swordtails, and platyfish (Meffe and Snelson, 1989). *X. birchmanni* is one of nine species belonging to the northern swordtail clade found in the Rio Panuco basin, Hidalgo, Mexico, and differs from other species in that males do not bear the classic sword-like caudal fin extension (Rauchenberger et al., 1990) (Fig. 1.3a, 1.4). The common name, sheepshead swordtail, derives from the presence of a nuchal (neck) hump in mature males, suggestive of a sheep skull in profile. *X. helleri* is one of four species belonging to the wider ranging southern swordtail clade, native to an area from Veracruz, Mexico, to the north-western Honduras (Fig 1.3b, 1.4). Although found in diverse colour forms in the ornamental fish trade, wild *X. helleri* is an olive-green colour, hence the common name, green swordtail.

Based on shared morphological features and some early genetic data *X. birchmanni, X malinche* and *X. cortezi,* were once considered to belong to a *cortezi* species complex within the northern clade, and potentially to be differing morphotypes of a single species (Rauchenberger et al., 1990). However, further genetic data coupled with the recognition that *X. birchmanni* x *X. malinche* hybrids had likely been among the original specimens examined in the 1990 study led to this idea being dismissed (Morris et al., 2001; Rosenthal et al., 2003). A more recent phylogenetic study of mitochondrial DNA from all three species confirmed the monophyly of *X. birchamnni* sampled (Gutiérrez-Rodríguez et al., 2008). On the other hand there has been very little doubt that *X.helleri* belonged to the southern swordtail clade (Rosen, 1979), subsequently confirmed by molecular techniques (Meyer et al., 1994; Borowsky et al., 1995; Hrbek et al., 2007).

Fig. 1.4 The Northern swordtail clade (top) with geographical distribution and the more widely distributed Southern swordtail clade (below, adapted from Rosenthal 2011)



1.7.2 REPRODUCTIVE BIOLOGY

Swordtails, as with most members of the Poeciliid family, are set apart from the majority of other fish species by a collection of interesting reproductive adaptations. Primary sexual characteristics become evident at around sixteen to twenty weeks, when males can be distinguished from females by the fusing of the nine rays of the anal fin to form a gonopodium (Fig. 1.5). This intromittent organ, controlled by a complex set of bones and muscles and adorned with asymmetrical species-specific hooks and claws is used to impregnate females (Rosen and Gordon, 1953).

Upon sexual maturity, male *X. birchmanni* and other sword-less species develop secondary characteristics including the replacement of the lateral line by vertical bars that become darker when exhibiting courting or threatening behaviour (Morris et al., 1995). Additionally the dorsal fin (and in *X. birchmanni,* the nuchal hump) becomes pronounced. Sexual maturity in male *X. helleri* and other sword bearing species triggers the extension of the brightly coloured long ornamental sword. Spermatogenesis and spermiation are dependent upon long- and short-term environmental variations respectively, for example temperature, day-length and mature female availability (Constantz, 1989). Females retain the dark ventral line, and often develop a gravid spot at sexual maturity, although this is not as apparent in *X. birchmanni* as other Poeciliid species (personal observations).

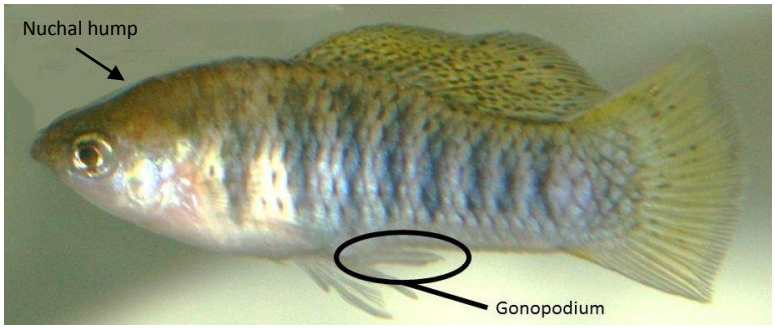


Fig. 1.5 Adult male *X. birchmanni*. The gonopodium, used to inseminate females is circled. Also visible are the (fading) dark vertical bars and lowering dorsal fin, indicating signs of a recent aggressive encounter. The nuchal hump, giving rise to the common name, sheepshead swordtail is also evident.

Copulation and insemination occur following the release from the female's urogenitary aperture of chemicals (probably oestrogen) that stimulate male sexual behaviour (Amouriq, 1964; 1967) and are probably perceived by taste (Parzefall, 1973) This is most likely to occur in a brief period before and after parturition, the time of maximum male interest (Parzefall, 1973). It is thought that the oestrogen acts as a pheromone to incite scramble competition among males and thus intensify sexual selection (Constanz, 1984). The entire female reproductive cycle lasts approximately 30 days (on average) under controlled conditions and, dependent on exact species, age, size and condition, females give birth to broods of varying sizes, with two-day parturition to fertilization intervals (Turner, 1937; Rosenthal, 1952; Constanz, 1984). Females are able to retain sperm in the folds of both ovary and gonoduct linings where it may persist for up to eight months, or eight broods, being nourished by female secretions, (Turner, 1937; Winge, 1937; 1989). Females can therefore give birth to consecutive broods from a single fertilization although they are non-superfetative (i.e. only one brood develops at any one time).

Caution is required in reaching conclusions regarding the exact detail of maternal provisioning in Poeciliids. Throughout the family, the placenta is thought to have evolved independently several times, either as a result of antagonistic co-evolution or as a means of adaptation to environmental pressures (Rosen and Bailey, 1963; Hrbek et al., 2007; Pollux et al., 2009). However, as no placenta exists in swordtails they are generally deemed to be lecithotrophic, with developing embryos nourished by egg-provisioning only. A degree of maternal provisioning (partial matrotrophy) may occur in some swordtail species, leading some authors to conclude that such species might best be classified as unspecialised matrotrophes (Scrimshaw, 1945; Depeche, 1976; Haas-Andela, 1976; Wourms, 1981), whereas more recent studies claim *Xiphophorus sp*. females to be viviparous, lecithotrophic and non-superfetating (Thibault and Schultz, 1978; Pollux et al., 2009).

1.8 THESIS OVERVIEW

In broad terms, the goal of this thesis is to explore some of the intricacies of the *dynamic of dominance*, focussing on the relationships between personality, stress response, morphology and life history. Chapters 2-5 detail a series of studies designed to address more specific hypotheses, but with this broad goal very much in mind.

In Chapter 2, personality variation in a captive population of swordtails is quantified. In particular the question of individual personality stability over long time periods is examined. Although there are some exceptions, the majority of studies finding evidence for repeatable (i.e. among-individual) differences in behaviour have used behavioural observations collected only over short time periods relative to expected lifetimes of the study organisms. Thus the stability of patterns of population-level behavioural variance and individual behavioural rank across longer timescales or multiple sampling periods is unclear. Since natural selection occurs through variation in lifetime fitness, the stability of personality over individual lifetimes is important. If personalities are not stable over long periods, then the possibility that they are generally under selection is greatly diminished (Smith and Blumstein, 2008) . Observations collected across two discreet time periods (long and short) are used to seek answers to these questions. By distinguishing between directly observed behavioural traits and underlying axes of personality (inferred from among-individual covariance between observed behaviours), the extent that one (or more) personality traits can adequately explain observed behavioural variation is assessed. The possibility that repeatability estimates from short time periods give an upwardly biased view of the importance of individual personality over longer periods is examined, and the long term stability of the axes of variation defining personality traits within a population is tested.

In Chapter 3 the link between stress, behaviour and competitive outcome is investigated using data from dyadic contests between male *X. helleri*. Previous work on dyadic contests, perhaps the simplest form of social competition, has emphasised the importance of prior experience on winner loser effects (Beaugrand and Goulet, 2000; Earley and Dugatkin, 2002; Earley et al., 2003; Smith and Blumstein, 2008) and resource holding potential (Moretz, 2003; Arnott and Elwood, 2008; Arnott and Elwood, 2009a; Arnott and Elwood, 2010). However, other factors are likely to be important, including the way that individuals cope with stress. Furthermore, in experimental studies where contests are staged, contest outcome may actually be dependent on an individual's response to acute stress caused by experimental protocol itself. In Chapter 3 this possibility is explored, with the specific hypothesis that a key determinant of contest outcome may be latency to recover behaviourally and physiologically from the stress of experimental protocol. If variance in stress response, or stress coping style (SCS), is an important determinant of observed contest behaviour

and/or outcome, then relationships should exist between these variables and both behavioural reaction to disturbance (prior to meeting an opponent) and physiological stress response as measured by cortisol levels.

In Chapter 4 the theme of exploring the correlation between behavioural and physiological stress responses is continued, extending the approach to include repeated measures of behaviour and endocrine state. Essential to properly evaluate the proactive-reactive model of SCS, this has seldom been attempted, with most studies relying on single measures of either physiological or behavioural responses or both. If the SCS model is valid, then not only should physiological and behavioural stress responses be repeatable, but (among-individual) correlation structure is expected between them. Physiological and behavioural response traits should also change in an integrated manner along a single major axis of among-individual variation.

Chapter 5 returns to the core theme of exploring the full *dynamic of dominance* using a quantitative genetic approach. A pedigreed population of *X. birchmanni* was raised under four experimental density treatments, designed to impose differing levels of social stress from competition for space. Fish were observed over a one year period, collecting data on morphology (growth), life history (male maturation, longevity (a proxy for fitness)) and personality (boldness, social dominance in males). Repeatability is first determined for those traits with repeated measures (growth, personality) then quantitative genetic models are used to fully partition the multivariate phenotypic (co)variance into genetic and environmental components. In this way not only the among-individual correlation structure of traits causal and consequent to competitive ability is scrutinised, but also the genetic relationships that underpin social dominance and ultimately set the potential for evolutionary responses to selection are characterised.

Finally, in Chapter 6, the main findings of this work are summarised and some concluding thoughts and reflections on the studies are presented. Recommendations for similar studies and avenues for further study are suggested.

CHAPTER 2

HOW STABLE ARE PERSONALITIES? A MULTIVARIATE VIEW OF BEHAVIOURAL VARIATION OVER LONG AND SHORT TIMESCALES IN THE SHEEPSHEAD SWORDTAIL, XIPHOPHORUS BIRCHMANNI

This chapter is published as: **Boulton K,** Grimmer AJ, Rosenthal GG, Walling CA and Wilson AJ (2014). How stable are personalities? A multivariate view of behavioural variation over long and short timescales in the sheepshead swordtail, *Xiphophorus birchmanni*. Behavioural Ecology and Sociobiology, 68:791-803.

2.1 ABSTRACT

Many studies have revealed repeatable (among-individual) variance in behavioural traits consistent with variation in animal personality; however, these studies are often conducted using data collected over single sampling periods, most commonly with short time intervals between observations. Consequently, it is not clear whether population-level patterns of behavioural variation are stable across longer timescales and/or multiple sampling periods, or whether individuals maintain consistent ranking of behaviours (and/or personality) over their lifetimes. Here we address these questions in a captive bred population of a tropical freshwater Poeciliid fish, *Xiphophorus birchmanni*. Using a multivariate approach, we estimate the among-individual variance-covariance matrix (**I**), for a set of behavioural traits repeatedly assayed in two different experimental contexts (open field trials, emergence and exploration trials) over long- (56 days between observations) and short-term (four day observation interval) time periods. In both long- and short-term data sets we find that traits are repeatable and the correlation structure of **I** is consistent with a latent axis of variation in boldness. While there are some qualitative differences in the way individual traits contribute to boldness, and a tendency towards higher repeatabilities in the short term study, overall we find population-level patterns of among-individual behavioural (co)variance to be broadly similar over both time frames. At the individual level we find evidence that short-term studies can be informative for an individual's behavioural phenotype over longer (e.g. lifetime) periods. However statistical support is somewhat mixed and, at least for some observed behaviours, relative rankings of individual performance change significantly between data sets.

2.2 INTRODUCTION

It is now apparent that, within animal populations, individuals often exhibit differences in behaviour that are repeatable across time and context. This repeatable variation is taken as evidence for animal temperament (e.g. Boissy, 1995; Réale et al., 2007), behavioural syndromes (Sih et al., 2004a), coping styles (Koolhaas et al., 1999), or personality, the latter term reflecting parallels with research in human psychology (Budaev, 1997a; Gosling, 2001). A number of axes of among-individual behavioural variation condensed into personality traits have been described, including boldness-shyness, exploration-avoidance and general activity (Réale et al., 2007). Understanding the evolution of personality has become a major field of study in behavioural ecology (Dall et al., 2004; Stamps and Groothuis, 2010). There is now growing evidence that traits relating to personality contribute to fitness variation and therefore may be both adaptive and generally under selection (Smith and Blumstein, 2008). However, if natural selection occurs through variation in lifetime fitness, then an important question arises: just how stable are personalities over individual lifetimes? Here we address this question in a captive population of fish. We do this using a novel multivariate approach that characterises personality variation as a latent character underpinning among-individual (co)variation in a suite of observed behaviours.

2.2.1 DEFINING PERSONALITY

While there remains considerable disagreement over how best to define individual personality traits (Réale et al., 2007; Toms et al., 2010; Carter et al., 2013; see below) there is broad consensus that among-individual behavioural variance is the statistical signature of animal personality. Typically this is quantified as the (among-individual) repeatability, defined as the proportion of observed variance explained by individual identity, of one or more observed behavioural traits. Thus partitioning of observed variance into among- and within-individual components (the latter arising from individual plasticity and/or measurement errors) from repeat observations on individuals is crucial to empirical studies of personality (Dingemanse et al., 2012a; Brommer, 2013; Araya-Ajoy and Dingemanse, 2014). In a meta-analysis, Bell et al (2009) concluded that on average, estimates of repeatability for observed behavioural traits decreased as the interval between sampling events increased. Consequently, it may be dangerous to assume that short-term studies reflect behavioural (and by implication, personality) differences that are stable over the lifetime of individuals. This is potentially important since short-term repeatability estimates predominate in

the literature, although the number of studies conducted over timeframes that may be considered more representative of natural life-spans is growing (for more recent examples, see Ronning et al., 2005; Bushuev et al., 2010; Chervet et al., 2011; David et al., 2012; Kanda et al., 2012). However, few studies have collected repeated observations over two distinct time periods from the same individual (but see for e.g. Carere et al., 2005) that would allow the repeatability of repeatability to be assessed. Here we do this, but also extend our analysis to the multivariate case to ask whether patterns of among-individual behavioural (co)variation reflect an underlying personality trait that is stable across distinct long- and short-term sampling periods.

In what follows we investigate the temporal stability of multiple behavioural traits in the freshwater Poeciliid fish, *Xiphophorus birchmanni* to answer two complementary questions. Firstly, at the level of the population, how stable are the patterns of among-individual trait (co)variance generated by underlying personality? Secondly, at the level of the individual, do short term studies reveal behavioural tendencies that are stable across lifetimes? To answer these questions we characterise behavioural variation along what we loosely consider to be an axis of shyness-boldness. Boldness is the most commonly studied axis of personality in fish (Toms et al., 2010), and positively correlates with fitness-related traits including reproductive success, parental provisioning, growth, aggression, social dominance, dispersal and proactive responses to stressors such as predation risk (Dingemanse et al., 2004; Brown et al., 2005a; Bell and Sih, 2007; Cote et al., 2010; Ariyomo and Watt, 2012; Rudin and Briffa, 2012; Mutzel et al., 2013). There remains, however, a lack of consensus on how best to define boldness and how it should be assayed (Toms et al., 2010). This raises obvious potential for misclassification of personality traits (Carter et al., 2013), and/or disagreement over appropriate experimental design (Toms et al., 2010).

The present goal is to investigate stability of a personality trait without adding further to existing debate over issues of definition. Consequently we do not attempt to define boldness or the best way to measure it *a priori*; rather, we follow the view of others that personality traits should be considered as latent variables that can best be uncovered by observing several measurable, correlated and potentially overlapping behaviours across contexts (Dochtermann and Jenkins, 2007; Dingemanse et al., 2010; Dochtermann and Roff, 2010). We therefore make a distinction throughout between behavioural traits that are observed directly, and personality (traits), inferred from among-individual (co)variance in observed behaviour(s). This exploratory approach

that follows Huntingford (1976) and others (Budaev, 1997a; Moretz, 2003) is becoming more mainstream and allows the avoidance of difficulties that can arise if a single behaviour is chosen *a priori* to assay boldness. For example, a fish that swims a long distance in a behavioural trial may be classified as willing to explore and therefore as bold; however, this behaviour could also plausibly be indicative of anxiety, with the animal's exploration being driven by a search for refuge.

Currently the most common experimental paradigm used to measure boldness is that of the open field trial (OFT), where an animal is placed in an open arena and its behaviour is monitored for a predetermined observation period. Initially developed for rodent studies (Hall, 1934; Walsh and Cummins, 1976; Moretz, 2003), OFTs have long been applied to fish models (Warren and Callaghan, 1975; Budaev, 1997a). Considered the most reliable way to assay boldness by some authors (Burns, 2008), others have argued that OFTs risk conflating boldness with other axes of variation that are distinct (if sometimes correlated) personality traits in their own right (e.g. exploration-avoidance, overall activity, Réale et al., 2007). If so, then simple modifications to OFTs such as providing a refuge that an animal can choose to emerge from and explore (emergence and exploration trial, EET) may be useful (Dingemanse et al., 2007).

2.2.2 HYPOTHESES

In what follows we use both types of behavioural trial mentioned above (OFT and EET) to observe how fish behave in these contexts and to characterise the repeatable component of multivariate behaviour. We then assess the extent that one or more major axes of variance adequately depict observed variation. In other words, we aim to describe the behavioural trait variation first, and then consider the extent that its repeatable component fits within the paradigm of a major axis of personality, i.e. the boldness-shyness axis (Dingemanse et al., 2010; Dochtermann and Roff, 2010). We then go on to address three specific questions regarding the temporal stability of personality. Firstly we ask whether repeatabilities estimated from repeated measures of individual behaviours over a short time period give a misleading view of the importance of among-individual variance over longer time periods. Secondly, by extending our analysis to the multivariate case we ask whether the structure of the between-trait among-individual covariance matrix, denoted **I**, following Wilson et al. (2013), is similar when estimated from short- and long-term data; i.e. do repeated empirical analyses of a single population actually reveal the same major axes of among-individual variation? If so, then a final question concerns the extent that individuals

retain the same relative ranking for repeatable behaviours, and hence personality, over their lifetimes.

2.3 METHODS

2.3.1 STUDY SPECIES AND HUSBANDRY

One hundred wild adult *Xiphophorus birchmanni* were caught in the Arroyo Coacuilco near the town of Coacuilco, municipality of San Felipe Orizatlán, Hidalgo, Mexico, (elevation 314 m lat/long 21.099 -98.587), and imported to the UK in February 2010. Between August 2010 and May 2011 we collected an offspring generation (n = 384) from 13 males and 27 females (mean (SE) brood size of 8.86 (0.541)). Gravid females were isolated and, following birth, broods were immediately netted and moved to one half of a partitioned 30 L tank; broods of more than six offspring were split with each half of the family placed in different tanks. Fry were fed twice daily on a mix comprising equal quantities of crushed ZM spirulina and brine shrimp flake and laboratory prepared brine shrimp nauplii. At an average of 17 weeks (range 12 to 27) juveniles were tagged with a single elastomer injection for individual identification purposes and transferred to mixed-family rearing groups of n = 8. Note it is not possible to determine sex at this age in this species and therefore the sex ratio was not controlled. Eight rearing groups were then kept within each of six sequentially set-up stacks of tanks, each stack sharing a common water supply and recirculating filtration system. As part of a parallel study of density effects on growth, rearing groups were initially housed under two different density regimes as follows. Within each stack, four groups were placed in 30 litre (37 x 37 x 22 cm) glass tanks (low density treatment) with the remaining four groups in 15 L half tanks (high density treatment). Half tanks were created by placing a black net covered Perspex-framed partition down the centre of a full – size tank. Thus, establishing a stack required 64 fish (i.e. 8 x 8) to be available for tagging simultaneously and this accounts for the variation in tagging age within stacks. Fish were fed twice daily with a standardised ration of flake food as above (morning) and a mix of previously frozen blood worm and daphnia (afternoon). On the days when behavioural data was to be collected, the morning feed was omitted in an attempt to encourage exploration tendencies. Temperature was maintained between 22 - 24^{o}C and a 12:12 hr light:dark cycle imposed. After being housed in this manner for 28 weeks, density was swapped for half of the tanks, thus creating four treatment effects with the total number of fish divided approximately equally between them as follows: Low/Low (n = 93), Low/High (n = 95), High/High (n = 87), High/Low (n = 93).

Observations from individuals failing to reach sexual maturity by the end of the long-term study (50 weeks), were excluded from the analysis and the above breakdown (n = 11).

2.3.2 BEHAVIOURAL DATA COLLECTION

Trials were of two types, open field (OFT) and emergence and exploration (EET) with multiple specific behavioural traits assayed in each trial type (Table 2.1). The trials were performed over two experimental study periods, denoted long-term (LT) and short-term (ST). All available fish contribute to the long-term data set (n = 373) while a random subset of 32 fish from each of the four density treatments (Low/Low n = 13, Low/High n = 4, High/High n = 9, High/Low n = 6) was used for the short-term study (Table 2.1). Overall, the long-term trials took 13 months to complete (May 2011 – May 2012), with data collected over an actual 32 week period for each fish (see Appendix 3). Each individual was subject to an OFT followed by an EET seven days later, a process that was repeated three times at 56 day intervals, thus yielding four OFT and four EET trials per fish. The short-term data set was collected in February 2013, with 32 individual fish subjected to alternating OFT and EET at 48 hour intervals (i.e. 2 days between trials, 4 days between repeated trials of the same type) with each animal undergoing five trials of each type. For those 32 individuals used in both study periods data was therefore collected over a timeframe with a mean (SE) of 531.4 (6.38) days. By comparison the mean (SE) longevity of individuals with known birth and death dates under our laboratory conditions is 450.3 (8.10) days.

Table 2.1 Data set for long-term (LT) and short-term (ST) studies. Number (N) and sex of individuals involved: male (M), female (F), total (T). Periods of data collection and intervals between trial pairs. Number of trials conducted: OFT (Open Field Trial); EET (Emergence & Exploration Trial); N_{LT} = 2448, N_{ST} = 320. Mean age of fish in days at the start of each trial pair with standard error in parentheses; "-" indicates trial not performed

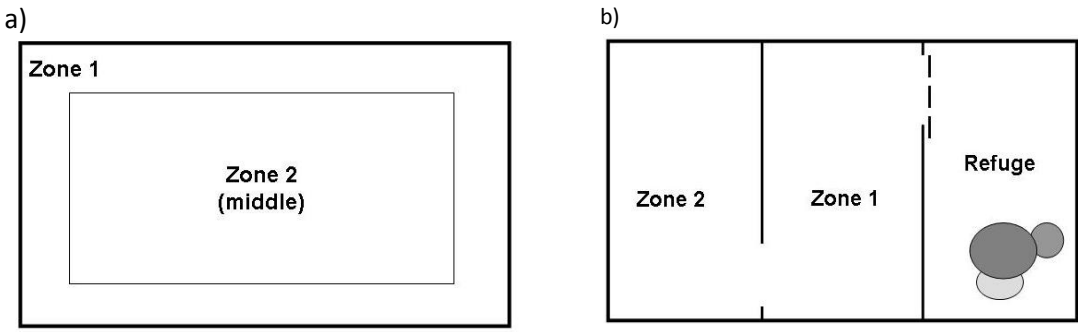
Study	N			Data collection period			Number of trials		Mean Fish Age (SE)				
	M	F	T	Start	End	Days between trials	OFT	EET	1	2	3	4	5
LT	223	150	373	May 2011	May 2012	56	1224	1224	203 (26.4)	259 (26.4)	372 (27.2)	427 (27.1)	-
ST	16	16	32	Feb 2013	Feb 2013	4	160	160	715 (13.4)	719 (13.4)	723 (13.4)	727 (13.4)	732 (13.4)

2.3.3 EXPERIMENTAL PROCEDURES

2.3.3.1 Open Field Trial (OFT)

An empty 45 x 25 x 25 cm tank was filled to a depth of 8 cm with room temperature water (22^{0}C). The tank was lit from below and visually screened by a cardboard casing to occlude external laboratory visual disturbance. Fish were caught individually from their home tank with a dip net, quickly examined for identification tags and immediately placed into the centre of the OFT tank. Following a 30 second acclimation period, behaviour was filmed for 300 seconds using a Sunkwang C160 video camera fitted with a 5 – 50 mm manual focus lens suspended above the tank. Data were then extracted from the video using the tracking software Viewer II (http://www.biobserve.com/products/viewer/index.html) that was set up to divide the tank basal area into two approximately equal halves (middle and perimeter zones) (Figure 2.1a). Water was changed between individual trials to prevent chemical cues affecting behaviour.

Figure 2.1 Set up of experimental tanks for a) open field trials (OFT) and b) Emergence and exploration trials (EET) as viewed from above. Both tanks measured 45 x 25 x 25 cm and were filled to a depth of 8 cm. For OFT two zones of equal area were defined for analysis. For EET the tank was divided into three equal zones with fixed opaque material. The refuge area contained a plastic plant and several small stones. A removable doorway (hatched line) provided a means of access from the refuge to the rest of the tank.



2.3.3.2 Emergence and Exploration Trial (EET))

A 45 x 25 x25 cm tank was physically divided into three sections with opaque Perspex, providing a right-hand, centre and left-hand chamber. A small (5 cm) opening was cut in each divider, starting two cm from the tank edge. The openings were positioned at opposite sides of the tank. The chamber on the right hand side was designated as the refuge, and equipped with a plastic plant and several small stones. A rising trapdoor was rigged to a pulley above the tank and positioned inside the refuge and covering the

exit into zone 1 (Figure 2.1b). Tanks were filled, emptied, lit and screened as above. Fish were individually caught and examined as before, and placed directly into the centre of the refuge where they were allowed 30 seconds to acclimate before the trapdoor was lifted. Filming then commenced for 300 seconds (as above), but only behaviour outside the refuge (i.e. in zones 1 and 2) was tracked and extracted for analysis.

2.3.4 BEHAVIOURAL TRAITS

The behavioural traits recorded in this study were selected as those likely to reflect variation along a bold-shy type personality axis (Table 2.2). For the OFT, we predicted that fish tending toward boldness would actively explore the novel environment of the OFT by leaving the tank sides and spending more time in the central zone than shy fish. OFT behaviour was therefore quantified by four traits; Track Length (TL), Activity (Act), Area Covered (AC) and Time in Middle of the tank (TIM) that we predicted would be positively correlated with one another. In the EET, we expected bold fish to locate the doorway in the refuge and leave through it. We recorded two traits from the EET: whether or not the individual emerged from the relative safety of the refuge (Emergence) and Latency in seconds to do so. We predicted positive within-individual correlations between Emergence from the refuge and the OFT traits, with negative correlations between Latency to Emerge and all other traits. Note that the EET tank was set up with the area outside the refuge further divided into two zones (1 and 2; Figure 2.1b). In the EET, we had initially planned to use latency to enter zone 2 (distal to the refuge) as an additional trait in our analyses; however, in practice this became a redundant trait due to a low frequency of fish entering this area.

Table 2.2 Behavioural traits recorded in OFT (Open Field Trials) and EET (Emergence and Exploration Trials).

Trial type	Measured trait		Definition
OFT	Track Length	(TL)	Distance swum (cm)
OFT	Activity	(Act)	Percentage of time moving at a minimum 1.5cm/sec (%)
OFT	Area Covered	(AC)	Area of tank floor covered (%)
OFT	Time in Middle	(TIM)	Time spent in Zone 2 (seconds, see Figure 2.1)
EET	Emergence	(Em)	Whether or not the fish emerged from the refuge (binary)

2.3.5 STATISTICAL ANALYSES

All data were modelled using restricted maximum likelihood mixed effects models implemented in ASReml V3 (Gilmour et al., 2009). Prior to analysis, data for the OFT trait Time in Middle were square root transformed to reduce positive skew. Visual inspection of residuals suggested that the assumption of residual normality was reasonable for the other traits recorded in OFT. All traits were rescaled to standard deviation units prior to analysis to prevent trait scale effects from influencing the structure of **I** (defined and estimated as described below). Given that a large proportion of fish did not emerge from the refuge (see results) the Latency to Emerge data were heavily censored and we elected to use only the binary variable of Emergence in subsequent analyses. Emergence was included in full multivariate models using REML under an assumption of (multivariate) residual normality. Statistical inferences on this trait should therefore be treated with obvious caution.

2.3.5.1 Analysis of binary data

While statistical approaches exist that allow non-Gaussian trait distributions to be used (e.g. MCMC Bayseian approaches implemented in theR package MCMCglmm; Hadfield, 2010b) they do not currently allow the error structures appropriate to our multivariate models (i.e. no definable or estimable residual covariance between OFT and EET traits – see below) and thus could not be used here. However, we checked the validity of REML-based conclusions regarding Emergence by fitting additional univariate and bivariate models using MCMCglmm. Specifically we fitted a univariate model of Emergence to estimate the repeatability of this trait and bivariate models of Emergence with all other OFT traits to estimate the covariance structure between these traits. All models in MCMCglmm modelled Emergence as a categorical trait with the residual variance fixed at 1 and all OFT traits as Gaussian. All MCMCglmm models were run for a total of 1050000 iterations with a burn-in of 50000 iterations and a thinning interval of 1000 iterations. The repeatability of Emergence from MCMCglmm models was defined as the intraclass correlation, calculated as $V_I / (V_I + V_R + \pi^{2/3})$, where V_I is the among-individual variance and V_R is the residual variance that in this case is fixed to 1 (Hadfield, 2010b).

2.3.5.2 Fixed effects

To test the hypothesis that among-individual variance for behavioural traits is both present and repeatable in our fish species, we first combined data from both collection periods and fitted a multivariate model of our observed behavioural traits. For each

trait we included fixed effects of the *mean*, *sex* (a two level factor determined from external morphology at maturation), home *stack* (a six level factor accounting for differences between sets of fish sharing water supplies), *trial number*, density *treatment*, and *day order*. *Trial number* is the cumulative number of trials experienced by an individual (fitted as a linear effect). *Density treatment* is a four level factor describing density conditions experienced in the rearing stacks. *Day order* was modelled as a linear effect of the number of preceding trials conducted on any day and was used as a proxy for time of day. This was included to control for potential diurnal rhythms in fish behaviour. We also fitted an interaction term of *trial number * density treatment,* in case any systematic changes in observed trait means across trials (due to e.g. age effects, habituation etc.) are themselves treatment dependent. Wald *F*-tests were used to test the significance of fixed effects in the models.

2.3.5.3 Random effects

By including individual identity as a random effect, we then partitioned multivariate phenotypic (co)variance not explained by the fixed effects into an among-individual and a within-individual (residual) component. The former is estimated as the variance-covariance matrix **I** that contains estimates of the among-individual variance (V_I) component for each trait on the diagonal and estimates of the corresponding covariance between trait pairs (COV_I) off the diagonal. The within-individual component is similarly estimated as a residual variance-covariance matrix (**R**). We make the standard assumptions that residual errors are normally distributed and uncorrelated across observations, and that (co)variance parameters in **I** and **R** are homogeneous across levels of the fixed effects (i.e. density treatments, trial number, stack etc). Although the two experiment-specific sets of traits are not observed in the same trials, we grouped the data by trial period, (e.g. OFT1 with EET1). Thus, we modelled a residual covariance term between OFT and EET traits observed within each trial period. Repeatability (R_I) was then estimated for each trait as the among-individual variance (V_I) divided by total phenotypic variance (V_p) (where V_P is the phenotypic variance conditional on the fixed effects; i.e. $V_P = V_I + V_R$). Between each pair of traits (1, 2) the among-individual covariance (COV_I) was rescaled to give the corresponding correlation r_I (where $r_{I(1,2)} = COV_{I(1,2)} / \sqrt{(V_{I1} * V_{I2})}$).

2.3.5.4 Testing model significance

To test the statistical significance of among-individual behavioural variation we compared the likelihood of our full multivariate model to two further models. In the

first of these, we fitted **I** as a diagonal matrix such that the model allows among-individual variance V_I for each trait, but assumes Cov_I is zero between all trait pairs. In the second, a null model, we removed the random effect of individual identity completely. Comparison of the diagonal model with the null model using likelihood ratio tests (LRT) allows a global test of the significance of among individual behavioural variance (Wilson et al., 2010a). Comparison of the full model with the diagonal model, again by LRT, allows a statistical test of whether **I** contains significant between-trait covariance structure (Wilson et al., 2013). LRT were performed by estimating $\chi^2_{\mathbf{n}DF}$ as twice the difference in model log likelihoods, with the number of degrees of freedom (n) equal to the number of additional parameters to be estimated in the more complex model.

The above analyses were then repeated using long- and short-term data subsets to estimate the corresponding matrices $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ and associated parameters. Note that, following the conclusion of the LT, the density treatments were no longer applied and the 32 fish used in the ST were housed together in the same stack. Therefore, the fixed effect *stack* was redundant and omitted from the models for the short-term subset analyses.

2.3.5.5 Eigen analysis

To further investigate the structure of **I**, $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$, we subjected each matrix to eigenvector (EV) decomposition. This allowed us to examine: a) how much variance is captured by the first axis (EV1) of multivariate behaviour in each case, b) whether factor loadings of individual traits onto EV1 are consistent with an interpretation of boldness-shyness and c) whether EV1 is similar in $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$. To provide a quantitative measure of how similar the multivariate behavioural axes emerging from the long- and short-term data sets were, we calculated the angle (θ) between the first eigenvectors of $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$. An angle of θ = 0° equates to the vectors being perfectly aligned, meaning that EV1, i.e. the axes of multivariate behavioural variation in $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ are identical. Conversely, an angle of θ = 90° would indicate the vectors are orthogonal (and thus maximally differentiated) to each other across the two different time periods (i.e. the major axis of behavioural variation across the two studies are independent).

2.3.5.6 Testing Eigen significance

Uncertainty around the factor loadings for individual traits on EV1 (for **I** matrix) and around θ was estimated using a parametric bootstrap approach (similar to that

outlined in the appendix of (Morrissey et al., 2012)). We simulated 5000 replicate draws of $\mathbf{I}$, $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ from multivariate normal distributions using the maximum likelihood estimates of these matrices as the means, and the variance-covariance matrices of their elements to define the variances. In each case the 5000 simulated matrices were subject to Eigen decomposition. Uncertainty around the point estimates of trait-specific factor loadings was then described using the 95% highest probability density (HPD) interval for the simulated values of these loadings (for $\mathbf{I}$, $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ respectively). Note that these intervals should be viewed as approximate as they are vulnerable to departures from multivariate normal assumptions. By comparing 5000 pairs of simulated LT and ST matrices we similarly estimated the uncertainty around our point estimate of θ. Note however that since θ cannot be less than zero, we also generated a null distribution for the estimator in the absence of any difference between (true) $\mathbf{I}$ matrices. This was done by comparing the leading eigenvector of each of the 5000 replicate draws of $\mathbf{I}_{LT}$ (simulated as described above), to the leading eigenvector of a second matrix, simulated with the same mean (i.e. the REML point estimate of $\mathbf{I}_{LT}$) but a variance equal to the estimated variance-covariance matrix from the short-term study. Thus the null distribution represents θ estimates given that i) the angle is zero since true $\mathbf{I}$ matrices are identical (and equal to the REML estimate of $\mathbf{I}_{LT}$, but ii) the second (short-term) matrix (and so its leading eigenvector) is estimated with greater uncertainty due to the lower sample size.

2.3.5.7 Data subset correlation

Finally, we compared V_I estimates in LT and ST data subsets, and tested the among-individual, across data subset correlations ($r_{I(LT,ST)}$). For each behavioural trait (x) we used a likelihood ratio test to compare a bivariate model of x_{LT} and x_{ST} where V_I is constrained to be equal, to a model where it is free to vary. This tests the hypothesis that among individual variance differs across data sets. (Note that since traits are analysed in observed standard deviation units V_I can also be interpreted as the repeatability estimate unconditional on fixed effects). We then expanded this model to estimate the among-individual, across data subset correlation ($r_{I(LT,ST)}$) and tested this against null hypotheses of both $r_I = 0$ and $r_I = +1$. Estimation of this correlation is possible since the 32 fish used in the short-term study were a subset of the long-term study. If $r_I = +1$, then this indicates that the ranking of phenotypic merits (i.e. each individual's repeatable component of the observed trait) is the same across data sets. However, if $r_I = 0$, then an individual phenotypic merit in the long-term study is

uncorrelated with the repeatable component of that same behaviour observed over a short time period in later life.

2.4 RESULTS

In total, 1235 sets of behavioural observations were conducted from a possible 1492, the difference being due to mortality of some fish over the study period. Summary data for all behavioural traits are presented in Appendix2, Figure A2.1. In EET, the number of fish emerging from the refuge within the observation period was lower than anticipated based on pilot data (LT = 526/2448, ST = 100/318), resulting in severe censoring of Latency to Emerge data. We therefore elected to use only the binary Emergence trait from this trial type in our analyses.

2.4.1 ANALYSIS OF FULL DATA SET

There was significant among-individual variance in multivariate behaviour (diagonal model versus null model, $\chi^2_5 = 125.6$, $P<0.001$), as well as among-individual covariance among traits (diagonal model versus full model, $\chi^2_{10} = 101.8$, $P<0.001$). Estimates of individual repeatability (R_I (±SE)) were low to moderate, ranging across traits from 0.055 (±0.024) for Emergence (on the observed scale, estimated by REML) to 0. 192 (±0.029) for Time in Middle (Table 2.3). Based on univariate models, V_I was statistically significant at $P<0.05$ for all traits (Table A1.2). The estimated fixed effects are not directly relevant to present objectives; however they are presented in full in Appendix 1 (Table A1.3).

Between traits, the signs of all among-individual correlations (r_I) were positive, consistent with our *a priori* expectations (Table 2.3). The OFT traits Track Length, Activity and Area Covered were all strongly correlated (and nominally significant based on $|r_I|$ > two standard errors); however while Time in Middle was strongly correlated with Area Covered ($r_I = 0.653 \pm 0.075$, Table 2.3), it was only weakly associated with the other OFT traits. The EET trait Emergence was positively correlated with each OFT trait (r_I estimates ranging from 0.304 with Track Length to 0.577 with Activity, Table 2.3).

Eigen analysis of **I**, estimated from the full data set revealed that the first two vectors explained 64 % (eigenvector 1, EV1) and 26 % (eigenvector 2, EV2) of the repeatable among-individual variation respectively (Figure 2.2). The trait loadings on the dominant vector EV1 are consistent with an interpretation of this axis of variation as boldness (or arguably exploration and/or general activity; see discussion). Thus

individuals that tended to emerge repeatedly in the EET, swim longer distances, are more active explore more area, and spend more time in the middle of the OFT tank. By comparison, EV2 trait loadings show this axis to be dominated by time spent in the middle of the tank. Track Length and Activity load on this vector to a lesser extent and with an opposing sign to Time in Middle, while the other traits show limited contributions to EV2 (Figure 2.2b).

Table 2.3 Among-individual variance/covariance matrix (**I**) from the multivariate analysis of a) all data, b) long-term study and c) short-term study. Estimates of variance (V_I, diagonal) with among-individual between-trait covariances (COV_I) below the diagonal and among-individual between-trait correlations (r_I; above the diagonal). Standard errors are shown in parentheses for all parameter estimates. Traits: Track Length (TL), Activity (Act), Area Covered (AC), Time in Middle (TIM), Emergence (Em).

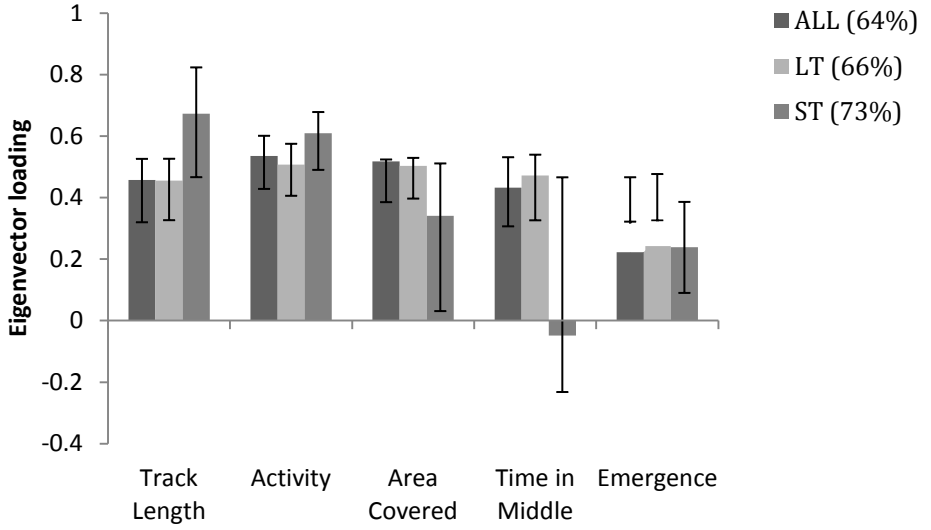
a) All Data	TL	Act	AC	TIM	Em
TL	0.130 (0.025)	0.865 (0.033)	0.750 (0.069)	0.162 (0.117)	0.304 (0.198)
Act	0.124 (0.024)	0.159 (0.026)	0.731 (0.065)	0.241 (0.106)	0.577 (0.182)
AC	0.097 (0.022)	0.104 (0.022)	0.128 (0.026)	0.653 (0.075)	0.414 (0.202)
TIM	0.026 (0.019)	0.042 (0.020)	0.102 (0.023)	0.192 (0.029)	0.540 (0.180)
Em	0.026 (0.017)	0.054 (0.018)	0.035 (0.018)	0.056 (0.019)	0.055 (0.024)
b) Long-term	TL	Act	Area	TIM	E
TL	0.143 (0.028)	0.892 (0.030)	0.777 (0.069)	0.238 (0.118)	0.272 (0.192)
Act	0.137 (0.026)	0.164 (0.028)	0.708 (0.072)	0.314 (0.106)	0.539 (0.180)
AC	0.108 (0.025)	0.106 (0.025)	0.136 (0.030)	0.704 (0.075)	0.458 (0.208)
TIM	0.041 (0.022)	0.058 (0.022)	0.118 (0.026)	0.207 (0.033)	0.607 (0.181)
Em	0.027 (0.020)	0.058 (0.020)	0.045 (0.021)	0.073 (0.022)	0.071 (0.028)
c) Short-term	TL	Act	Area	TIM	E
TL	0.458 (0.155)	0.926 (0.041)	0.640 (0.182)	-0.247 (0.256)	1.070 (0.513)
Act	0.381 (0.137)	0.369 (0.134)	0.812 (0.112)	0.017 (0.274)	1.001 (0.502)
AC	0.188 (0.095)	0.214 (0.097)	0.188 (0.089)	0.492 (0.222)	0.545 (0.524)
TIM	-0.083 (0.089)	0.005 (0.084)	0.106 (0.079)	0.248 (0.101)	-0.667 (0.557)
Em	0.165 (0.080)	0.139 (0.073)	0.054 (0.056)	-0.076 (0.059)	0.052 (0.066)

As noted earlier, our REML analysis makes an assumption of (multivariate) residual normality that is violated by inclusion of the binary trait Emergence. Univariate analysis of Emergence using MCMCglmm, calculated following equation 15 of Nakagawa and Schielzeth (2010), yielded a slightly higher estimate of repeatability (on the liability scale) with a posterior mode of R = 0.090, 95% HPD interval 0.024 – 0.177, Appendix 1, Table A1.1). While noting that the interval will never span zero since R is constrained to lie in positive parameter space, the posterior mode is nonetheless distinct from zero (Appendix 2, Figure A2.2). Bivariate models (i.e. the use of one OFT trait plus Emergence as the phenotypic variates) also confirmed the presence of strong positive among-individual correlations (r_I) between Emergence and OFT traits. Thus,

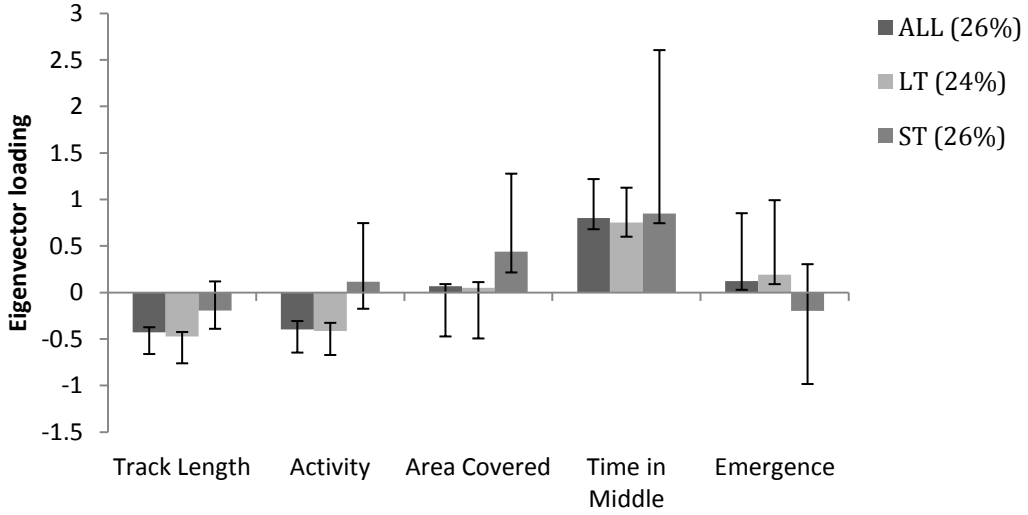
the MCMCglmm analyses corroborate the results of the REML analysis for Emergence (Appendix 1, Table A1.1).

Figure 2.2 Eigenvector decomposition of **I** for all data combined (ALL), long- (LT) and short-term (ST) data sets, with percentage of variance explained in parentheses. Shown are the trait loadings in standard deviation units for the first (a) and second (b) eigenvectors. Error bars show 95% HPD intervals from the parametric bootstrap (see text for details). Note that the point estimates of EV1 loadings on Emergence in ALL and LT datasets actually lie outside the simulated intervals. This reflects sensitivity of intervals estimates to departures from multivariate normality assumed in the bootstrap.

a)



b)

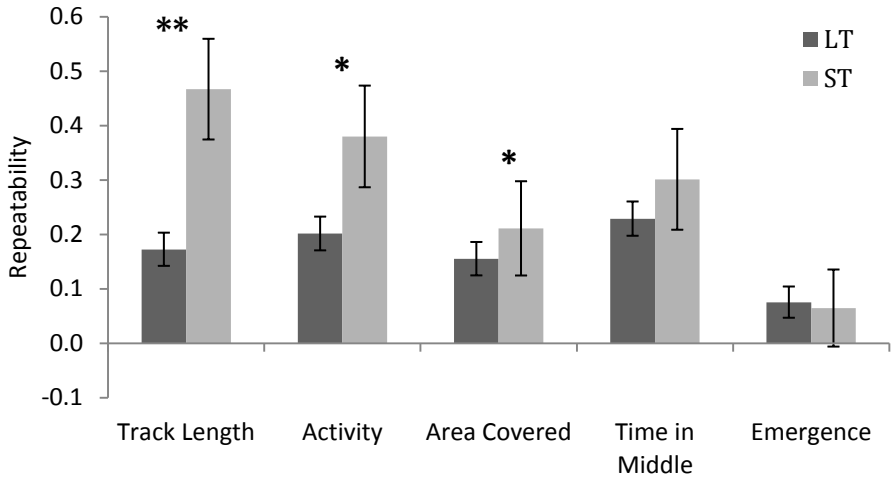


2.4.2 *COMPARISON OF LONG- AND SHORT-TERM RESULTS*

In both the long- and short-term studies, the presence of repeatable variance was statistically supported (comparisons of null and diagonal model: LT $\chi^2_5 = 77.0$, $P < 0.001$; ST $\chi^2_5 = 29.7$, $P < 0.001$) as was the presence of between-trait among-individual covariance structure (comparisons of diagonal and full multivariate model: LT $\chi^2_{10} = 95.0$, $P < 0.001$; ST $\chi^2_{10} = 54.9$, $P < 0.001$). Univariate models confirmed that V_I was statistically significant for all OFT traits in both LT and ST, but not for Emergence in ST (Appendix 1, Table A1.2).

The estimate of $\mathbf{I}_{LT}$ is very similar to that obtained using all data (as described above), not unexpected given that the long-term study contributes the bulk of the total data set. However, comparison of $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ (and derived parameters thereof) indicates some differences in the structure of among-individual behavioural variation as estimated from our long- and short-term studies (Table 2.3). Note that the smaller size of the short-term data set means that the estimates are less precise for this study; this is reflected in the larger standard errors associated with the parameters. Repeatability estimates (R) were higher in the short term study across all traits. However the increased R from ST was particularly striking for Track Length (Table 2.3, Figure 2.3). For this trait, along with Activity and Area covered the null hypothesis of equality of (V_I) across data sets could be rejected (comparison of bivariate models with homogeneous and heterogeneous V_I, $P < 0.05$, Figure 2.3).

Figure 2.3 Estimated trait repeatabilities from long- (LT) and short-term (ST) studies. Error bars specify one standard error. P-values (** = P <0.01; * = P <0.05) indicate significant differences between V_I based on likelihood ratio tests (see text for detail)



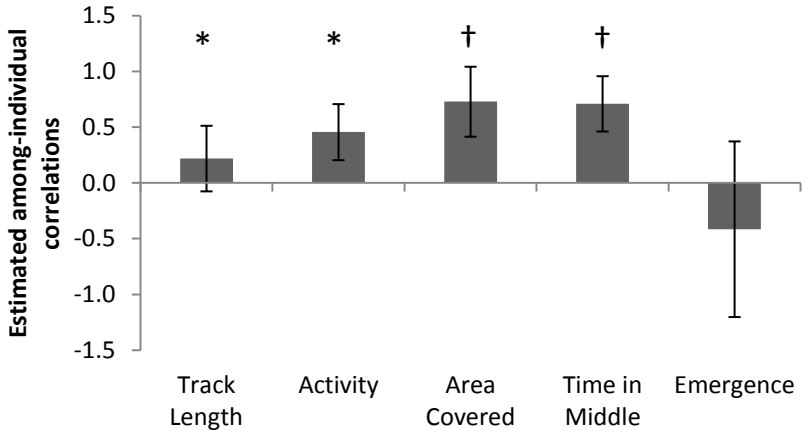
The among-individual between-trait correlations (r_I) reveal a broadly similar structure for the long- and short-term studies (Table 2.3). Thus estimates for ST largely confirm our *a priori* expectation of positive correlation structure between the OFT traits and Emergence. One qualitative exception to the expected pattern is provided by Time in Middle. In LT this trait is positively correlated with all other traits as expected; however, in ST the sign of r_I is negative (but not significant) between Time in Middle and Track Length and Emergence (Table 2.3).

Eigen decomposition confirms the view that qualitative differences between $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ are largely related to Time in Middle. Thus, in both data sets the first eigenvector again dominates the variance in **I** (accounting for 66% and 73% in long- and short-term respectively), consistent with an important latent character underlying behavioural variation (Figure 2.2a). Time in Middle has a strong positive loading on $EV1_{LT}$, consistent with our *a priori* expectation that a bold fish would spend more time in the middle of the open field arena, the corresponding loading coefficient is close to zero (in fact slightly negative) on $EV1_{ST}$. The angle (θ) between $EV1_{LT}$ and $EV1_{ST}$ is 34.63° (95% HPD interval, 5.03- 53.09°). While the point estimate of 34.63° indicates at least some divergence between the leading eigenvectors on a scale from 0 (no difference) to 90 (axes are orthogonal), it is not significantly greater than the angle expected by chance if the true matrices are identical (95% HPD of the null distribution for θ generated by our parametric bootstrap is from 1.54 – 69.14°). While we acknowledge that our null distribution indicates low statistical power to reject the null hypothesis that θ = 0 (see Appendix 2, Figure A1.3), our conclusion is however that $EV1_{LT}$ and $EV1_{ST}$ are broadly similar, with qualitative differences largely attributable to the decreased loading of TIM on $EV1_{ST}$. This is further evidenced by a drop in θ from 34.63° to just 11.15° for the corresponding comparison of **I** estimates excluding Time in Middle. There are also some qualitative inconsistencies evident between $EV2_{LT}$ and $EV2_{ST}$ for the OFT traits, due to greater loadings on Track Length (changes sign), Activity, Area Covered and Time in Middle, while the loading on Emergence is reduced (also changes sign) (Figure 2.2b). The angle (θ) between $EV2_{LT}$ and $EV2_{ST}$ = 48.32° (95% HPD interval 25.75- 86.48°) that again is not significantly different from null expectations.

For those individuals tested in both long- and short-term studies, the among-individual correlations between LT and ST data sets were positive (although not always significant based on likelihood ratio tests) for OFT traits (Figure 2.4) ranging from

0.219 (± 0.294) to 0.729 (± 0.314). Estimates were significantly greater than zero for Area Covered and Time in Middle. However, we also found that the correlation was significantly less than 1 for the traits Track Length and Activity. Thus, while phenotypic performance of an individual in one data set may be predictive of its behaviour in the other, there is also evidence that the ranking of individuals, at least for Track Length and Activity, significantly differs between long and short term studies. For Emergence the corresponding among-individual correlation estimates between long- and short-term were actually negative, though not significantly so. In fact the estimate was characterised by so much uncertainty that despite being negative it was not possible to reject the null hypothesis of r = +1. We suggest this is a result of the low repeatable variation of Emergence and thus little weight should be placed on this result.

Figure 2.4 Estimated among-individual correlations (r_I) between LT and ST data sets for each observed trait, with standard error bars. Each correlation was tested against two null hypotheses of interest: i) $r_I = 1.0$ (* = P<0.05), and ii) $r_I = 0.0$ († = P<0.05), using likelihood ratio tests to compare unconstrained and constrained models (see text for details)



2.5 DISCUSSION

Data from our long-(LT) and short-term (ST) studies provide evidence of among-individual variance in behaviour, both when considered separately and in combination. Of the five traits assayed in the two distinct types of behavioural trial - open field (OFT) and emergence and exploration (EET) - repeatabilities were statistically supported in all cases. In addition our analyses support the presence of a significant among-individual correlation structure for behavioural traits in **I**. Correlation structure is found both within- and across-contexts (i.e. trial types), indicating behavioural variation among fish that is consistent with accepted definitions of animal personality. We found that repeatabilities of OFT traits were higher than the EET though not

significantly so in all cases. Our results therefore support the assertion of Burns (2008) that the OFT is a good and reliable test of boldness and exploratory behaviour in small fish, although it is certainly possible that the EET could be better optimized to target the among-individual component. We discuss the biological interpretation of (multivariate) variance within these two trial types further below. However, here we note the pragmatic consideration that the binary distribution of Emergence data obtained from the EET is more difficult to analyse and interpret while the censoring of Latency to Emerge created a data distribution not readily modelled in any software. Although such problems are likely surmountable by modification of the behavioural assay (e.g. using an extended observation time to eliminate or at least reduce censoring), at least in this case it is not clear to us that the EET provides additional biological insight.

2.5.1 Comparison of long- and short- term data sets

Comparison of long- and short-term data sets suggested that the patterns of individual (co)variance between traits frequently used to define boldness are relatively stable. Nevertheless, as predicted *a priori* we found a tendency for the magnitude of R_I to decrease with a higher interval between observations, at least in OFT trials. For example, repeatabilities for OFT traits ranged from 0.188 to 0.458 in the short term data (with repeat observations at an average interval of four days) but 0.136 to 0.207 in the long term data (average interval of 56 days). In a meta-analysis of behavioural repeatability studies that included either long- (i.e. > 1 year) or short-term (i.e. < 1 year) intervals between observations, the average (median) across all estimates was 0.37 (Bell et al., 2009). Here our repeatability estimates pertain to correlated traits and are therefore not independent. Nevertheless, apart from our short-term study estimates for Track Length and Activity, we note that our estimates for all other traits were lower than those of the meta-analysis average. Repeatability estimates from short-term studies in the meta-analysis (Bell et al., 2009) outnumbered those from long-term studies by 11:1; however, our study considers observations collected within two distinctly separate periods across individual lifetimes.

Arguably the more important question to be asked of our long- and short-term data sets concerns the stability of correlation structure within the multivariate **I** matrix and the interpretation of boldness from its eigenvector decomposition. As seen with the single trait repeatabilities, the structure of $\mathbf{I}_{LT}$ mirrored that of **I** estimated from all data combined. This is unsurprising given that the long-term data comprised a much

greater number of individuals and will thus dictate patterns in the combined dataset. $\mathbf{I}_{LT}$ is dominated by a single vector that is broadly consistent with our expectations of boldness. Significant within- and between- trial type correlations indicate that individuals emerging from the EET refuge are more likely to have high scores for all OFT traits, thus matching our expectation of bold behaviour.

Though not statistically significant, qualitative differences between $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ were apparent. These differences were focussed around the sign and strength of correlations between Time in Middle and traits from both trial types, indicating that both bold and shy individuals from the short-term study spent a similar amount of time in the middle, whereas in the long-term study shy fish behaved in a more thigmotaxic manner. This pattern was reflected in comparisons of the major eigenvectors of long- and short-term data, where a moderate, albeit not statistically significant, angle (θ) between the first long- and short-term axes was estimated. Furthermore, if Time in Middle is dropped from the calculation, the estimated angle is reduced by more than half. Thus our interpretation is that both data sets reveal a major vector of among-individual (co)variance in observed behavioural traits. This vector is similar in the two data sets and can be interpreted as a latent personality trait - namely boldness. In both data sets bolder individuals tend to swim longer distances, be more active and explore more area (in the OFT), and are more likely to emerge from a refuge (in the EET). However, tendency to spend more time in the middle of the OFT arena appears not be a reliable indicator of boldness as it was only associated with this vector in the LT study. Indeed this trait was the major source of qualitative difference between the two matrices.

In the current study it is not possible to distinguish whether higher repeatabilities and the changing structure of **I** with regard to Time in Middle are a consequence of the sampling period (long- vs. short-term) or potentially reflect interesting, possibly even species-specific, biological changes that happen with age and/or trial experience. Note, however, that our analyses control for any habituation effects on mean behaviour, and that we found little statistical support for individual-by-trial-number interactions (results not shown). More generally some authors have argued that individual behaviour is likely to become more rigid and follow more set patterns over time (Roberts and DelVecchio, 2000). If so we would predict increasing repeatabilities with age (here confounded with time scale of data collection). Conversely, others suggest that in the absence of any disturbance (e.g. in a constant laboratory environment),

expectations of changes to individual patterns of behaviour formed in early life are ill-founded (Stamps and Groothuis, 2010). While no overall differences were found between juvenile and adult behavioural repeatabilities in the Bell et al. (2009) meta-analysis, a subset of data suggested juvenile behaviour to have higher repeatability than that of adults. However, the metanalysis contained only three studies that included observations following individuals through from juvenile to adult status. Thus direct comparison of age classes is not straightforward. Clearly more empirical studies of how repeatability changes with age would be valuable, as indeed would parallel studies exploring environmental dependence. Here we assumed homogeneous variance structures across environments (density treatments, stacks) and other fixed effects (sexes, day order) for simplicity. However, these assumptions can be relaxed in the statistical models to test for and quantify individual by environment interactions (IxE) as changes in the among-individual variance (or structure of I in the multivariate case, Dingemanse et al., 2010). Here *post hoc* analyses of the LT data set provides some evidence of heterogeneous repeatabilities across density treatment classes (see Appendix 1, Table A1.4). Though not expected to bias current conclusions (parameter estimates presented are effectively averaged across treatments), if robust this effect may certainly be biologically interesting.

The population level patterns of among-individual (co)variances between traits were broadly similar between LT_I and ST_I, albeit with some differences as described above. However, by using the same individuals in both long- and short-term studies we were able to address the question of whether the relative ranking of individuals with respect to their behavioural tendencies was stable. The estimates of r_I for each observed behavioural trait between the long- and short-term datasets provide a mixed answer to this question. Positive correlations for the OFT traits do show a degree of stability in (repeatable) behavioural tendencies across the data sets though statistical support was mixed and it appears individuals were more likely to maintain a consistent ranking for some traits (e.g. Area Covered) than others (e.g. Track Length).

2.5.2 CONCLUSION

We previously stated it is not our intention to be prescriptive about what boldness is or how it should be assayed. Nevertheless, *a priori*, we anticipated that in the OFT, bold fish would travel long distances and be willing to visit a large area of the tank including the central zone, and that these traits would correlate significantly with whether individuals emerged in the EET. However, this depiction requires that the bold

individual is also active and/or exploratory. Above we have noted that the major axis of variation in **I** is largely consistent with expectations of a bold-shy continuum as the terminology is used in the literature; however, the strength of among-individual correlations suggests that it could equally be called exploration or general activity in a novel environment. Nevertheless, as qualitatively almost all the variance loads onto this single axis of variance, we conclude that these continuums (personality axes) are, at least in our study species, either the same entity or so tightly correlated that attempting to distinguish between them may have little practical value. Indeed, Burns (2008) concluded that emergence from a refuge was difficult to interpret strictly as either boldness or exploration, even though it has been described as boldness only by others (e.g. Budaev, 1997b; Brown et al., 2005a). Exploring the functional significance of the consequences of this behavioural variance in wild populations is likely to yield more insight than further debate with regard to terminology (e.g. Brown et al., 2005a; Dingemanse et al., 2012b; Kurvers et al., 2012; Carvalho et al., 2013). Nonetheless, we have sufficient statistical support in our results to conclude that both trial types revealed behaviours characteristic of boldness, evident from the strong among-individual correlations between all the observed traits. This again leads us in the direction of Burns' (2008) view that in practice, the OFT offers the most useful test arena for this axis of personality.

Here we have obtained repeated measures of multiple behavioural traits during two test types and across two distinct sampling periods (long- versus short-term), something that has seldom been accomplished in the literature. In practical terms, we conclude that the open field trial is preferable to the emergence and exploration trial as an experimental test for investigating boldness, and we show how eigen decomposition of an **I** matrix can usefully identify latent personality traits. This multivariate approach is broadly similar to that used in several other recent studies (Budaev, 2010; Carter et al., 2013; Araya-Ajoy and Dingemanse, 2014). Our study also provides information about the stability of personality, both in terms of population level patterns and individual differences. We find that observed behavioural traits are repeatable over long time periods as well as when observations are made over only a few weeks, although there is a tendency for short term estimates to be higher. Taking a multivariate approach we show that **I** is dominated by a single vector through phenotypic space that is similar across the two study periods and can be interpreted as boldness. We note however, there are at least some qualitative differences in the

relationships of observed behaviours to this vector. At the individual level we also find qualified support for the proposition that short-term studies are informative for an individual's behavioural phenotype over longer (e.g. lifetime) periods.

CHAPTER 3

HE WHO DARES ONLY WINS SOMETIMES: PHYSIOLOGICAL STRESS AND CONTEST BEHAVIOUR IN XIPHOPHORUS HELLERI

This chapter is published as: **Boulton K**, Sinderman B, Pearce RM, Earley RL, and Wilson AJ (2012). He who dares only wins sometimes: physiological stress and contest behaviour in *Xiphophorus helleri*. Behaviour 149:977-1002.

3.1 ABSTRACT

While many factors influence contest outcome and social dominance in animals, there is increasing interest in behavioural-physiological stress coping styles (SCS). Causality, however, is often ambiguous – is physiological state determined by contest outcome or *vice versa*? Furthermore, experimental protocols may themselves induce stress responses that impact individual behaviour and thus potentially contest outcome. Here we test whether latency to recover from acute stress, measured both physiologically and behaviourally, predicts who initiates and who wins dyadic contests between pairs of male green swordtails *(Xiphophorus helleri)*. In line with our predictions, animals that recovered faster (behaviourally) from disturbance created by the experimental protocol prior to meeting an opponent were more likely to initiate contests; however, they were not more likely to win and, contrary to expectations, had higher pre-contest cortisol levels than their opponents. They also showed greater physiological stress responses to the experiment as determined from the difference between pre- and post-contest cortisol levels. Moreover, stress response was independent of whether a contest escalated. In contradiction to evidence found in other taxa and fish systems, the suite of traits that we measured were not correlated in a manner that allowed classification of the animals into the usual reactive and proactive stress coping styles. Our results suggest that coping style may play a key role in determining exact individuals that initiate contests, but that other factors govern contest outcome.

3.2 INTRODUCTION

Competition for resources such as food, mates or territory, often involves contests where winners, or dominant individuals, improve their fitness at the expense of losers (Brockelman, 1975). Many factors are expected to influence contest outcome and so determine dominance status. While these are known to include size (e.g. Huntingford et al., 1990) and behavioural traits such as aggression (Francis, 1988), individual styles

of coping with stress may also be important (Koolhaas et al., 1999; Pottinger and Carrick, 2001; Øverli et al., 2004). Stress threatens homeostasis that is re-established by both physiological and behavioural responses. Importantly, when studying behaviour, experimental protocols may induce stress responses that impact individual behaviour, thus indirectly influencing eventual contest outcome. Here we explore the hypothesis that latency to recover from stress, as measured both behaviourally and physiologically, is a key determinant of contest initiation and outcome. In animals, physiological stress-coping mechanisms are highly conserved and governed by the hypothalamic-pituitary-adrenal (HPA) axis. In fish this role is assumed by the hypothalamic-pituitary-interrenal (HPI) axis with water-borne cortisol being a good physiological indicator of HPI activity (for a review, see Scott and Ellis, 2007; Scott et al., 2008).

Classically, much research on animal conflict has focused on the concept of resource holding potential (RHP; Parker, 1974). Commonly used measures of RHP (e.g. body size) often predict contest initiation and outcome, although resource ownership, individual motivation and social processes such as eavesdropping and prior fighting experience are also important (Hsu et al., 2006; Arnott and Elwood, 2008). Studies that attempt to control for RHP, for example by size matching and using neutral arenas, have suggested that individuals initiating contests tend to win them (Jackson, 1991). However, this is not always the case (Moretz, 2003), suggesting that factors other than the initial motivation to fight may affect contest outcome especially during escalated contests (Hsu and Wolf, 2001).

The relationship between physiological stress (HPA/HPI axis activity) and social dominance has received increasing attention and has been well studied across many taxa, including rodents (Bronson, 1973), primates (Abbott et al., 2003), birds (Verbeek et al., 1996), mammals (Young et al., 2006), domestic livestock (Bergsma et al., 2008) and fish (Øverli et al., 2007). Moreover, causality is often ambiguous and it is unclear whether physiological state is determined by outcome, or outcome is determined by physiological state. For example, faster recovery of baseline cortisol levels following aggressive contests is associated with dominance (Netherton et al., 2004), while individuals with higher baseline cortisol levels are less likely to win contests or to obtain dominance status in a hierarchy (Hannes, 1984; DiBattista et al., 2005). Other types of behavioural variation may be linked to physiological stress, particularly an

individual's stress coping style (Earley et al., 2006). In a study focusing on both behavioural and neuroendocrinological parameters, Koolhaas et al. (1999) contrasted *proactive* and *reactive* coping styles and suggested a proactive/boldness link (boldness is here described as a willingness to explore novel environments, Budaev, 1997a). Many studies have demonstrated correlations between boldness and aggression (for example, Bell and Sih, 2007), and of specific interest is that in fishes, empirical measurements of HPI activity, aggression and boldness have been associated with differences in coping style (Schjolden et al., 2005; Aubin-Horth et al., 2012).

The majority of studies investigating contest behaviour and dominance in domestic and wild fishes use experimental designs that require netting individuals to facilitate periods of isolation. This is usually followed by some form of disturbance, such as the removal of partitions between isolated contestants in novel environments (for example, Wilson et al., 2011a). Could it be that contest winners are those that better cope with stressors imposed by the experimental protocol prior to even encountering an opponent? If so, then aspects of personality (e.g. boldness) and/or stress coping style may predict observed aggression and contest outcome.

Here we test the effect of disturbances imposed by the experimental protocol on contest behaviour and outcome using the male green swordtail (*Xiphophorus helleri*), a small, tropical freshwater, live-bearing fish. Due to their readily aggressive nature, species from the *Xiphophorus* genus are commonly used as behavioural models in studies of dominance and many such studies have focused on visual and social cues as explanations for conflict resolution (Earley, 2006). However, we hypothesize that if coping style is important in the determination of observed contest behaviour under experimental situations, then relationships should exist between the behavioural reaction to disturbance prior to meeting an opponent, the likelihood of initiating a contest, contest outcome and the physiological stress response as measured by cortisol levels. Specifically, we predict that a short latency to resume normal swimming behaviour following disturbance will be associated with fish that initiate and win contests; such animals are predicted to be less stressed, i.e. have lower baseline (pre-contest) cortisol levels and a smaller stress response (post contest minus pre-contest cortisol level) than the eventual losers.

3.3 METHODS

Green swordtails (*Xiphophorus helleri*) obtained from a commercial distributor were housed in heterosexual groups in 152 and 208 L aquaria equipped with gravel substrate (3 cm), filtration, and aeration. Water temperature was maintained between 23 - 25° C, pH between 7.2-7.6, and fish were kept on a 12 h light: 12 h dark photoperiod. Stress Coat™ (94 µl/L) and freshwater aquarium salt (2 g/L) were added to the tanks prior to fish arrival to mitigate the loss of fish mucus and to reduce osmotic stress, respectively; each of these is a common response of fish to shipping and handling.

3.3.1 DYAD ESTABLISHMENT

Males were netted from the aquarium and placed in a plastic bag with a small amount of water to keep the gills and body moist and to immobilize the fish for measurement; measurements were taken with Vernier calipers accurate to 0.1 mm. Measurements of standard length (SL, snout tip to caudal peduncle), total body length (snout tip to caudal fin tip), body depth (BD, anterior portion of dorsal fin to origin of gonopodium), and sword length (SwL, caudal fin tip to sword tip) were obtained. Pairs of males for dyadic trials were matched for lateral surface area (LSA; < 20 units difference) because LSA has been shown to be a better predictor of fighting ability than any one measure of size alone (Beaugrand et al., 1996). LSA (mm^2) was determined as:

(standard length * body depth) + (sword length * sword depth)

assuming a sword depth of 1.0 mm. Body markings and coloration were also noted for purposes of identification. Macromelanophore patterns and sword characteristics were used to discriminate the two opponents (Franck et al., 2001; Basolo and Trainor, 2002). A total of 30 pairs were formed.

3.3.2 CONTESTS AND HORMONE COLLECTION

Immediately after measurements, fish were transferred directly from the plastic bag to 1000 ml polypropylene holding beakers containing 1000 ml of aerated freshwater. Stress Coat™ (94 µl) and freshwater aquarium salt (2 g) were added to the holding container to replace fish mucus and reduce osmotic stress associated with handling during measurement. The holding beakers were outfitted with a fine mesh net bottom and placed inside another 1000 ml polypropylene beaker; this design made it possible to transfer the fish between beakers gently, quickly (< 5 seconds) and without the handling typically associated with capture (e.g. chasing, netting). The fish remained in

the holding beaker for 2 days to acclimate before being transferred to new 1000 ml sampling beakers containing 1000 ml of freshwater (with 4g freshwater salt) for 2 h, with hormones being released into the water during this time (Scott et al., 2008). Stress Coat™ was not added to the hormone collection beaker because it is not known whether the chemical interferes with hormone extraction and assay; freshwater salt, however, can be purged from hormone extraction columns (see below). After 2 h in the pre-fight sampling beaker the fish were transferred using a net to 38 L experimental fight tanks, separated into two equal compartments by an opaque divider. Each compartment was equipped with an aeration device and the water was treated with Stress Coat™ and freshwater aquarium salt. The two fighters were placed on opposite sides of the same fight tank and acclimated for 22 h. After this time the dividers were lifted (remotely) and the air stones were also removed. This physical disturbance typically resulted in frantic swimming behaviour by both fish, characterized by fast, erratic movements both horizontally and vertically before the fish settled to the gravel bottom. We therefore consider it to be a response imposed by the experimental protocol itself. We determined the latency of behavioural recovery from this event as the time (from lifting of partition) to resume normal swimming, defined as swimming slowly in a horizontal orientation with fins often erect or semi-erect.

The fish then interacted until a dominance relationship was established, defined as the point when one individual retreated 10 consecutive times without reciprocating aggression or displayed typical submissive posturing, such as folding fins upon approach from the opponent (Franck and Ribowski, 1989; Beaugrand, 1997). Contests lasted for an average of 2286 ± 441 seconds and were recorded digitally using a Sony PC110 Digital Video camera then burned to DVD. The identity of the animal that first began swimming normally following partition removal, initiated the contest (approached within one body length of the opponent), and won the contest was recorded using JWatcher version 1.0 (Blumstein and Daniel, 2007; http://jwatcher.ucla.edu/). Latency to contest initiation, as well as contest duration (from initiation to settlement) was calculated in seconds from partition lifting. Additionally, we classified each contest as being escalated or not. Escalated contests were defined as those involving high intensity reciprocal attacks, where the opponents would alternate attack-bite sequences often while circling one another, and/or mouthwrestling, where contestants would lock jaws in an apparent test of strength.

Immediately after contest resolution, fighters were netted and placed in individual 1000 ml sampling beakers for 2 h for a post-fight hormone collection.

3.3.3 Hormone Extraction and Radioimmunoassay

C18 SPE columns (Extract-Clean®, 500 mg, 4.0 ml; Alltech Associates, Inc.) were primed with 2 x 2 ml of 100% ethanol (EtOH) and 2 x 2 ml distilled water. Tygon tubing (formulation 2275) was attached to the C18 columns and placed in a beaker containing a 250 ml water sample taken from the original 1000 ml, the vacuum was engaged and water-borne steroid hormones isolated. Total hormone (free and conjugated fractions) was eluted from the columns with 2 x 2 ml 100% ethanol collected in 6 ml (12 x 75 mm) borosilicate vials. Samples were stored at 4°C overnight and the ethanol was evaporated in a Savant AES 1010 speedvac for 1.5 h (45 min at 40°C) one day prior to radioimmunoassay. Hormone residues were resuspended in 60 µl of 0.1 M phosphate buffer. Cortisol radioimmunoassay was conducted using a coat-a-count kit purchased through Diagnostic Products Corporation (Los Angeles, CA). Samples were run in duplicate in three separate assays conducted on three consecutive days. Briefly, 25 µl of each sample was pipetted into antibody-coated polypropylene tubes followed by the addition of 1 ml of I^{125}-labeled cortisol. Samples were incubated in a 37°C water bath for 45 min. Liquid in all samples was then decanted, and the tubes were blotted and allowed to air dry for 30 min prior to quantification. The sensitivities of the three assays were 0.0268 µg/dl, 0.033 µg/dl, and 0.0624 µg/dl. Pooled low-, medium- and high-level human serum (CON6 Multivalent Control Module, Diagnostic Products Corporation) were used as intra-assay controls; intra-assay coefficients of variation (assay 1, 2, and 3) were: tri-level low (6.2%, 3.8%, 2.1%), tri-level medium (2.8%, 12.0%, 4.3%), and tri-level high (4.8%, 5.0%, 7.2%). Inter-assay coefficients of variation were 6.4%, 7.5%, and 7.3% for tri-level low, tri-level medium and tri-level high, respectively.

The kit was validated for *X. helleri* by assessing parallelism and by calculating expected versus observed cortisol concentrations from known samples cold-spiked with standards. Twenty non-experimental swordtails (males and females) were transferred to collection beakers filled with 400 ml freshwater for 8 h (0800-1600 h). Hormones were extracted and processed as described above, except that they were resuspended in 120 µl and combined to form a pool of 2.4 ml stored as 55 µl aliquots at –80°C. 240 µl of the pooled control was used for serial dilutions. Briefly, 120 µl of this sample was

transferred to a 1.5 ml Eppendorf tube and mixed by vortexing with 120 µl of 0.1 M phosphate buffer to create a 1:2 dilution; 120 µl of 1:2 dilution was mixed with an equal volume of 0.1 M phosphate buffer to create a 1:4 dilution, and so on until 1:16. The serial dilutions were run in quadruplicate using the RIA protocol described above with the Cortisol Coat-a-Count kit from DPC. The log-logit transformed dilution curve was parallel to the standard curve (comparison of slopes: $t_7 < 0.01$, $P > 0.05$; (Zar, 1996), P. 355). A 385 µl sample of pooled hormone extract was used to assess recovery. 110 µl was pipetted into a tube to constitute the 'neat' (1:1) control. 55µl of the large sample was then pipetted into 5 additional tubes and mixed with an equal volume of each standard provided with the DPC Cortisol coat-a-count kit (1, 5, 10, 20, 50 µg/dl). Expected recovery concentrations were based on the known amount of cortisol in the *X. helleri* control sample. Minimum recovery was 90.3% and the slope of the observed vs. expected curve was 0.97, demonstrating a highly linear relationship between observed and expected recovery.

3.3.4 DATA ANALYSIS

One fish died during the period of post-contest cortisol collection and therefore data relating to the participating trial were excluded from analysis. A further pair was eliminated because they did not interact on any level. A total of 28 contests from the original 30 pairs of fish were therefore observed, with 25 producing clear winners and losers and 15 classified as escalated. The first individual to swim normally following partition removal and the individual that initiated the contest was unambiguously determined in all 28 cases (see Appendix 1, Table A1.5 for raw data on all contests).

In order to summarize associations between the full set of morphological, behavioural, and endocrine traits measured we generated a correlation matrix using Genstat 14.1 (Payne et al., 2005; Blumstein and Daniel, 2007). Correlations between morphological and physiological traits were estimated using the full set of observations (i.e. one record per individual, n=56) for body depth (BD), standard length (SL), sword length (SwL), lateral surface area (LSA), pre-contest (PreF) and post-contest (PostF) cortisol (F) levels and physiological stress response (SR). Endocrine assays before and after the trial were $\log_{10}$ transformed to yield PreF and PostF respectively, while we defined SR as the change in cortisol expression on a $\log_{10}$ scale (i.e. SR = PostF – PreF). For those traits where the phenotypic value of one individual within a trial necessarily determines that of the second, we used observations from one randomly chosen focal individual per trial only (n = 28). These traits include the binary variables of Swimfirst

(whether the focal fish was first to resume normal swimming after disturbance), Initiate (whether the focal fish initiated the contest) and Status (whether the focal fish was the winner). For these randomly chosen focal individuals we also determined a relative measure of size difference (LSAdiff), defined as the difference in phenotypic values (focal LSA – opponent LSA). Correlations with two further traits, latency to swim (LatSwim) and latency to initiate (LatInit) were also estimated. However, these traits are only meaningfully observed for the individual within each trial that either swims first or initiates the contest, respectively. Thus estimated correlations with these variables are conditional on moving first or initiating the contest as appropriate (n = 28).

To more directly test the hypothesized causal relationships between behavioural recovery from disturbance, contest initiation and outcome (i.e. status) and stress response, we formulated a set of linear models that were solved by restricted maximum likelihood using ASReml (Version 3, Gilmour et al., 2009). In particular this allowed us to test our hypotheses while properly accounting for any influence of body size (LSA) on endocrine traits and/or contest behaviour. Note therefore that our phenotypic measures of the endocrine traits (PreF, PostF, SR) are not corrected in any way for the expected influence of fish size (Scott and Ellis, 2007) prior to analysis; rather, the linear model framework allows us to control for these effects statistically within the analysis.

As described above, each contest provides only a single phenotypic observation for the binary traits of Initiate (Model 1) and Status (Model 2) and these response variables were analysed using generalized linear models (with logit link function). Thus we modelled probability (on the logit scale) of initiating a contest as a function of being first to adjust to normal swimming behaviour following removal of the partition (*Swimfirst*), as well as baseline cortisol (*PreF*, size (*LSA*), and all two-way interactions of these explanatory variables such that:

$$Initiate_{ik} = \mu + Swimfirst + PreF + LSA + Swimfirst.PreF + Swimfirst.LSA + PreF.LSA + \varepsilon_k$$

(Model 1)

Where $Initiate_{ik}$ is the probability (on the logit scale) of individual *i* initiating contest *k*, μ is an overall mean, and ε is a residual error term (assumed to be uncorrelated across trials). The probability of winning a contest (*Status*, 0/1) was modelled in a similar

way, but with the addition of fight *Escalation* (as a two-level categorical variable, i.e. whether a fight did or did not escalate) fitted as a factor, and its interaction terms as additional explanatory effects. Escalation is included here because Swimfirst may only predict contest winners when fights do not escalate (e.g. see Hsu and Wolf, 2001).

$$Status_{ik} = \mu + Swimfirst + PreF + Escalation + LSA + Swimfirst.PreF + Swimfirst.LSA + PreF.Escalation + PreF.LSA + Escalation.LSA + \varepsilon_k$$

(Model 2)

Finally we modelled stress response (*SR*) to test the hypothesis that it would be lower for those individuals that had won contests, and particularly so in the absence of contest escalation. Values of SR can be assigned to both individuals within a trial but may not be fully independent. We therefore analysed SR using a linear mixed effect model (with normal error structure) that included a random effect of trial to account for non-independence (Model 3).

$$SR_{ik} = \mu + Swimfirst + Status + Escalation + LSA + Swimfirst.LSA + Swimfirst.Status + Swimfirst.Escalation + Status.LSA + Status.Escalation + Escalation.LSA + Trial_k + \varepsilon_k$$

(Model 3)

For each of the models shown above we adopted a model reduction strategy where explanatory terms were dropped if they were statistically non-significant at $P \geq 0.1$ under a two-tailed conditional *F* - test. Main effects were retained in the model if one or more of their interactions were retained on this basis. Note that we chose to use a threshold of $\alpha = 0.1$ rather than 0.05 in our model reduction strategy and therefore our final models can contain marginally non-significant explanatory terms (i.e. $0.1 \leq P \leq 0.05$). We adopted this strategy as, since available sample sizes are fairly small we expect power will be limiting. However, we deem that it is instructive to consider whether marginally non-significant terms are at least qualitatively consistent with hypothesized biological processes, i.e. it may not be sensible to equate non-significance with an effect size of zero.

Table 3.1 Phenotypic trait correlation matrix. The full data set was used to estimate correlations between the morphology and physiology traits of body depth (BD), standard length (SL), sword length (SwL), lateral surface area (LSA), Pre- (PreF) and post-contest (PostF) cortisol levels, and stress response (SR). The randomly selected half data set was used to calculate correlations between traits with only one phenotypic observation per trial: Status, Swimfirst, Initiate and differences in lateral surface area between opponents in the same contest (LSAdiff). Correlations for the traits latency to swim (LatSwim) and latency to initiate (LatInit) are calculated using one observation per trial, conditional on swimming first or initiating the contest. Bold font denotes a significant correlation (2-tailed P <0.05). Bold italic font denotes a marginally non-significant correlation (2-tailed P <0.1).

	BD	**SL**	**SwL**	**LSA**	**PreF**	**PostF**	**SR**	**Status**	**Swim first**	**Initiate**	**LSA diff**	**Lat Swim**	**Lat Init**
BD	-												
SL	**0.949**	-											
SwL	**0.308**	**0.429**	-										
LSA	**0.987**	**0.984**	**0.391**	-									
PreF	**0.432**	**0.477**	**0.418**	**0.453**	-								
PostF	**0.422**	**0.434**	**0.296**	**0.425**	**0.639**	-							
SR	-0.013	-0.050	-0.144	-0.033	**-0.425**	**0.425**	-						
Status	0.021	0.044	-0.200	0.037	-0.157	-0.191	-0.009	-					
Swimfirst	0.040	0.014	0.001	0.040	0.006	0.273	0.315	-0.116	-				
Initiate	0.016	0.059	0.172	0.042	**0.449**	***0.370***	-0.181	-0.131	**0.559**	-			
LSAdiff	0.104	0.164	-0.164	0.128	-0.181	-0.133	0.092	***0.344***	0.016	-0.202	-		
LatSwim	-0.163	-0.204	***-0.382***	-0.187	**-0.450**	**-0.695**	-0.086	0.157	*	-0.178	***0.344***	-	
LatInit	-0.228	-0.321	**-0.464**	-0.286	***-0.392***	**-0.474**	-0.035	0.359	*	*	***0.412***	**0.642**	-

*Correlation not available

3.4 RESULTS

3.4.1 BETWEEN TRAIT CORRELATIONS

The estimated correlation structure provided evidence of significant associations between a number of the traits measured (Table 3.1). Phenotypic correlations were close to unity between the morphological traits of BD, SL and LSA ($r_{BD.SL}$ = 0.95, $r_{BD.LSA}$ = 0.99, $r_{SL.LSA}$ = 0.98; all P<0.001), perhaps unsurprising given that these all capture aspects of body size. Sword length (SwL) was also positively correlated with body size traits although less strongly. Body size traits were significantly and positively correlated with both pre- and post-contest cortisol levels (r ranging from 0.42 - 0.48, all P ≤ 0.001; Table 3.1) although again the correlation between PostF and sword length (SwL) was lower(r = 0.30, P = 0.03). Given that endocrine traits are not standardised for size variation prior to analysis these results are consistent with the expectation of a positive association between body size and cortisol release into the water (Scott et al., 2008), controlled for in our model based hypothesis testing (as discussed above). Note that stress response (SR) is auto-correlated with pre- and post-contest cortisol levels as a consequence of its definition (i.e. SR = PreF – PostF $r_{SR.PreF}$ = - 0.43, and $r_{SR.Post}$F = 0.43, both P = 0.001). Cortisol levels before and after the contest are also significantly correlated within individuals ($r_{PreF\ PostF}$ = 0.64, P = < 0.001). However, correlations between SR and size (as measured by the various morphology traits) are weak and non-significant.

We found a significant positive correlation between the behavioural traits swim first and contest initiation as we hypothesized (r = 0.56, P = 0.004). For the set of individuals that both swam first and initiated the contest, latency to swim was also strongly correlated with latency to initiate (r = 0.64, P = 0.003). However, although non-significant, swimming first was not positively correlated with status (i.e. winning, r = -0.12, P = 0.58), and among those fish that did swim first the correlation between latency to swim and status was close to zero (r = 0.16, P = 0.45). Thus the correlation structure is consistent with our hypothesis that individuals more rapidly resuming normal swimming after partition removal are more likely to initiate contests. However, these individuals are not more likely to win the subsequent contest.

The correlation structure provided only limited statistical support for relationships between behavioural and endocrine traits. Contrary to our expectation that individuals exhibiting lower baseline cortisol, i.e. presumably less stressed prior to the trial, would

move first, we actually found a positive, albeit weak and non-significant, correlation between preF and Swimfirst ($r = 0.006$, $P = 0.98$). Higher PreF was significantly associated with an increased tendency to initiate the contest ($r = 0.45$, $P = 0.025$). Both PreFand PostF levels were negatively correlated with latency to swim (among fish that swam first) and the relationship was significant in both cases ($r_{PreF,LatSwim} = -0.45$, $P = 0.024$, $r_{PostF.LatSwim} = -0.70$, $P < 0.001$). Negative correlations of similar magnitude were found between PreF and PostF and the latency to initiate a contest; however, only the PostF correlation was significant ($r_{PreF.LatInit} = -0.39$, $P = 0.10$, $r_{PostF.LatInit} = -0.47$, $P = 0.04$) (Table 3.1).

3.4.2 Model based hypothesis testing

Model 1 supported our hypothesis that individuals that swim first would also initiate contests more often ($P = 0.029$); however, contrary to our *a priori* expectation that contest initiators would have lower levels of pre-contest cortisol, higher PreF levels were in fact associated with contest initiators ($P = 0.036$, Table 3.2). These patterns are qualitatively consistent with the significant correlation structure between initiate, PreF and Swimfirst as reported above. The estimated effect of PreF on tendency to initiate was more convincing in the reduced model (3.03 ± 1.37 µg/dl) than in the full model (-7.34 ± 15.64 µg/dl). This could reflect the fact that the latter estimate of the PreF effect is conditioned on the putative dependence on body size (although neither LSA nor its interactions were statistically significant). Model 2 provided no evidence that contest winning is predicted by swimming first or by baseline physiological stress (i.e. PreF). These findings are counter to our second *a priori* hypothesis, but again consistent with the simple correlation analysis. Although we also tested for dependency of these effects on contest escalation and/or size effects, in fact no explanatory variables were retained in the reduced version of Model 2. Thus we were unable to predict contest outcome from size, behaviour, or baseline physiological stress. Finally, although stress response was lower in contest winners as we had predicted, the difference between losers and winners was not significant in the full model (-0.40 ± 0.46 µg/dl, $P = 0.90$) and therefore status was not retained in our reduced model (Model 3). However, based on a marginally non-significant interaction of Swimfirst and size (LSA) ($P = 0.085$, Table 3.2) both variables were retained in the reduced model. Under the full model for stress response, 5 (± 23)% of the observed variance not explained by fixed effects was explained by Trial. Under the reduced model, the corresponding estimate was 14 (±

19)% of the variance. The random effect of Trial is not significant in either the full (P = 0.83) or the reduced (P = 0.49) models.

Table 3 2 ANOVA table of fixed effects fitted in full and reduced linear models of Initiate, Status and Stress Response. Indicated are estimated effect sizes for explanatory terms fitted (with SE in parentheses), and conditional *F*-tests. Initiate and Status are modelled as binary response variables while a normal error structure was fitted for Stress Response. Where used as explanatory variables Swimfirst, Status and Escal were fitted as two level factors with the estimated coefficients denoting the effect of factor level 1 (fish swam first, fish won the contest, contest was escalated) relative to factor level 0. Models of Stress Response also included a random effect of Trial (see text for details).

		FULL MODEL				**REDUCED MODEL**			
Trait	**Fixed Effect**	**Coefficient (SE)**	**DF**	***F***	**P**	**Coefficient (SE)**	**DF**	***F***	**P**
Initiate	mean	-2.37 (8.28)	1,21	0.59	0.449	-1.52 (0.802)	1,25	0.04	0.838
	Swimfirst	9.35 (9.92)	1,21	3.71	0.068	2.48 (1.07)	1,25	5.35	0.029
	PreF	-7.38 (15.60)	1,21	4.61	0.044	3.03 (1.37)	1,25	4.90	0.036
	LSA	0.001 (0.019)	1,21	1.49	0.235				
	PreF.LSA	0.033 (0.036)	1,21	0.84	0.368				
	Swimfirst.LSA	-0.015 (0.023)	1,21	0.42	0.522				
	Swimfirst.PreF	-0.781 (6.50)	1,21	0.01	0.906				
Status	mean	-7.33 (9.19)	1,14	0.84	0.375	-0.080 (0.400)	1,24	0.04	0.843
	Swimfirst	4.27 (8.90)	1,14	0.08	0.778				
	PreF	-2.42 (7.21)	1,14	1.43	0.252				
	Escal	2.32 (7.66)	1,14	0.06	0.804				
	LSA	0.018 (0.021)	1,14	0.86	0.371				
	Swimfirst.PreF	3.72 (3.94)	1,14	0.89	0.361				
	Swimfirst.Escal	1.68 (2.07)	1,14	0.66	0.431				
	Swimfirst.LSA	-0.013 (0.021)	1,14	0.38	0.548				
	PreF.LSA	-0.007 (0.014)	1,14	0.25	0.625				
	PreF.Escal	1.91 (3.18)	1,14	0.36	0.557				
	Escal.LSA	-0.007 (0.018)	1,14	0.15	0.707				
Stress	mean	0.094 (0.455)	1,39	1.99	0.167	-0.220 (0.295)	1,26	0.62	0.438
Response	Swimfirst	0.824 (0.456)	1,19	0.13	0.720	0.758 (0.383)	1,26	0.42	0.525
	Status	-0.402 (0.456)	1,19	0.02	0.903				
	Escal	-0.467 (0.500)	1,20	0.42	0.522				
	LSA	-0.002 (0.001)	1,20	0.02	0.885	0.001 (0.001)	1,26	0.05	0.824
	Swimfirst.Status	0.038 (0.211)	1,20	0.03	0.858				
	Swimfirst.Escal	-0.081 (0.202)	1,19	0.16	0.694				
	Swimfirst.LSA	-0.002 (0.001)	1,19	3.31	0.085	-0.002 (0.001)	1,26	3.56	0.071
	Status.Escal	-0.120 (0.202)	1,19	0.35	0.559				
	Status.LSA	0.001 (0.001)	1,19	1.10	0.307				
	Escal.LSA	0.001 (0.001)	1,20	1.15	0.295				

3.5 DISCUSSION

The primary goals of this study were to determine firstly whether the latency to recover behaviourally from an acutely stressful event commonly employed in behavioural experiments – lifting partitions - could explain variation in contest behaviour and outcome. Secondly, we wanted to test whether this latency was related to endocrine measures of physiological stress obtained from water-borne cortisol assays. Our first prediction was that fish more rapidly resuming normal swimming behaviour following removal of a partition in a dyadic behavioural trial would tend to initiate and win contests. These relationships between behavioural traits were not supported by our data, suggesting that a proactive coping style is associated with readjusting to experimental protocol disturbances; however, it is not associated with initiating or winning contests. Although many studies on fish have found a strong positive association between initiating and winning contests (e.g. Jackson, 1991; Hsu et al., 2009), our data suggest that we should be careful in assuming this pattern will always hold.

Both the correlation analysis and the linear models, where potentially confounding effects of body size could be statistically accounted for (Scott and Ellis, 2007), revealed some associations between behavioural and endocrine traits. However, these associations were not consistent with our *a priori* predictions. For example, we predicted that behavioural recovery following a partition being lifted would be faster for fishes with lower baseline (pre-contest) cortisol levels; however, the reverse pattern was seen. While this effect was non-significant, pre-contest cortisol level was significantly and positively associated with tendency to initiate contests. Pre-contest cortisol level did not predict contest outcome, and there was no significant effect of status on stress response. Although SR was lower in winners as we predicted the effect size was small and non-significant.

Overall our results do not fit comfortably into the proactive-reactive framework that has been used to interpret suites of correlated traits as reported in mammalian, avian, and other fish systems (Koolhaas et al., 1999; Øverli et al., 2007; Carere et al., 2010). Some recent studies provide evidence consistent with this framework, testing the hypothesis that differences in behaviour are associated with differences in stress response (Øverli et al., 2002; Øverli et al., 2005; Øverli et al., 2007). These studies found that those individuals more rapidly resuming normal behaviour in novel

environments or following acute stress were socially more dominant and in addition, had lower baseline cortisol levels and stress - responsive cortisol levels than those taking longer to resume normal behaviour. Thus, individuals have been argued to lie along a continuum of coping styles ranging from proactive to reactive, respectively. It should be noted that these fish studies were carried out using lines of domestic rainbow trout (*Oncorhynchus mykiss*) specifically selected for divergent cortisol responses; however, more recent work focussing on variation within populations has reached similar conclusions in a range of wild and domestic fish species (see Conrad et al., 2011 for a comprehensive review).

The swordtails used for our study were captive bred and, although they had wild-type colours, have an unknown history of artificial selection under conditions of high resource availability with environmental stressors likely to differ substantially from those of wild fish. We certainly acknowledge that relaxed natural selection in captivity might result in increased phenotypic variance and/or behaviour-physiology correlations that are either unexpected or that would be maladaptive in the wild (e.g. Lee and Berejikian, 2008; Conrad and Sih, 2009). We also acknowledge that our sample size was relatively small, thus limiting statistical power, and that control experiments to examine physiological responses to barrier removal without a subsequent dyadic contest would be useful. Nevertheless, it is equally true that other studies conducted under both laboratory and field conditions have reported deviations from the expected trait correlation structure between proactive – reactive coping style extremes, suggesting that the categorization is too simplistic (Brelin et al., 2008; Archard and Braithwaite, 2011; Vaz-Serrano et al., 2011; Archard et al., 2012). Environmental context can dissolve or generate trait correlations (e.g, Bell and Sih, 2007), and even completely reverse relationships between behaviour and physiology (Ruiz-Gomez et al. 2008). These studies suggest considerable plasticity in trait associations and the involvement of multiple, perhaps independently operating mechanisms that shape associations between behaviour and endocrine state.

Evidence from studies of behaviour in male tree lizard morphs (Thaker et al., 2009) suggests that animals with elevated cortisol levels are more prepared for an immediate response to predators. Koolhaas et al., (1997) suggested that elevations of glucocorticoids at appropriate times can be adaptive, in that they prepare the animal for immediate environmental unpredictability. Speculatively, it is possible that in our study we have uncovered a similar finding: animals with already elevated cortisol

levels recover more quickly from stressors and therefore behave, at least initially, in a proactive manner. Similarly, contest winners may simply be reacting more quickly on a physiological level both to the disturbance from the experimental protocol and the attack from the proactive opponent. If this were indeed the case then a higher overall stress response for the reactive individual would seem to be appropriate.

3.5.1 CONCLUSION

Variation in endocrine traits did not match all our a priori expectations. *Post hoc* analyses revealed significant variance among-individuals that may have important functional consequences. Specifically, a *post hoc* mixed model analysis showed that after conditioning on size (LSA) and sampling point (i.e. pre- or post-trial) $\log_{10}$ transformed cortisol levels were repeatable (interclass correlation of 0.26 (± 0.13), $\chi^2_{1DF} = 6.16$, P = 0.013). This highlights the fact that there is among-individual variation (and within-individual consistency) in assayed cortisol levels, beyond that attributable to size variation). This model also confirmed the expected increase in cortisol levels with LSA (0.002 (± 0.0004), $F_{1,54DF} = 11.38$, P = 0.002), and also that average cortisol levels were higher post-trial (difference of 0.125 ((± 0.046) on the $\log_{10}$ scale, $F_{1,55DF} = 7.52$, P = 0.008) consistent with a positive physiological reaction, i.e. stress response, to the contest and/or experimental protocol. However, there was variation in SR and indeed 18 of 56 fish actually had lower cortisol release rates (i.e. SR < 0) in response to barrier removal and social challenge.

Furthermore, neither the causes nor the consequences of this among-individual variance are known at present. Such differences could emerge if individuals experience size- and status-dependent shifts in gill permeability to steroid hormones (e.g. Scott et al., 2008), i.e. a change in stress responsive release rates reflects the ability of steroids to leak across the gills for water-borne hormone measurement. Alternatively, given the inherent lag between spikes in plasma and water-borne hormones, we could be observing the confluence of status- and size-dependent differences in within-contest cortisol production. Acute elevations of stress hormone have been associated with increased aggression during social interactions (Kruk et al., 2004; Earley et al., 2006). Although we do not know if the association between acute stress responses, aggression, and social dominance is size-dependent, it is possible that large winners mounted a stronger within-contest stress response than is detectable in the water-borne sample.

CHAPTER 4

HOW INTEGRATED ARE BEHAVIOURAL AND ENDOCRINE STRESS RESPONSE TRAITS? A REPEATED MEASURES APPROACH TESTING THE COPING STYLE MODEL

4.1 ABSTRACT

It is widely expected that physiology and behaviour will be integrated within divergent stress coping styles (SCS) that may represent opposite ends of a continuously varying reactive-proactive axis. If such a model is valid, then stress response traits should be repeatable and physiological and behavioural responses should also change in an integrated manner along a major axis of among-individual variation. While there is some evidence of association between endocrine and behavioural stress response traits, few studies incorporate repeated observations of both. To test this model we use a multivariate, repeated measures approach in a captive bred population of *Xiphophorus birchmanni*, quantifying variation among individuals in behavioural stress responses and measuring water borne steroid hormone levels (cortisol, 11-ketotestosterone). Under a mild stress stimulus, (multivariate) behavioural variation among individuals was consistent with a strong axis of personality (shy-bold) or coping style (reactive-proactive) variation. However, behavioural responses to a moderate stressor were less repeatable and robust statistical support for repeatable endocrine state over the full sampling period was limited to 11-ketotestosterone. Although *post hoc* analysis suggested cortisol expression was repeatable within individuals over short time periods, qualitative relationships between behaviour and glucocorticoid levels were counter to our *a priori* expectations. Thus, while our results clearly show among-individual differences in behavioural and endocrine traits associated with stress response, the correlation structure between these is not consistent with a simple proactive-reactive axis of integrated stress coping style.

4.2 INTRODUCTION

When challenged by adverse environments, animals use behavioural and physiological components of the stress response to maintain homeostasis (Selye, 1973; Johnson et al., 1992; Chrousos, 1998). Stress response may vary among individuals within a population (Huntingford, 1976; Verbeek et al., 1996; DeVries, 2002), a phenomenon that has led researchers to postulate the existence of stress coping styles (SCS) (Benus et al., 1991; Koolhaas et al., 1997; Koolhaas et al., 1999; Korte et al., 2005). It is widely expected that behaviour and physiology will be integrated within divergent SCS typically characterised as being either *proactive* or *reactive* (Koolhaas et al., 1997). Proactive individuals actively challenge stressors and present behavioural profiles consistent with bold personalities (e.g. Brown et al., 2007; Thomson et al., 2011), rapidly develop rigid routines and usually have low hypothalamic-pituitary-axis (HPA) activity. In contrast, reactive individuals demonstrate low levels of aggression and appear to be more flexible in their behavioural responses, tending toward raised HPA activity (e.g. Øverli et al., 2007; Carere et al., 2010). Although often presented as dichotomous, proactive and reactive coping styles may actually represent opposite ends of a continuously varying axis of SCS. If the SCS model is valid, then stress response traits should not only be repeatable, but physiological and behavioural responses also ought to change in an integrated manner along a major axis of among-individual variation (Wechsler, 1995). Here, using a freshwater fish population, we investigate among-individual variation in behavioural and endocrine stress response traits to test these predictions and thus evaluate the SCS model.

In general, studies of vertebrate stress responses have focused primarily on endocrine physiology. Despite this, comparatively few studies to date have directly tested for repeatable, among-individual variance in stress related endocrine traits (but see e.g. Andrade et al., 2001; Ferrari et al., 2013). Nonetheless, genetic studies have provided evidence of heritable variation for endocrine response to stress in many taxa (e.g. Silberg et al., 1999; Evans et al., 2007), and a trait cannot be heritable without being repeatable. In fishes, genetic variation for plasma cortisol (F) levels has been demonstrated widely (e.g. Pickering and Pottinger, 1989; Fevolden et al., 1993; Barton, 2002; Pottinger, 2010). Artificial selection on rainbow trout (*Oncorhynchus mykiss*) has successfully generated high and low post-stress cortisol lines (Pottinger and Carrick, 1999), while quantitative trait loci (QTL) for endocrine stress response traits have been mapped in several aquaculture species (Massault et al., 2010; Boulton et al., 2011).

Even though endocrine processes may be important for coping with acute stress challenges, it should also be recognised that behavioural responses such as freezing, fighting or fleeing may be more critical in some contexts (e.g. response to predation attempt) (Blanchard et al., 1998). There is evidence for alternative behavioural stress response profiles in rodents (Benus et al., 1991; Sgoifo et al., 1998; Koolhaas et al., 1999; Veenema, 2009) , birds (e.g. Carere et al., 2003; Fraisse and Cockrem, 2006), and livestock (Hessing et al., 1994). In many cases associations between behaviour and HPA activity have been found, consistent with SCS (e.g. Sutherland and Huddart, 2012; Wesley et al., 2012). More generally, empirical studies in the burgeoning field of animal personality (Sih et al., 2004a; Réale et al., 2007) have emphasised that among-individual variation in behaviour is taxonomically widespread. This is certainly true for behaviours associated with stress exposure (e.g. Wilson, 1998; Gosling and John, 1999; Briffa et al., 2008; Rudin and Briffa, 2012), leading some authors to argue that SCS and personality can sometimes be synonymous (Øverli et al., 2007).

Along a reactive-proactive axis of SCS, behaviour is expected to change in a manner broadly corresponding to the axis of shyness-boldness described in the personality literature (Wilson et al., 1994; e.g. Budaev, 1997a; Winberg et al., 2007; Huntingford et al., 2010; Raoult et al., 2012). Empirical studies demonstrating variation in boldness have been conducted in many taxa including fishes (e.g. Budaev et al., 1999; Bell et al., 2009). While there is some evidence of association between endocrine and behavioural stress response traits in a range of taxa (e.g. Andrade et al., 2001; Creel, 2001; Thaker et al., 2009; Archard et al., 2012), few studies have incorporated repeated observations on both traits (but see Ellis et al., 2004; Sebire et al., 2007; Ferrari et al., 2013). This is an important limitation because repeated measures are required to partition the among-individual differences expected under the SCS model from sources of within-individual (i.e. observation specific) variation (Dingemanse et al., 2010; Dochtermann and Roff, 2010; Dingemanse and Dochtermann, 2013). Therefore two key questions remain largely unanswered. Firstly, to what extent are endocrine stress responses a repeatable phenotype of the individual? Secondly, assuming that correlations between behavioural and endocrine stress responses are apparent, to what extent are they actually driven by among-individual differences in SCS?

Here we aim to address these questions using a small tropical freshwater fish, *Xiphophorus birchmanni*. In this species we have previously demonstrated a strong axis

of among individual variation in boldness that is stable over long periods, i.e. representative of expected life span (Boulton et al., 2014). We now expand on this previous work to ask whether there is also among-individual variation in endocrine physiology, and whether behavioural and endocrine responses to a stressor are integrated within SCS. To investigate behavioural response we subject fish to a modified open field trial (OFT, a mildly stressful novel situation) coupled with a simulated predator attack to provide a moderate acute stress stimulus. To investigate endocrine state we quantify cortisol (F), the principal and most frequently measured glucocorticoid in fishes released by activation of the hypothalamic-pituitary-interrenal (HPI) axis on exposure to stressors (Mommsen et al). In addition we quantify 11-ketotestosterone (11KT), an important androgen in teleosts (Mayer et al., 1990; Mommsen et al., 1999). Although not normally considered a stress hormone *per se*, many studies point toward a link between gonadal steroids and personality traits such as aggression and boldness (Pellis and McKenna, 1992; Borg and Mayer, 1995; Oliveira et al., 2002; Taves et al., 2009; Koolhaas et al., 2010). Here we seek to test three specific predictions: 1) that fish exposed to stressors differ consistently in behavioural responses thus aligning with expectations under a shy-bold personality paradigm; 2) that there is repeatable variation for pre-stressor endocrine state and/or change in hormone levels following stress exposure; 3) that behavioural and endocrine stress response traits (co)vary and correlation exists at the among-individual level as predicted by the SCS model.

4.3 METHODS

4.3.1 ANIMAL HUSBANDRY

Twenty male *Xiphophorus birchmanni* were selected at random from a tank containing second-generation captive-bred fish. Animals were of unknown age but of similar size (1.16 ± 0.073 g) and developmental stage. All were sexually mature based on external assessment of gonopodium development. Fish were housed individually in half sections of ten 30 L (37 x 37 x 22cm) tanks, divided by opaque, water permeable dividers constructed from Perspex frames covered with dark-coloured fine-gauge nylon net. Ten half-tanks were contained within a stack sharing a common recirculating water supply; thus, within a stack fish were physically and visually, though not chemically isolated. Fish were maintained at 21 – 23^{0}C on a 12:12 light:dark cycle. Fish were fed twice per day, using a mix comprising equal quantities of crushed spirulina (ZM systems, UK: http://www.zmsystems.co.uk/) and brine shrimp

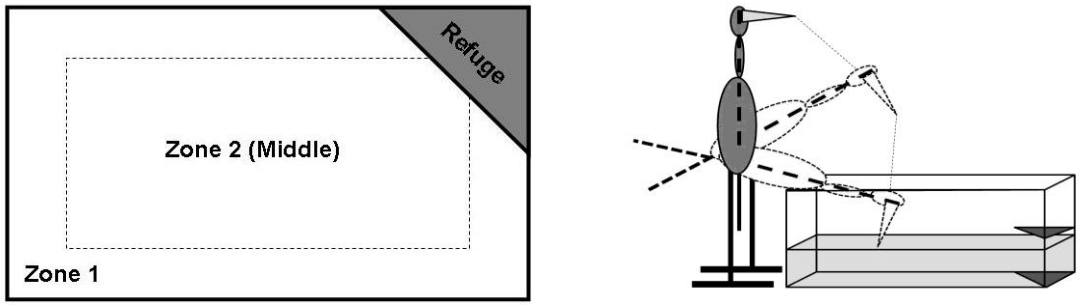
flake in the morning followed by a previously frozen mixture of bloodworm, brine shrimp nauplii and daphnia in the late afternoon. Fish were not fed on the morning of days when they underwent trials.

4.3.2 BEHAVIOURAL TRIALS:

Following collection of a pre-trial water sample for hormone assay (see below), each fish was placed in an empty 45 x 25 x 25 cm glass tank filled to a depth of 8 cm with nine litres of clean water. The tank was positioned on an illuminated light box, increasing contrast to allow data extraction using video-based tracking software. A small refuge was created in the tank by attaching a triangular piece of aquarium filter foam (10x10x14cm) just above the water in one corner (Figure 4.1a). An equally sized piece of card was placed below the tank in the same corner. Thus when within the refuge the fish was not visible from above, and was shielded from light coming from below. A Sunkwang C160 video camera fitted with a 5 – 50 mm manual focus lens was suspended above the apparatus.

Figure 4.1 Setup of experimental arena for behavioural trials showing a) an overhead view, and b) the position of the decoy heron used to simulate an avian predation event. Zones 1 and 2 are defined for scoring by tracking software only and were of equal area. The refuge comprises a triangle of aquarium filter foam taped just above the water level to give the impression of a bank to hide under. A piece of card (of similar size and shape) was also placed under this corner of the tank. The decoy heron was positioned so as not to cast a shadow over the arena, its downward swing constrained to stop with the beak at water level.

a) b)



Following introduction to the tank, each fish was allowed 300 sec to acclimate to the experimental arena then behaviour was recorded for 120 sec on video (described below). Note that being placed in a novel environment is considered to be a mild stress stimulus in small fishes (Burns, 2008). A further (moderate) acute stress exposure was then imposed, using a decoy heron on a swinging stand to simulate an avian predation

event (Barber et al., 2004) (Figure 4.1b). Members of the Ardeidae family are known to predate the river where the parental generation of fish were captured (Arroyo Coacuilco near the town of Coacuilco, municipality of San Felipe Orizatlán, Hidalgo, Mexico; GG Rosenthal, personal communication). The decoy was positioned in such a way that it did not create a shadow over the arena in the upright position. When released, the decoy swung down rapidly towards the tank. The swing was limited to stop the decoy abruptly (with the beak at water level), causing a loud percussive sound and vibration that disturbed the tank. A further 120 sec of behaviour was recorded before the fish was removed for collection of the post-trial water sample. Water in the experimental tank was replaced prior to the next trial. The entire sampling process was repeated five times at four day intervals. All fish were sampled on each occasion (in variable order, to avoid confounding any diurnal effects with individual identity) with the exception of one individual that died between the fourth and fifth trials. Two 119 litre glass tanks (122 x 38 x 30 cm) were used to store water at room temperature to supply the behaviour trials and hormone collection beakers (see below).

4.3.2.1 Behavioural traits

Data were extracted from videos using tacking software from Biobserve (http://www.biobserve.com/products/viewer/index.html). Specifically, for the 120 second period before the heron strike we measured: Track length (TL, total distance moved in cm); percentage of time spent Active (ACT); percentage of tank basal Area Covered (AC); Time in Middle of tank (TIM, in sec, Figure 4.1). Using a slightly different experimental arena (with no refuge) we have recently shown these traits to be repeatable in *X. birchmanni*, and characterised by among-individual covariance structure consistent with a major axis of boldness variation (Boulton et al., 2014). In addition, we recorded time spent out of the refuge (TOR), our *a priori* expectation being that this would be consistently higher in bold individuals. Based on pilot data, we had expected all fish to respond to the acute stressor (simulated predation event) by immediately entering the refuge and indeed this was observed in all but two trials. However, while we had planned to use a continuous measure of latency to re-emerge as a further metric of behavioural stress response, in approximately two thirds of trials the fish did not re-emerge within the subsequent two minute observation period. Due to this data censoring we used Emergence from the refuge (emREF) as a binary behavioural response to the acute stressor (1 the fish re-emerged, 0 it did not).

4.3.3 ENDOCRINE ASSAYS

We used a non-invasive method to assess individual endocrine state from holding water samples (Ellis et al., 2004). This allows repeated sampling of small fish that would not survive invasive collection of blood plasma for assay. Water samples were collected pre- and post- behavioural trial as follows. First, fish in home tanks were captured using non-PET plastic beaker inserts made by cutting the neck from cylindrical 500 ml opaque Nalgene bottles and drilling drainage-holes into the base, (following Archard and Braithwaite, 2011). The insert was then gently lifted from the tank, allowing water to drain, before being placed in a glass beaker containing 500ml clean water. Capture and handling time, i.e. transfer to beaker of clean water, was not recorded, but was estimated to take no longer than 60 sec in each case. The beaker was covered with a dark net and left for 60 min to obtain the pre-trial sample. The insert was then used to gently transfer the fish to the behavioural trial arena tank. After the behavioural trial a clean insert was used to transfer the fish to a second beaker of 500ml water for a further 60 min period to collect the post-trial sample. Fish were then placed onto a dry paper towel and weighed (to the nearest 0.01 g) before being returned to home tanks. Nitrile gloves were worn throughout all procedures requiring contact with fish or holding water. After use, all beakers and inserts were rinsed thoroughly with distilled water then ethanol and allowed to dry overnight.

4.3.3.1 Solid phase extraction

Each 500 ml water sample was filtered to remove any debris (Whatman Filter paper, grade 1, 24cm) and steroids were extracted to C18 solid phase columns (SepPak® Vac 3 cc/500mg; Waters Inc., Milford, MA, USA) previously primed (2 x 2 mL HPLC-grade methanol followed by 2 x 2 mL distilled water). Solid phase extraction was conducted under vacuum pressure using a twenty-port manifold (Waters, as before) and Tygon tubing (Saint Gobain, Formulation 2275) to transfer samples from beaker to column. Columns were stored at –20°C until the end of the behavioural data collection, when all columns were packed in dry ice and despatched to CCMar, Universidade do Algarve, Faro, Portugal, for quantification of water borne hormone levels by radio-immunoassay (RIA). Columns were defrosted at 4°C and activated by washing with 2 x 2 ml deionized water to purge any salts. Steroids were eluted into glass tubes with ethanol (3 x 1 ml). The ethanol was evaporated at 42°C under nitrogen gas and the residue re-suspended in 1 ml RIA buffer (gelatine phosphate 0.05 M, pH 7.6).

4.3.3.2 Radioimmunoassay

RIA was used to quantify levels of free F and 11KT. For the cortisol RIA we used an antiserum raised in rabbit against cortisol-3-CMO-BSA (ref 20-CR50 Fitzgerald Industries International, Concord, USA). Cross-reactivities were 54% for 11-desoxycortisol, 10% for cortisone, 16% for 17,21-dihydroxy-5β-pregnan-3,11,20-trione, 5% for 11β,17,21-trihydroxy-5β-pregnan-3,20-dione, 0.05% for 11β-hydroxytestosterone and less than 0.001% for testosterone. The 11-ketotestosterone antiserum cross-reactivities are given elsewhere (Kime and Manning, 1982). To verify the specificity of the RIAs towards the samples, a pool of water extracts was first separated by normal phase thin-layer chromatography and fractions assayed for the two steroids. The two RIAs were shown to be highly specific, only cross-reacting with single fraction co-migrating, respectively, with F and 11KT. Inter-and intra-assay variability for the two assays was below 12%.

4.3.3.3 Validation of water borne steroid assays

That water borne steroid assays may predict plasma and/or whole body concentration has been demonstrated in a number of fish species, (e.g. Scott and Liley, 1994; Ellis et al., 2007; Sebire et al., 2007). However, the method has not previously been used in *X. birchmanni* and we therefore tested the relationship between steroid concentrations in water and whole fish. Twenty-six randomly selected stock fish of mixed sex, age and size were held separately in 500 ml glass beakers for 60 min as described above. They were then immediately euthanized by transfer to a beaker containing an MS22 solution (50 g/l) buffered with an equal quantity of sodium bicarbonate. Fish were weighed (to the nearest 0.01 g), then frozen whole at –20°C before being shipped to CCMar. Water borne samples were processed as described above. Whole fish samples were individually pulverised in liquid nitrogen with a mortar, transferred to glass extraction tubes, mixed with 5 ml absolute ethanol (Merck 1.00983.5000), vortexed for 10 min and centrifuged. The supernatant was aspirated to a second extraction tube, evaporated, and resuspended in 200 µl distilled water. Free steroids were extracted twice with 3 ml diethyl ether (VWR 23811.292), the solvent dried with nitrogen gas and the extracts resuspended in radioimmunoassay buffer. Steroid release rates (pg/hr) determined from pre- and post-trial collections and sacrificed fish were natural log (Ln) transformed for analysis.

4.3.4 STATISTICAL ANALYSIS

Data were analysed using (multivariate) linear mixed effect models parameterised by restricted maximum likelihood with the statistical package, ASReml V3, (Gilmour et al., 2009). Since this software does not readily accommodate non-Gaussian traits we analysed the binary behavioural response trait emREF using a Bayesian approach implemented in MCMCglmm (Hadfield, 2010a; Hadfield, 2010b). In all models, the inclusion of fish identity as a random effect allowed the observed phenotypic (co)variance structure to be partitioned into among-individual (**I**) and residual (**R**, within-fish) between-trial components. Prior to analysis data were square root (all behaviours except emREF) or natural log transformed (endocrine traits) to meet assumptions of normality. After transformation, all data were rescaled to standard deviation units. This rescaling was done for two reasons: firstly, it simplifies the interpretation of results since the estimated among-individual variance (V_I) for any (transformed) trait corresponds to the repeatability (R; defined as the proportion of observed phenotypic variance explained by individual identity); secondly, for the inference of a latent personality trait, this prevents any single observed behaviour from dominating due to scaling effects alone (Wilson et al., 2013). For all traits we fitted fixed effects of *mean*, *trial number* (the cumulative number of trials experienced by an individual), home *stack* (a two level factor accounting for sets of fish sharing the same water supply), and *day order* (used as a proxy for time of day and modelled as a linear effect of the number of preceding trials performed that day). For endocrine traits we also included *mass* as an additional fixed effect. This allowed us to account for the expected increase in hormone release rate with size due to diffusion into the holding water across a larger gill area (Ellis et al., 2004). The covariates *day order* and *mass* were both mean-centred. For models fitted using REML the significance of fixed effects was tested by Wald *F*-tests, while likelihood ratio tests (LRT) were used to assess the significance of the random effect of fish identity. For models fitted using MCMCglmm statistical inference was based on the posterior distributions of estimated parameters.

4.3.4.1 Estimating behavioural coping style

First we modelled the set of baseline behavioural traits observed prior to the simulated predation event. This was to test our *a priori* expectation that there would be among-individual variance and covariance structure consistent with the presence of an axis of boldness variation. We initially fitted a multivariate model with no random effects, such that all variance was allocated to the residual (within-individual) component **R**,

specified as a diagonal matrix (model 1) by restraining all among-trait covariance terms to equal zero. This model was compared to a second model (model 2), where fish identity was fitted as a random effect, and the among-individual component **I** was specified as a second diagonal matrix structure. This allowed a global test (i.e. across all baseline behaviour traits) of among-individual variance by comparing models 1 and 2 with a likelihood ratio test (LRT) following Wilson et al (2010a). For comparing multivariate models in this way we conservatively assume that twice the difference in model log-likelihoods is distributed as $\chi^2_{\mathbf{n}}$, where the DF (n) is equal to the additional number of parameters to be estimated in the more complex model, in this case five. Note that for univariate model comparisons (Appendix 1, Table A1.6) we modify the test following recommendations presented by Stram and Lee (1994) and Visscher (2006). We then modelled between-trait covariance in **R** (model 3) and in both **I** and **R** (model 4), allowing us to test whether behaviours covary (model 3 vs 2) and whether among-individual differences contribute significantly to this covariance (model 4 vs 3). In model 4, **I** is therefore estimated as a fully unstructured matrix, with trait specific variance ($V_{\mathbf{I}}$) estimates on the diagonal (equal to the trait repeatabilities) and the among-individual covariance ($COV_{Ix,y}$) between each pair of traits (x,y) off the diagonal. Among-individual correlations (r_I) were then calculated by rescaling the among-individual covariance (COV_I) so that $r_{x,y} = COV_{I(x,y)} / \sqrt{(V_{Ix} * V_{Iy})}$.

Eigenvector (EV) decomposition was then used to evaluate whether **I** among this set of traits (as estimated under model 4) was dominated by a single major axis interpretable as boldness. Specifically, based on previous findings in an independent data set (Boulton et al., 2014) we predicted that the first eigenvector of **I** ($EV1_{\mathbf{I}}$) would capture most of the among-individual behavioural variance and would be characterised by trait-specific loadings of equal sign and similar magnitude. We used parametric bootstrapping (Boulton et al., 2014) to simulate 5000 replicate draws of **I** from a multivariate normal distribution with means and variances defined by the REML estimate of **I** and its sampling variance-covariance matrix respectively. Each matrix was then subjected to eigen analysis and we used the 95% highest probability density (HPD) interval of parameter distributions to describe uncertainty around the trait loadings on $EV1_{\mathbf{I}}$.

We then estimated the repeatability of emREF (univariate model) and its among-individual correlations with the baseline behaviours (using bivariate models) observed prior to the predator strike using MCMCglmm (Hadfield, 2010a; Hadfield, 2010b).

Emergence was treated as a categorical trait with residual variance fixed at 1. All (transformed) OFT traits were treated as Gaussian. MCMCglmm models were run for 1050000 iterations with a burnin of 50000 iterations and a thinning interval of 1000 iterations. The repeatability of emREF on the liability scale was determined as the intraclass correlation, calculated as $V_I / (V_I + V_R + \pi^{2/3})$, where V_I is the among-individual variance and V_R is the residual variance (i.e. 1) (Hadfield, 2010b).

4.3.4.2 (Co)variance structure between endocrine traits and behaviour

To validate the assumption that water borne steroid levels were representative of biological processes, we first estimated the correlations between the water borne and entire body levels of cortisol (F) and 11KT from the sacrificed fish (n = 26). Correlations were estimated between natural log transformed rates of hormone release scaled by mass. Following this, to characterise patterns of variance and covariance in endocrine traits, mixed model analyses similar to those described above were applied to the (natural log transformed) endocrine traits collected across the five trials, expressed in standard deviation units. For these analyses, rather than dividing by mass, we included mass as an additional fixed effect for all endocrine traits. Thus, we tested for repeatable variation in pre- ($_{PRE}$) and the post-stressor ($_{POST}$) hormone levels of F and 11KT, estimated the covariance structure between these endocrine traits and partitioned it into within- and among-individual components as for the behavioural traits above.

To test the primary hypothesis predicted by the SCS paradigm, that among-individual differences in behaviour are correlated with among-individual differences in endocrine physiology, we then fitted additional multivariate models to estimate the among-individual correlation (r_I) between endocrine and behavioural traits (ACT, emREF). Note that Activity (ACT, percentage time active) was used here as a univariate proxy for baseline behavioural variation based on the eigen decomposition of the **I** matrix between behaviours (see results below for details).

Table 4. 1 Mixed model comparisons to test for among-individual variance, between-trait covariance, and among-individual covariance (between traits) in a) the set of baseline behaviours, b) endocrine traits measured pre-and post-stressor and c) activity and pre-stressor endocrine state. Models 1-4 were fitted to each set of traits to partition observed (co)variance into residual (within-individual, **R**) and among-individual (**I**) components. **R** and **I** were modelled as either diagonal (DIAG) or unstructured (US) matrices (except in model 1 where **I** was not fitted at all; NF). Shown are the log-likelihoods (LogL) and statistical comparisons to preceding model using likelihood ratio tests.

Set of traits	**Model**	**Covariance structure**			**Comparison to previous model**			**Effect being tested**
		R	**I**	LogL	χ^2	DF	P	
a) Baseline behaviours	1	DIAG	NF	-290.5				
	2	DIAG	DIAG	-274.0	32.9	5	<0.001	Among-individual variance
	3	US	DIAG	151.7	851	10	<0.001	Between-trait covariance
	4	US	US	163.0	22.6	10	0.013	Among-individual covariance
b) Endocrine traits (pre- and post-stressor)	1	DIAG	NF	-195.3				
	2	DIAG	DIAG	-190.5	9.57	4	0.048	Among-individual variance
	3	US	DIAG	-179.7	21.6	6	0.001	Between-trait covariance
	4	US	US	-176.8	5.83	6	0.433	Among-individual covariance
c) Activity & pre-stressor endocrine traits	1	DIAG	NF	-146.9				
	2	DIAG	DIAG	-138.2	17.3	3	<0.001	Among-individual variance
	3	US	DIAG	-137.8	0.806	3	0.855	Between-trait covariance
	4	US	US	-134.3	6.98	3	0.073	Among-individual covariance

4.4 RESULTS

4.4.1 AMONG-INDIVIDUAL VARIANCE IN BEHAVIOUR

Across the full set of baseline behaviour traits there was evidence for significant among-individual variance (comparison of models 1 & 2, $\chi^2_5 = 32.9$, $P < 0.001$), as well as covariance structure between traits (model 2 vs. 3, $\chi^2_{10} = 851.4$, $P < 0.001$) that included an among-individual component (model 3 vs. 4, $\chi^2_{10} = 22.6$, $P = 0.013$). Thus we conclude that these behavioural traits are repeatable and covary among-individuals (Table 4.1). From model 4, repeatabilities (± SE) for baseline behaviours ranged from 0.101 (± 0.105) for Time in Middle to 0.305 (± 0.153) for Activity (Table 4.2). Univariate analyses, assuming the test statistic to be asymptotically distributed as a mix of $50{:}50\chi^2_0$ and χ^2_1 (following Visscher, 2006), were statistically significant at $P<0.05$ for all individual traits except Time in Middle (see Appendix 1, Table A1.6). Though not directly relevant to the present objectives, fixed effects estimated from these univariate models are also presented for completeness (see Appendix 1, Table A1.7).

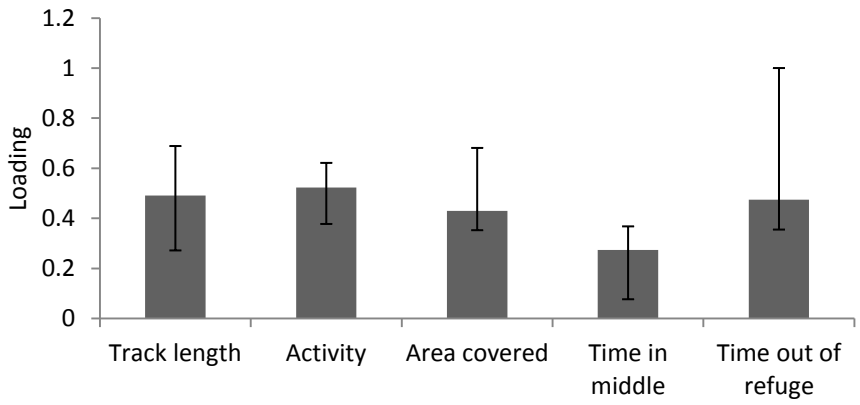
Table 4.2 Estimated **R** (residual, within-individual) and **I** (among-individual) matrices for the set of baseline behavioural traits: Track-length (TL); Activity (ACT); Area Covered (AC); Time in Middle (TIM); Time Out of Refuge (TOR). Trait specific variances are shown on the diagonal (shaded), with between-trait covariances (below diagonal) and correlations (above diagonal). Variances on the diagonal of **I** can be interpreted as repeatabilities since (transformed) traits were scaled to standard deviation units. Standard errors are provided in parentheses.

R	TL	ACT	AC	TIM	TOR
TL	0.722 (0.118)	0.984 (0.004)	0.913 (0.02)	0.632 (0.070)	0.942 (0.014)
ACT	0.696 (0.115)	0.695 (0.114)	0.901 (0.022)	0.663 (0.065)	0.961 (0.009)
AC	0.680 (0.116)	0.658 (0.113)	0.769 (0.125)	0.801 (0.042)	0.881 (0.026)
TIM	0.502 (0.107)	0.516 (0.107)	0.656 (0.120)	0.872 (0.141)	0.672 (0.064)
TOR	0.681 (0.114)	0.682 (0.113)	0.658 (0.114)	0.534 (0.109)	0.726 (0.118)
I	TL	ACT	AC	TIM	TOR
TL	0.274 (0.145)	0.986 (0.011)	0.975 (0.034)	0.838 (0.249)	0.959 (0.034)
ACT	0.285 (0.148)	0.305 (0.153)	0.957 (0.046)	0.902 (0.223)	0.992 (0.013)
AC	0.237 (0.134)	0.246 (0.136)	0.217 (0.131)	0.855 (0.184)	0.931 (0.064)
TIM	0.140 (0.106)	0.158 (0.111)	0.127 (0.106)	0.101 (0.105)	0.927 (0.205)
TOR	0.253 (0.139)	0.277 (0.145)	0.219 (0.130)	0.149 (0.108)	0.256 (0.141)

Between baseline traits, the among-individual correlations (r_I) were positive and strong, ranging from 0.838 (± 0.249) between Track-length and Time in Middle, to

0.986 (± 0.011) between Track-length and Activity (Table 4.2). Consistent with this correlation structure, we found that 96.2% of the variance in **I** was explained by the first eigenvector of **I** (Figure 4.2, Appendix 1, Table A1.8). Trait loadings on this vector are all positive and broadly similar in magnitude (bootstrapped 95% confidence intervals overlap for all traits (Figure 4.2)), commensurate with our *a priori* expectations of boldness. This result provides independent experimental confirmation of our previous finding that a strong axis of boldness variation exists in this population (Boulton et al., 2014). Based on the confidence intervals we conclude that trait loadings do not differ significantly from each other, but are greater than zero (Figure 4.2).

Figure 4.2 Loadings (in Standard Deviation units) on the first eigen vector of **I**, representing 96.2% of the total estimated variance for the baseline behaviour traits. Error bars indicate 95% highest probability density intervals estimated by parametric bootstrap (see text for details).



Statistical support for among-individual variance in tendency to emerge after the acute stressor (predator strike) was less compelling. Using MCMCglmm the estimated repeatability for emREF (on the liability scale) was moderately high (intraclass correlation (IC) = 0.406, 95% higher probability density (HPD) 0.074 – 0.790). Note however that the IC estimate is constrained to be positive (i.e. the HPD interval cannot span zero) and the posterior mode of IC was not clearly distinct from zero (Appendix 2, Figure A2.2). For comparison, we estimated a repeatability (± SE) for emREF of the observed scale of 0.160 (± 0.107) using REML. Although nominally significant (P = 0.04; see Appendix 1, Table A1.6), the likelihood ratio test applied assumes residual normality that is clearly not the case for this binary trait. MCMCglmm estimates of $r_{\mathbf{I}}$ (95% CI) between emREF and baseline behaviours were all positive but not statistically significant, ranging from 0.172 (-0.479 – 0.830) for Track-length to 0.508 (-0.452 –

0.839) for Area Covered (Table 4.3). We therefore interpret variation in emREF cautiously. Some variance among individuals in response to the acute stressor may be present. If so, individuals more likely to re-emerge following the simulated predator strike tend to be the bolder fish, as indicated by baseline behaviours. However, this qualitative pattern is not statistically robust in our data.

Table 4.3 MCMCglmm estimates of intraclass correlations (r_I) between pre-strike behaviours and post-strike Emergence, with 95% upper and lower higher probability density values.

		95% HPD interval	
Emergence with:	r_I	lower	upper
Track-length	0.172	-0.479	0.830
Activity	0.508	-0.452	0.839
Area Covered	0.337	-0.421	0.930
Time in Middle	0.279	-0.639	0.962
Time Out of Refuge	0.214	-0.599	0.827

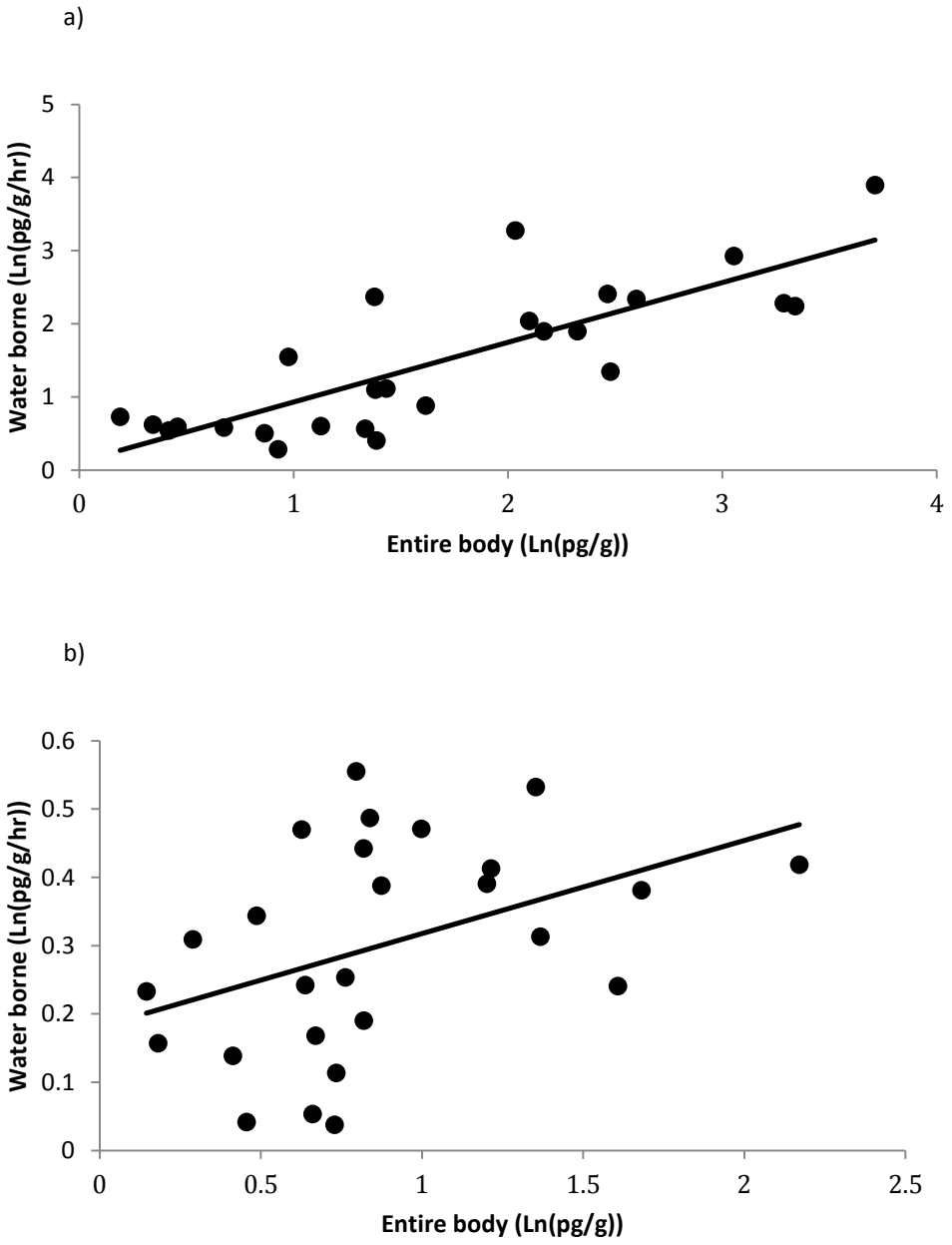
4.4.2 AMONG-INDIVIDUAL VARIANCE IN ENDOCRINE TRAITS

Our validation sample confirmed significant positive correlations between mass-adjusted water borne release rate and whole body hormone concentrations. For cortisol the relationship was strong (r = 0.815, ± 0.067, P<0.001) and linear on a (natural) log-log scale (Figure 4.3a). For 11KT the relationship was weak, but nonetheless positive and significantly greater than zero (r = 0.420 ± 0.165, P = 0.028; Figure 4.3b). Thus we consider water borne endocrine levels to be an appropriate proxy for whole body measures in this species. In our experimental samples, absolute cortisol release rates were actually higher in the pre- than post-stressor collection periods (mean F_{PRE} (SE) = 1871 (± 176) pg/hr, mean F_{POST} (SE) = 669 (± 64.9) pg/hr). Comparison of paired samples confirmed that individuals released significantly less cortisol in the post-trial collection period (paired sample t-test, t_{98} = 7.17, P<0.001). There was no evidence for a difference in 11KT levels between pre- and post-sampling periods (pre- mean (SE) = 105.56 (± 4.21) pg/hr, post-mean (SE) = 99.69 (± 3.63) pg/hr, paired sample t-test, t_{96} = 1.169, P = 0.123).

Multivariate models provided evidence of among-individual variance in endocrine phenotype (comparison of models 1 & 2, χ^2_4 = 9.57, P = 0.048). Covariance between traits was also present (model 2 vs. 3, χ^2_6 = 21.6, P = 0.001), although an among-individual component to this was not statistically supported (model 3 vs. 4, χ^2_6 = 5.83, P = 0.443), (Table 2.1b). Under the full model (4), repeatabilities (SE) varied from 0.039

(± 0.087) for F_{POST} to 0.202 (± 0.113) for $11KT_{PRE}$ (Table 4.4). Univariate models yielded similar repeatability estimates (Appendix 1, Table A1.6), and revealed significant effects of *day order* and *Trial* on endocrine state (Appendix 1, Table A1.7). However, V_I was only statistically significant for $11KT_{PRE}$. Thus we conclude that robustly supported among-individual variance in endocrine state is limited to $11KT_{PRE}$, although we note that the estimate of V_I for F_{PRE} was marginally non-significant in the univariate analysis.

Figure 4.3 Relationships between water borne and entire body levels of a) cortisol and b) 11-ketotestosterone. Solid lines show ordinary least squares regressions.



Examination of **I** and **R** matrices between endocrine traits (Table 4.4), showed that the significant covariance structure detected was likely driven by a single positive relationship between F_{POST} and $11KT_{POST}$. 90% of the covariance between these traits was partitioned into **R**, yielding a within-individual correlation r_R (SE) of 0.356 (± 0.101). Given no evidence of significant covariance structure in **I** we do not further consider pairwise estimates of r_I except to note that the estimate between F_{PRE} and $11KT_{PRE}$ was strongly positive and approaching significance (r_I = 0.768 (± 0.389)). Thus to the extent that F_{PRE} is actually repeatable, individuals with higher release rates are also characterised by higher 11KT levels, not lower as we expected *a priori*.

Table 4.4 Estimated **R** (residual, within-individual) and **I** (among-individual) matrices for release rates of cortisol (F) and 11-Ketotestosterone (11KT) pre- and post-behavioural stressor trial. Trait specific variances are shown on the diagonal (shaded), with between-trait covariances (below diagonal) and correlations (above diagonal). Variances on the diagonal of **I** can be interpreted as repeatabilities since (transformed) traits were scaled to standard deviation units. Standard errors are provided in parentheses.

R	F_{PRE}	$11KT_{PRE}$	F_{POST}	$11KT_{POST}$
F_{PRE}	0.594 (0.097)	0.051 (0.116)	0.066 (0.115)	-0.205 (0.111)
$11KT_{PRE}$	0.030 (0.069)	0.589 (0.097)	0.104 (0.115)	0.083 (0.115)
F_{POST}	0.049 (0.085)	0.076 (0.085)	0.903 (0.147)	0.356 (0.101)
$11KT_{POST}$	-0.138 (0.080)	0.056 (0.078)	0.296 (0.102)	0.766 (0.124)
I	F_{PRE}	$11KT_{PRE}$	F_{POST}	$11KT_{POST}$
F_{PRE}	0.091 (0.077)	0.768 (0.389)	0.854 (1.102)	0.881 (1.284)
$11KT_{PRE}$	0.104 (0.071)	0.202 (0.113)	0.552 (0.807)	0.867 (0.872)
F_{POST}	0.051 (0.059)	0.049 (0.072)	0.039 (0.087)	0.815 (1.210)
$11KT_{POST}$	0.054 (0.056)	0.078 (0.071)	0.033 (0.064)	0.041 (0.081)

4.4.3 Correlation structure between activity, Cortisol and 11KT

To test among-individual correlation between boldness and endocrine state we fitted trivariate models of activity (ACT), F_{PRE} and $11KT_{PRE}$. Using a univariate proxy for boldness is appropriate given the strong correlation structure in **I** among baseline behaviours (see above, Table 4.2 and Figure 4.2). Such an approach also avoids the issue of carrying forward uncertainty surrounding principal component estimates. ACT was chosen since it has the highest loading (with the narrowest confidence interval) on the estimated vector of boldness. F_{POST} and $11KT_{POST}$ were not included in these multivariate models given the lack of repeatable variation for these traits. Model

comparisons (Table 4.1c) confirmed among-individual variance (model 1 vs. 2, χ^2_3 = 17.3, P <0.001); however, the model was not significantly improved by inclusion of between-trait covariance in **R** or **I** (model 2 vs. 3, χ^2_3 = 0.086, P = 0.848; model 3 vs. 4, χ^2_3 = 6.98, P = 0.073) (Table 4.1). Under Model 4, estimated repeatabilities were similar to those already reported (Table 4.5). While reiterating that our model comparisons indicate non-significant between-trait covariance structure (within- and among-individuals), it is perhaps worth noting that r_I estimates are positive, and strong in some cases (e.g. Table 4.5). Thus the qualitative result is that, counter to our expectations, individuals characterised by higher (pre-stressor) release rates of F and 11KT are the bolder individuals as measured by ACT.

Table 4.5 Estimated **R** (residual, within-individual) and **I** (among-individual) matrices between pre-trial cortisol and 11-ketotestosterone (F_{PRE}, $11KT_{PRE}$) and activity (ACT). Trait specific variances are shown on the diagonal (shaded), with between-trait covariances (below diagonal) and correlations (above diagonal). Variances on the diagonal of **I** can be interpreted as repeatabilities since (transformed) traits were scaled to standard deviation units. Standard errors are provided in parentheses.

R	F_{PRE}	$11KT_{PRE}$	ACT
F_{PRE}	0.594 (0.097)	0.056 (0.116)	-0.026 (0.116)
$11KT_{PRE}$	0.033 (0.069)	0.591 (0.098)	-0.052 (0.116)
ACT	-0.017 (0.075)	-0.034 (0.075)	0.697 (0.115)
I	F_{PRE}	$11KT_{PRE}$	ACT
F_{PRE}	0.090 (0.076)	0.743 (0.396)	0.785 (0.391)
$11KT_{PRE}$	0.099 (0.070)	0.198 (0.111)	0.383 (0.350)
ACT	0.129 (0.081)	0.093 (0.094)	0.300 (0.151)

4.5 Discussion

Overall our results provide limited support for among-individual (co)variation consistent with an integrated stress coping style (SCS) in *Xiphophorus birchmanni*. Individuals did differ consistently in their behavioural responses to mild stress imposed by the modified open field trial. Furthermore, this behavioural variation is consistent with an underlying shy-bold axis of personality. However, it is less clear that individuals differ in behavioural response to the simulated predator attack. Additionally, while there is some evidence of repeatable variation in endocrine state, robust statistical support was limited to pre-trial 11KT levels. Though not statistically significant, there was a tendency for bolder, or more behaviourally proactive,

individuals to release more cortisol. Although potentially indicative of some degree of integration between behavioural and endocrine stress response components, this pattern is actually counter to the SCS model's prediction of lower HPA activity in proactive individuals (Koolhaas et al., 1999). In what follows we discuss first the behavioural, and then the endocrine data in more detail before commenting further on the relationship between the two. In addition to presenting our biological conclusions we also highlight a number of methodological issues and difficulties of interpretation that warrant further consideration.

4.5.1 BEHAVIOURAL RESPONSE

We found partial support for our first hypotheses that fish would differ consistently in behavioural response to stress exposure. Analysis of behavioural data collected under the mild stress stimulus showed that individual traits assayed were repeatable, and the **I** matrix contained significant among-individual correlation structure consistent with a single latent axis, or personality trait, underpinning the observed variation. Moving along this axis, hereafter interpreted as shyness-boldness, trait expression changes in a concerted manner. Thus a fish that consistently swims further is also more active, explores a greater area, spends more time in the centre of the experimental arena, and spends less time hiding in the refuge. This finding confirms our earlier report of a strong axis of boldness variation in *Xiphophorus birchmanni* that is broadly stable over long time periods (i.e. representative of expected lifespan under natural conditions; Boulton et al., 2014), and adds to rapidly accumulating evidence of personality variation in fishes (Burns, 2008; Toms et al., 2010; Wilson et al., 2013). However, we note that our data do not clearly support the expectation that boldness (as inferred from the baseline data) leads to faster re-emergence following the moderately stressful simulated predation event. To some extent this could reflect a lack of statistical power caused by reliance on the binary emREF variable and we acknowledge that a longer post-strike observation period (to avoid censoring latency to emerge) may have afforded greater biological insights. Nonetheless, our findings do highlight an interesting question for future empirical studies: to what extent are among-individual behavioural stress response profiles consistent across stress stimuli of varying type or intensity?

4.5.2 ENDOCRINE RESPONSE

Our second hypothesis regarding repeatable among-individual variation of endocrine state was also supported only partially. We found significant variation among

individuals for pre-trial androgen levels, with a repeatability of approximately 10%. However, the repeatability of pre-trial cortisol levels was only half that and (marginally) non-significant. We found no support whatsoever for repeatable variation of either F_{POST} or $11KT_{POST}$. Note that we analysed pre- and post-trial hormone levels rather than defining the change (i.e. response) as the trait of interest, since reducing two traits to one inevitably leads to a loss of information. Nonetheless, consideration of the response offers a complementary and intuitive viewpoint. Additional models (not shown) provided no statistical evidence of repeatable variation in endocrine responses, defined as the log transformed post- minus log transformed pre-hormone release rates (results not shown).

Repeatabilities of labile traits are typically expected to decline with the inter-observation time period (Bell et al., 2009) and/or over the total period of time that observations are made (Boulton et al., 2014). Given that the repeatability of F_{PRE} was approaching significance, we carried out additional *post hoc* analysis that revealed significant (positive) correlations among trial specific measures (Table 4.6), being strongest between successive trials in the first half of the study period (i.e. 1 and 2, 2 and 3). Consistent with this finding, fitting a univariate mixed model to data from the first three trials yielded a much higher repeatability for F_{PRE} than our estimates using all data (R = 0.323 (± 0.155), P = 0.027).

Table 4.6 Estimated correlations of F_{PRE} trials (T1-T5). Estimates are conditioned on effects of weight and day order. Standard errors are shown in parentheses and significant correlations (inferred from $|r| \geq 2SE$) are denoted by bold font.

	T1	T2	T3	T4
T2	**0.845** (0.074)			
T3	**0.521** (0.191)	**0.717** (0.142)		
T4	**0.562** (0.180)	**0.530** (0.197)	0.323 (0.229)	
T5	-0.213 (0.269)	-0.314 (0.274)	-0.297 (0.275)	-0.022 (0.262)

Thus, we conclude that there are some real differences among individuals in pre-trial cortisol synthesis but that, relative to $11KT_{PRE}$ (and baseline behaviours as discussed above), these differences were less stable over the time course of our study. Our study does not address the biological reasons why this may be the case, although Table 4.6 indicates that the relatively low estimate of R overall is driven particularly by a lack of correlation between trial 5 and other observations. We note that significant effects of

Trial on mean F_{PRE} were detected (Appendix 1, Table A1.7), with an initial increase from trials 1-3 (Appendix 2, Figure A2.4b) followed by a decline across the final two observations. This is potentially indicative of habituation (on average) to stress caused by the endocrine assay procedure itself, or to an increase in the rate of negative feedback resulting in a decreased rate of cortisol output (Wong et al., 2008; Fischer et al., 2014, see discussion below). If the degree or rate of habituation or change in rate of negative feedback differs among individuals then this could contribute to the low correlations between F_{PRE} at trial 5 and the earlier observations.

4.5.3 INTEGRATION OF STRESS RESPONSES

Our third hypothesis was that behavioural and physiological stress response pathways would be integrated within individuals. Specifically, under the SCS model we predicted bolder individuals would be characterised by consistently lower glucocorticoid release but higher androgen levels (Earley and Hsu, 2008; Glenn et al., 2011). Statistical support for among-individual covariance in our trivariate analysis of boldness (Activity), F_{PRE} and $11KT_{PRE}$ was marginally non-significant but, in light of our conclusion that some among-individual variation in F_{PRE} is present, we consider two aspects of the estimated correlation structure to be noteworthy. Firstly, the among-individual correlation between F_{PRE} and $11KT_{PRE}$ was strongly positive. Although within- and between-individual covariance cannot be partitioned from a single observation, it was also the case that (mass adjusted) levels of the two hormones were positively correlated in validation samples (water borne $r = 0.624$ (0.122), $P<0.001$; entire body $r = 0.846$ (0.047), $P<0.001$). Thus, while we had predicted a negative relationship between (repeatable) levels of cortisol and 11KT, our results point towards it being positive. Secondly and again counter to our predictions, we found a strong positive among-individual correlation between activity and F_{PRE}. Thus, it is the bold (or proactive) behavioural types that exhibit higher rates of glucocorticoid release prior to undergoing the trial, commensurate with our findings in Chapter 3 of this thesis.

Many empirical studies have reported negative correlations between bold or proactive behaviours and HPA/HPI activity consistent with predictions of the SCS, (Sloman et al., 2001; Brown et al., 2005b; Verbeek et al., 2008; Raoult et al., 2012). However, exceptions to this pattern are also found, particularly in studies that have used repeated measures to quantify relationships at the among-individual level (e.g. Van

Reenen et al., 2005; Ferrari et al., 2013). The present results therefore add further weight to the suggestion that the SCS model, at least as originally proposed, may be overly simplistic. One possibility is that a model with two (or more) independent axes of behavioural response variation, for example locomotion and fearfulness (Van Reenen et al., 2005), might be more appropriate. Equally, this may be true for endocrine response, with variation in the degree of the endocrine response, habituation and negative feedback all having the potential to be independent axes of endocrine response variation. Recently, an argument has been put forward that distinguishing between the qualitative (coping style) and quantitative (stress reactivity) components of among-individual variation is important (Koolhaas et al., 2010). Koolhaas et al. (2010) also suggest that widespread support for the proactive-reactive SCS model in domesticated species may be an artefact of strong selection on either physiological or behavioural traits in captive-bred populations. If so then relationships between these traits will likely be more variable in wild populations. Although the fish used in our study were captive bred, they were only two generations removed from the wild and can therefore be considered broadly genetically representative of their natural source population.

The water borne endocrine assay has been verified in many fishes including a number of Poecillids, (e.g. Netherton et al., 2004; Archard et al., 2012; Gabor and Contreras, 2012). Here we were able to validate its use as a non-invasive proxy for whole-body hormone levels in the sheepshead swordtail, *Xiphophorus birchmanni*. Nonetheless, some patterns in our data pose challenges for interpretation. In particular, we found a significant decline in mean cortisol released between paired (i.e. individual and trial specific) pre- and post-trial samples. Thus on average, the cortisol response to stress imposed by the trial was negative, not positive as expected. It is possible that our 60 minute steroid collection period was too long resulting in capture of the cortisol surge released as a result of handling stress in the F_{PRE} levels, and saturation of the HPI axis due to negative feedback and/or reabsorption of cortisol during the F_{POST} collection (Scott and Ellis, 2007). Arguments that water borne collection procedures are stressful, despite being non-invasive, have been put forward (Wong et al., 2008). Thus, rather than being baseline measures, our F_{PRE} may indeed be indicative of a stress response. Suggestions of habituation to the technique also have been made, thus rendering the repeated measures approach difficult to interpret (Wong et al., 2008; Fischer et al., 2014). Certainly we found a (non-significant) decrease in mean for F_{PRE} levels after the

third trial (Appendix 2, Figure A2.4). Suggestions that a flow-through system for steroid collection may be a better method of hormone collection as fish do not then encounter confinement stress are valid; however, necessarily water borne collection requires physical and chemical isolation, and, if studies on both behavioural and physiological components of SCS are to be carried out, then these necessitate capture, handling and confinement.

4.5.4 CONCLUSION

In conclusion, our multivariate repeated measures approach allowed us to characterise physiological and behavioural response to an acute stressor in a second generation captive bred population of *X. Birchmanni*. Although there was evidence for among-individual variance in F and 11KT, the lack of significant repeatability for cortisol and the positive correlations between physiological and behavioural traits did not lend support to the SCS paradigm. The fact that repeatabilities of endocrine levels were stronger when observations were closer together suggests the potential for experimental design to have a strong influence on biological conclusions regarding whether or not a trait is repeatable. Our findings add weight to the suggestion that cortisol measures in wild (or recently wild derived) populations may be less stable than those measured in laboratory adapted populations (Koolhaas et al., 2010). In line with other recent studies, our results also suggest that the water borne collection procedure used is a mild stressor, and thus that interpretation of pre-contest levels as baseline levels may not be appropriate. We conclude that stress coping style and personality (certainly boldness) are two separate axes of latent variation, that may (or may not) converge, that this may be species and/or environment dependent, and that high correlations between physiological and behavioural traits are likely to be an artefact of laboratory studies.

CHAPTER 5

QUANTIFYING THE GENETIC BASIS OF SOCIAL DOMINANCE IN THE POECILIID, XIPHOPHORUS BIRCHMANNI

5.1 ABSTRACT

Competition for resources is an important source of environmental effects on individual phenotype. While growth, life history traits, and fitness (i.e. survival and fecundity) are typically dependent on acquiring resources (e.g. food, territory, mating opportunities), individuals within a population are expected to vary in their competitive ability (or social dominance). Because winners gain resource, and therefore fitness, at the expense of losers, social dominance should be under strong selection. The evolutionary consequences of this selection will depend on the extent of genetic (co)variation in and between traits that determine and/or depend on dominance. Although body size and weaponry are known to predict dominance status in many animals, it has been widely hypothesised that aspects of animal personality such as boldness may also be important. In this study we investigated the effects of competition for space on growth, life history, longevity and personality traits in a pedigreed population of the sheepshead swordtail, *Xiphophorus birchmanni*. Fish were reared in mixed family social groups and subjected to different density treatments. Repeated observations were made on morphology and behaviour over a 50 week period, with timing of non-repeated life history events (e.g. maturation age and size, death) also recorded. As expected, growth and longevity were reduced at high levels of competition. We also found that male dominance score was repeatable, and that dominant individuals grew faster, lived longer and matured at a later age. However, while our analyses demonstrated significant among-individual correlation between boldness and dominance, the association was strongly negative, not positive as we had predicted based on expectations in the literature. Quantitative genetic modelling provided evidence for genetic contributions to phenotypic variance and, in some cases, between-trait covariance. Nevertheless, the estimated heritability of dominance itself was low and not statistically significant. We therefore conclude that correlations between dominance and the other aspects of phenotype and fitness considered here are driven primarily by environmental rather than genetic effects.

5.2 INTRODUCTION

An individual's phenotype is determined by its genotype and the particular environmental conditions it experiences during development. Competition for resources (e.g. food, space) from conspecifics is one important environmental factor known to have large effects on phenotypic traits including growth and life history traits (e. g, maturation, fecundity, longevity). Competition is also found across many different social contexts. For example, sibling competition often impacts growth rates in animals that provision their young (e.g. Nilsson and Svensson, 1996) while overt aggression associated with male-male competition for mates can sometimes be an important source of mortality (e.g. Liker and Szekely, 2005). Importantly, by producing winners and losers, competition generates variation in resource dependent traits and, ultimately in fitness. Since winners increase their (relative) fitness at the expense of losers (Brockelman, 1975), those traits that contribute to competitive ability are also expected to be under strong selection. If so, then the evolutionary consequences of this selection will depend on the genetic covariance structure between traits causal to social dominance (Wilson, 2014). Here, taking a quantitative genetic approach, we characterise the genetic basis of social dominance in a population of the freshwater Poeciliid, *Xiphophorus birchmanni* and explore the extent that genetic and environmental effects (including the degree of competition itself) shape the multivariate phenotype.

The overall effect of high competition in a population is to reduce mean (absolute) fitness. While therefore recognised as one of the most important ecological mechanisms regulating population growth (Schoener, 1983; Sih et al., 1985; Chase et al., 2002), fitness effects of competition are themselves driven by impacts across a potentially wide range of phenotypic traits. Most obviously if (on average) individuals obtain less of a resource (e.g. food), then declines in the mean of resource dependent traits (e.g. growth) are expected as competition increases. Plastic responses to changing levels of competition will typically be multivariate, involving concerted change across whole suites of correlated traits and fitness components (Baur and Baur, 1992) that depend on resource acquisition and may sometimes be adaptive. For example, if competition changes the optimal resolution of resource or time allocation trade-offs (Stearns and Koella, 1986; Roff, 2000), then animals may alter their allocation strategies accordingly. Adaptive plasticity is particularly well documented for behavioural traits with, for example, parental investment in offspring adjusted

according to need (Smiseth et al., 2008). Increases in behaviours positively linked to resource acquisition (e.g. activity, boldness, aggressiveness; discussed further below) have also been reported (Relyea, 2004; Thomson et al., 2012).

However, from an evolutionary perspective perhaps the most important role of competition is as a mechanism that generates variation in phenotype and fitness. Within a population, individuals can vary in competitive ability or *social dominance,* that can be defined for current purposes as an individual's (repeatable) tendency to win (or hold) resources under competition (following Wilson et al., 2011b). Dominant individuals win resources and therefore can increase their (relative) fitness at the expense of subordinates, for example by growing faster, maturing earlier, or investing more in reproduction (Cobb and Tamm, 1975; Bernstein, 1976; Huntingford et al., 1990; Fox et al., 1997; Bell et al., 2012). The question of what factors determine social dominance can be addressed at two levels, firstly by seeking to understand the phenotypic features of an individual that make it more (or less) likely to succeed in competition, and secondly by asking to what extent social dominance is dependent on genetic versus environmental effects?

Much of our understanding of how dominance is determined comes from studies of dyadic contests where there are usually clearly defined winners and losers. Studies have highlighted the importance of morphological traits such as body size and/or weapons (e.g. horns, Preston et al., 2003) that are widely used as measures of resource holding potential (RHP, i.e. an individual's absolute fighting ability, Parker, 1974). Although these morphological traits are frequently found to predict contest behaviour and outcome, there is growing recognition that social dominance can also depend on an individual's (repeatable) behavioural phenotype or personality (Reale et al., 2010). Two personality traits in particular have been widely linked to social dominance – aggressiveness and boldness. Since competitive environments tend to promote agonistic behaviour the link with aggressiveness is perhaps unsurprising (e.g. Bernstein, 1976; Francis, 1988; Ostner et al., 2008; Magellan and Kaiser, 2010b). While it is important to recognise that aggression (actual, threat or signal of attack, Hand, 1986; Francis, 1988) and dominance are not equivalent, the former is a behavioural strategy often used to assert the latter (Bernstein, 1976). Moreover, patterns of aggressive behaviour expressed among individuals in a group situation are often a good predictor of dominance hierarchies (Francis, 1988; Jackson, 1991). Thus, the likelihood of a successful competition outcome may typically be higher for individuals with

aggressive personalities (Earley, 2006; Magellan and Kaiser, 2010a; Wilson et al., 2011a). Boldness is somewhat more difficult than aggression to define (Carter et al. 2012, Boulton et al. 2014), but can be loosely described as a willingness to take risks, especially in novel situations. If the functional significance of variation in boldness remains to be fully understood, evidence from a wide range of taxa shows that it is often (positively) correlated with aggression (Sih et al., 2004a; Johnson and Sih, 2005; Pintor et al., 2008; Ariyomo and Watt, 2012). Furthermore, both these personality traits are commonly associated with social dominance, resource-dependent life history traits and fitness measures (e.g. Dingemanse et al., 2004; Brown et al., 2005a; Bell and Sih, 2007; Stamps, 2007; Biro and Stamps, 2008; Cote et al., 2010; Ariyomo and Watt, 2012; Rudin and Briffa, 2012; Mutzel et al., 2013).

Where fitness is tightly linked to competitive outcome, it has been argued that traits determining RHP (and thus dominance) will be under strong directional selection (Huntingford et al., 1990; Campton, 1992; Kruuk et al., 2002; Benson and Basolo, 2006; Prenter et al., 2008). Simple evolutionary theory predicts that, all else being equal, this selection should erode genetic variance. If so, then phenotypic variation observed at equilibrium will largely be due to environmental effects (Kruuk et al., 2002; Benson and Basolo, 2006). However, in practice quantitative genetic studies typically reveal significant genetic variation in traditional RHP traits (e.g. body size) and there is growing evidence that aggression and boldness are also heritable (e.g. Guhl et al., 1960; Bakker, 1986; Benus et al., 1991; Drent et al., 2003; Sinn et al., 2006; Ariyomo et al., 2013b). Thus it seems plausible that dominance will often be determined by genetically variable traits. If so, this will have important implications for our understanding of life history evolution (Wilson 2014). This is because genes that increase dominance will also allow individuals to succeed in competition, gain more resources, and so invest more in all resource dependent life history traits. In this way genetic effects on dominance could play a major role in shaping the genetic-variance-covariance matrix (**G**) between life history traits. Ultimately genetic variance is the raw material for evolution, and understanding the genetic (co)variance structure between traits associated with competitive success is thus important for predicting phenotypic evolution.

Here we test the genetic basis of dominance, and characterise both genetic and environmental contributions to (co)variance in and between dominance, personality, growth, life history traits and fitness (longevity) in a captive population of the

sheepshead swordtail, *Xiphophorus birchmanni*. Swordtails have been widely used in studies of social dominance (e.g. Borowsky, 1973; Bao and Kallman, 1978; Beaugrand and Zayan, 1985; Franck and Ribowski, 1989; Earley, 2006; Brown et al., 2007; Walling et al., 2007; Boulton et al., 2012), while previous work on this particular population has found evidence of stable personality traits including aggressiveness (Wilson et al., 2013) and boldness (Boulton et al., 2014). In adult males aggressiveness has been shown to be a better predictor of dyadic contest outcome than size, while dominant individuals tend to grow faster (as measured by relative weight gain; Wilson et al., 2013).

By manipulating density, we subject a captive bred generation of fish to contrasting low (L) and high (H) competition treatments, in both early and later life, to test for effects on growth, personality and life history traits. We hypothesise that high competition, particularly if experienced in early life, will reduce growth rates and negatively impact fitness (measured here as longevity). After testing the direct effects of competition on phenotypic expression, we use a multivariate modelling approach to estimate the relationships between traits at the individual and additive genetic levels. We predict that among-individuals, social dominance will be correlated with personality (boldness), growth and life history (size, age and condition at maturity), and fitness (longevity). Similar correlation structure is expected at the genetic level, provided heritability for dominance is present. This is because genotypes predisposing to contest winning should positively influence growth, life history, and survival. However, if selection on dominance has been strong, then we predict it will have a low heritability, and the among-individual phenotypic correlations will be driven by environmental, rather than genetic, effects. Finally, we test for genotype by environment (i.e. competition treatment) interactions to see whether the genetic (co)variance parameters are themselves sensitive to the competitive environment experienced. If so then this implies that the plastic responses to competition level are themselves genetically variable, and so could evolve if subject to selection (Scheiner, 1993).

5.3 METHODS

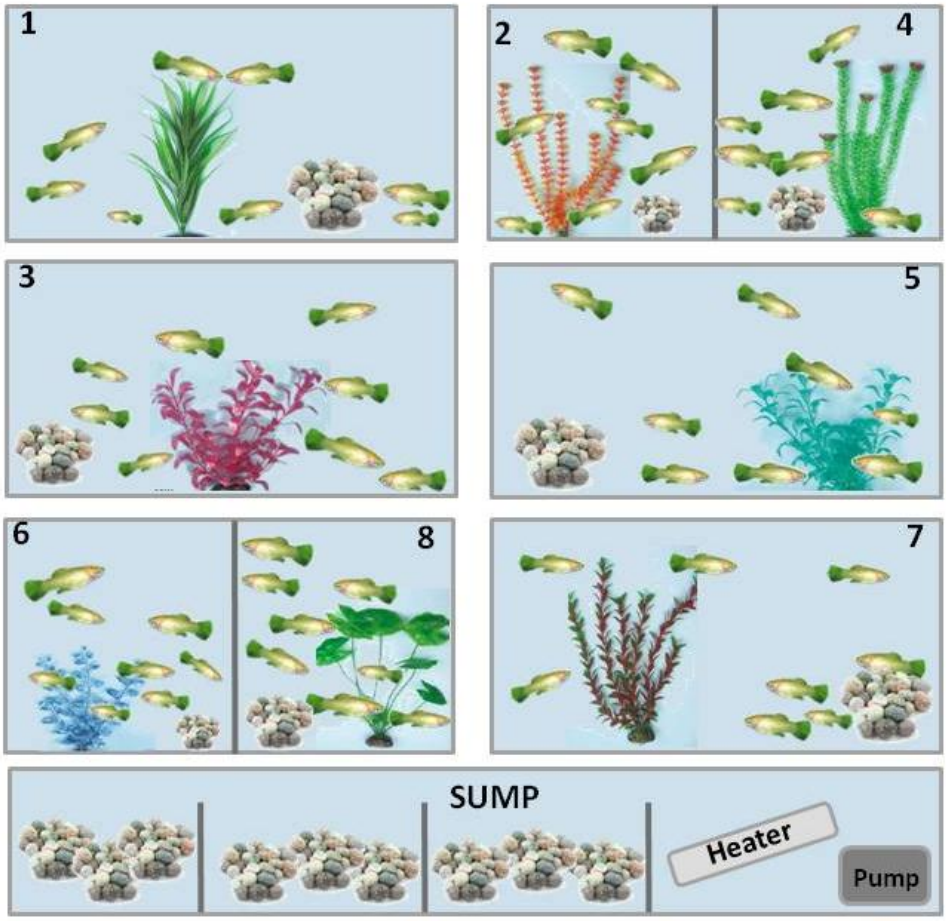
5.3.1 STUDY SPECIES AND PRODUCTION OF FAMILIES

In the spring of 2010, one hundred wild adult fish (60 female and 40 male) were caught from the Arroyo Coacuilco, near the town of Coacuilco, municipality of San Felipe

Ortizalan, Hidalgo, Mexico, and imported to the UK. Fish were tagged with visible implant elastomer (http://www.nmt.us/products/vie/vie.shtml) to allow individual identification, then randomly allocated to breeding groups (1 male:3 females) and housed in glass aquarium tanks (37 x 37 x 22 cm) enriched with small stones and living plants. Water temperature was maintained at 21-23°C, and light was provided on a 12:12 hour light:dark cycle. Breeding group fish were fed twice daily on proprietary flake food (ZM foods, http://www.zmsystems.co.uk/) and previously frozen bloodworm and daphnia.

Between August 2010 and May 2011, a captive bred generation of *Xiphophorus birchmanni* (n = 384, comprising 77 families nested within a half-sib structure with 15 males and 28 females parents represented) was produced as follows. Breeding groups were inspected daily and obviously gravid females were isolated in a separate tank enriched with stones and artificial plants made from nylon netting. The latter were to provide additional refuge for new born fry since cannibalism is well known in Poeciliid *sp*. Females were inspected daily and returned to their breeding group tanks after giving birth. All offspring from each family produced were individually measured on the day of birth, (standard length (SL) from the tip of the snout to the caudal peduncle) to the nearest 0.01 mm using digital callipers. Families were then transferred together to one half of a brood tank. Brood tanks were prepared by inserting a Perspex framed divider covered with fine-gauge black nylon net into a 37 x 37 x 22 cm tank, thus separating into two equal volumes. Two families were therefore kept in each brood tank. Where a family comprised >eight individuals it was divided equally between multiple brood tanks across different stacks (maximum 8 individuals per brood tank). Brood tanks (and experimental tanks, see below) were grouped in stacks, each comprising six tanks connected to a single sump and therefore sharing a common recirculating water supply (Figure 5.1). This set up was to reduce the potential for bias in genetic parameters due to common environment effects (i.e. confounding of family structure with any water quality effects on traits of interest). Offspring were fed twice daily on fresh brine shrimp nauplli and ZM spirulina and brine shrimp flake (as above).

Fig. 5.1 Stack design. Four full- (1, 3, 5, and 7) and four half-size tanks (2, 4, 6, 8) were housed on a shelving unit. A Sump tank positioned below controls water quality and temperature. Each tank contained a few pebbles, a living plant, and n = 8 fish. 64 fish of a similar age were required to set up the stack.



5.3.2 DENSITY TREATMENTS

At an average of 16 weeks (range 12-27 weeks), fish entered the experiment. After tagging to allow identification, individuals were assigned to mixed family rearing groups (n = 8 fish per group) and subjected to one of two density treatments. Low density rearing groups were housed in a full tank, high density groups were housed in a half tank (partitioned as described above). Six stacks (each comprising four low and four high density groups on a recirculating water supply; Fig. 5.1) were sequentially established. Since a stack of rearing groups could only be set up when 64 fish (eight groups x eight fish per group) reached a size suitable for tagging, variation in age of fish entering the experiment was unavoidable. The sex ratio of rearing groups was not controlled since juvenile *X. birchmanni* cannot be sexed from external morphology. Fish were fed twice daily, with low (L) and high (H) density groups receiving the same

ration (a mixed diet of flake, previously frozen blood worm, daphnia and brine-shrimp nauplli). After 28 weeks at this regime (part 1 of the study), the density treatments were reversed for four randomly chosen groups within each stack, and groups were maintained for a further 22 weeks (part 2 of the study) when the experiment was terminated. Thus within each of the six stacks, four density regimes were experienced (LL, LH, HL, HH), with two groups (of initial n = 8 fish) per regime. Natural mortality over the course of the experiment resulted in variation in group size through time, although survival was high (368 of 384) over the first density treatment period (i.e. part 1 of the study).

5.3.3 GROWTH AND LIFE HISTORY

Fish were measured (standard length (SL) and mass (WT)) at the start of the experiment and subsequently at four-weekly intervals (see Appendix 3 for exact dates). These data were used to calculate growth rates and condition factor (see below). Up to 13 measures were made on each fish (with measure eight corresponding to the end of part 1, i.e. the 28 week initial density treatment). For males, maturation age was recorded as the age at the first measure when gonopodium formation was apparent. We did not assign maturation ages to females since clear morphological indicators of female maturity are not available in this species. Longevity was recorded as individual age at death in days (regardless of whether death was natural or the animal was euthanized for welfare reasons), or age at the end of the experiment (for fish alive at measure 13).

5.3.4 BEHAVIOURAL PHENOTYPING

Behavioural data were collected to provide information on two different personality traits: boldness and dominance. We ascertained boldness for all fish in the study using open field trials (OFT). Individuals were subjected to up to four trials: two at the initial density treatment in part 1 (weeks 13 and 21) and two in part 2 of the study (weeks 33 and 41, see Appendix 3). The OFT have been described in full elsewhere (Chapter 2) and the genetic modelling here is a re-treatment of data already published in relation to testing the temporal stability of personality (see Boulton et al., 2014). Briefly, a 45 x 25 x 25 glass tank was filled to a depth of 8 cm with room temperature water (22°C), and individual fish were introduced directly from their experimental rearing tanks. After a thirty second acclimation period, behaviour was filmed for five minutes using a Sunkwang C160 video camera suspended above the tank. A suite of traits putatively indicative of boldness were extracted from the video using tracking software (Viewer

11http://www.biobserve.com/products/viewer/). These were track length (total distance travelled, cm), activity (percentage time spent moving), area covered (percentage tank base moved over) and time in middle (time spent in zone 2, mins (see Chapter 2, Fig. 2.1 for zone definitions)). Previous analysis using multivariate linear mixed models revealed that the among-individual (i.e. repeatable) component of multivariate behavioural variation was dominated by a single major axis of variance, broadly commensurate with expectations of a shy-bold continuum (Boulton et al., 2014). For current purposes we selected a single trait, activity (percentage time in trial spent active) to act as a proxy for boldness.

Dominance scores were determined for males only using in-tank observations (ITO). Behaviour of each male in each rearing group was recorded for five minutes, on a maximum of five occasions, two at the initial density treatment during part 1 of the study (weeks 18 and 25) and three at the final density treatments during part 2 of the study (30, 38, 44 weeks). Within groups, focal males were watched sequentially in a haphazard order by a recorder who was seated in front of the tanks in full view of the fish. Fish were accustomed to researcher presence and activity in the laboratory and our assessment was that this did not impact behaviour. Within groups, individual males were readily and individually identifiable from phenotype (size, melanophores) and elastomer tags. Previous work has shown that aggression positively predicts feeding dominance among male *X. birchmanni* (Wilson et al., 2013) while male dominance hierarchies are known to determine access to females in swordtails generally (Magellan and Kaiser, 2010b). For each five minute observation period we assigned a within-group dominance score to each focal male as the total number of aggressive actions toward other males (attacks, dorsal fin displays, chases), plus the number of courting attempts (displaying to female, shepherding away from other males), minus the number of submissions (retreating or fleeing from another male) and aggressive acts received.

5.3.5 STATISTICAL ANALYSIS

We first tested the hypothesised density treatment and genetic effects on growth, life history and behaviour using (univariate) linear mixed effect models. These were fitted by restricted maximum likelihood (REML) using the program ASReml V3 (Gilmour). To model growth we derived three response variables from the primary WT and SL data. These were:

i) Percentage daily increase in SL, $(SGR_{SL}) = \frac{[(ln(SL_{x+1}))-(ln(SL_x))]}{Age_{x+1}-Age_x} * 100$

ii) Percentage daily increase in WT $(SGR_{WT}) = \frac{[(ln(WT_{x+1}))-(ln(WT_x))]}{Age_{x+1}-Age_x} * 100$

iii) Change in condition factor (CCF) = $(CF_{x+1} - CF_x)$, where $CF = 10000 * WT/SL^3$

where *x* = denotes measure (1-13) and Age = age (days).

Life history traits modelled were age at maturity ($ageMAT_M$), standard length at maturity ($SLmat_M$), mass at maturity ($WTmat_M$), condition factor at maturity ($CFmat_M$) and longevity (LONG), (where subscript M denotes traits measured for males only). Behaviours modelled were activity (ACT) and dominance score (DOM_M). Note that while several traits are measured as percentages, visual inspection of residuals from univariate models indicated that assuming normal error structures was reasonable. However, we square root transformed the activity data to reduce skew and maintain consistency with its treatment in Chapter 2 of this thesis.

5.3.5.1 Fixed effects

In addition to testing the effects of our experimental density treatments (see below) we included a number of additional fixed effects to control statistically for putative sources of phenotypic variation not directly relevant to our main hypotheses. Wald *F*-tests were used to assess the significance of all fixed effects.

For all traits, models were formulated with fixed effects of *the mean*, *stack* (as a six level factor) *sex* and *sex ratio*. The latter two terms were not necessarily known at the time when a trait was actually observed (if fish were yet to mature) although can be inferred retrospectively. *Sex* was fitted as a three level factor since eleven fish could not be unambiguously determined as either male or female at any time during the experiment. Note also that *sex ratio* was defined for any observation on any fish as the proportion of that individual's tank mates that eventually become male (i.e. number of males in group excluding self/(number in group -1)). We also included a linear effect of *geometric mean group size* (geometric mean of group size at measures up to and including the present). This was to account for any effects of declining group size caused by mortality. We used the geometric rather than arithmetic mean as this should better capture the expected cumulative effect of any decline in competitor (i.e. tank mate) numbers over time. For growth and behaviour traits we modelled the average age trajectory using a 3rd order (i.e. cubic) function. For behavioural traits we also included

trial (as a factor with up to five levels to account for any habituation or sensitisation effects) and *day order* (as a linear effect). *Day order* (the number of trials run that day) was also fitted as a proxy for time of day. This was included since diurnal rhythms are known to affect some behavioural traits (Boulton et al., 2014; Thesis Chapter 4). All fixed covariates were mean centred prior to fitting.

Effects of density treatments were then statistically modelled as follows. First we modelled an effect of *early life density treatment* (ELD) as a two level factor (low, high) on observations made in part 1. We then fitted an interaction between ELD and *late life density* (LLD) on observations made in part 2. The ELD*LLD interaction actually defines a four level factor, and thus accounts for all possible treatment regimes (LL, LH, HL, HH). For clarity we subscript treatment effects that are conditional on the time of phenotypic observation as ELD_{p1} and $(ELD*LLD)_{p2}$. These conditional effects are appropriate since the second (part 2) treatment cannot influence observations made in part 1; however it is possible that both first and second treatments influence observations made in part 2.

The great majority (84%) of males that matured during the study did so during part 1 (Fig. 5.2g). Since the sample size of maturation ages observed in the second part of the study was so small (n = 10) we elected to include treatment effects of ELD, but not the (ELD*LLD) interaction on male maturation traits (see Table 5.1). Conversely, there was little mortality in the first part of the study (13%), and therefore we modelled density treatments slightly differently to assess the main effects early and late life density (ELD and LLD) and test for an interaction between the two (ELD*LLD). Note then that for maturation and longevity, the treatment effects are not conditional on the time of phenotypic observation (and there is only one observation per individual) therefore we do not subscript them.

5.3.5.2 Random effects

With fixed effects as described above we first ran models with the single random effect of fish identity to test for repeatable (among-individual variation) in those traits with repeated measures (i.e. SGR_{SL} SGR_{WT}, ACT, DOM_M). We tested the significance of the among-individual variance term (V_I) by likelihood ratio test (LRT) comparison to a reduced model (i.e. without the random effect), assuming the test statistic to be asymptotically distributed as a mix of 50:50 χ^2_0 and χ^2_1 (following Visscher, 2006). We estimated trait repeatabilities, (R, conditional on the fixed effects) as the ratio of V_I to

total phenotypic variance $V_{P,}$ with the latter calculated as the sum of V_I and V_R (the residual variance).

For all traits we then used animal models (Kruuk et al., 2008) to fully partition the variance into additive genetic (V_A), brood tank (V_{BT}) and residual (V_R) variance components. In addition to individuals additive genetic merit, brood tank (BT) was fitted as a random effect to account for any common environment effects experienced in the brood tank (i.e. when fish were housed in family groups) that might otherwise bias genetic parameters. For traits with repeat measures we also fitted a permanent environment (PE) effect and partitioned the corresponding variance component (V_{PE}) that is expected to include non (additive) genetic sources of among-individual variance. We tested the significance of random effects using likelihood ratio tests as described above. We determined heritability (h^2) as V_A/V_P and similarly calculated the ratios of all other variance components to V_P to provide standardised effect sizes.

5.3.5.3 Testing for evidence of univariate GxE

We then tested each trait for evidence of genotype-by-environment interaction (GxE). We did this by adding an interaction between individual genetic merit and ELD treatment to the random effect structure described above. This model was compared to that of the simpler model (no GxE term) using LRT. Note that we elected to test for GxE interactions using ELD only, as we deemed our data set too small to meaningfully test for differences in genetic variance across the four ELD*LLD categories.

5.3.5.4 Multivariate analysis

Multivariate models were then used to test the covariance structure between traits and partition it into components arising from genetic and environmental effects. Fixed effects on all traits were as described above for the univariate analyses. Since achieving stable model convergence is difficult for large multivariate models we chose to reduce the number of traits from ten to seven. Specifically we dropped SGR_{WT} due to very high correlations between this and SGR_{SL}, (r_{PE} = 0.998 ± 0.101; r_G = 0.996 ±.0.061; r_{BT} = 0.999 ± 0.090). Similarly strong correlations were found between $SLmat_M$ and $WTmat_M$, (r_{PE} = 0.977 ± 0.056; r_G = 0.8909 ± 0.014). Thus we conclude that growth and maturation size traits based on SL vs. WT contain essentially equivalent biological information and we therefore elected to reduce complexity by including only SGR_{SL} and $SLmat_M$ in our multivariate models. The inclusion of CCF in multivariate models prevented

convergence of log likelihoods due to very small variance components; therefore this trait was also omitted from the final models.

Phenotypic (co)variance was then partitioned into an additive genetic component (presented as the additive genetic (**G**) matrix containing estimates of the additive genetic variance V_A for each trait and covariance COV_A between each pair of traits), and corresponding permanent environment (**PE**) and residual (**R**) structures. Since statistical support for V_{BT} was provided by univariate models for SGR_{SL}, we included a brood tank effect only on this trait in the multivariate modelling. Note that for traits with multiple measures, environmental (or strictly non-additive genetic effects) are partitioned into within-individual (V_R) and among-individual (V_{PE}) components. For traits measured once only, we model the environmental variance in the **PE** structure (as among individual variation) since, for example, within-individual variance in longevity is not observable (i.e. each fish only died once). All among-trait covariance terms in **G** and **PE** were rescaled to give the corresponding genetic (r_G) and environmental correlations. In practice, despite dropping three traits and simplifying the brood tank effects by fitting them only on SGR_{SL} we were still unable to obtain stable convergence for the seven trait animal model. Instead, our **G**, **PE** and **R** structures were estimated from a series of smaller models fitting three to six traits simultaneously.

We tested the significance of between trait covariance in **PE** and **G** using likelihood ratio tests to statistically compare a series of models: (A) with diagonal **PE** but no **G**, allowing estimation of only the permanent environment variance V_{PE}; (B) full **PE** but no **G,** allowing estimation of COV_{PE}; (C) full **PE** and diagonal **G**, allowing V_A to be estimated; (D) full **PE** and **G**, allowing all parameters to be estimated. Comparison of model (A) with model (B) tests for significant phenotypic covariance between traits (at the among-individual level). Comparison of models (B) and (C) tests the hypothesis that additive genetic effects explain at least some of the phenotypic variance observed, while comparing (C) to (D) tests whether genes also contribute to the phenotypic covariance (Wilson et al., 2013). Finally, comparison of models (B) and (D), provided an overall test of whether the **G** matrix explains a significant part of the phenotypic (co)variance structure of the traits modelled (Wilson et al., 2010a).

5.4 RESULTS

In total, for each of 222 males, 151 females and 11 immature individuals (sex undetermined at time of death or end of data collection period) from the pedigreed generation (n = 384), we collected 4992 morphological measures (SL and WT). From these the three growth-associated traits (specific growth rate of SL and WT and change in CF per unit time – SGR_{SL}, SGR_{WT} and CCF respectively) were derived. Longevity was recorded for all fish (n = 384) and traits pertaining to male maturation (age, SL, WT) were obtained and derived (CF) for males surviving to maturity (n = 193). A total of 1235 open field trials (OFT) were conducted (both sexes) and 864 in tank observation completed to determine male dominance scores. Mean observed phenotypic trajectories across measures are shown by density treatment category for size, growth, and life history traits in Fig. 5.2.

5.4.1 ENVIRONMENTAL EFFECTS

5.4.1.1 Environmental effects on growth (SGR_{SL} and SGR_{WT}) and change in condition factor (CCF)

Across all observations, the two measures of growth rate (SGR_{SL} and SGR_{WT}) were strongly positively correlated ($r = 0.851$, $P < 0.001$). Unsurprisingly, univariate models yielded qualitatively similar results for these traits (Table 5.1). Significant stack effects were found (Table 5.1), as were non-linear declines in growth rate with age (Table 5.1) that tended to plateau from around measure 7 (Fig. 5.2). Our data also show that growth rates, as measured by SGR_{SL} and SGR_{WT}, are higher in females than males or individuals of unknown sex (effect of being male on $SGR_{SL} = -0.012 \pm 0.002$, $P < 0.001$; on $SGR_{WT} = -0.060 \pm 0.010$, $P < 0.001$; Table 5.1).

We also found evidence of significant social environment effects on growth. While sex ratio had no statistically significant effect, growth rate increases with declining number of tank mates as expected (effect of gmCOMP on $SGR_{SL} = -0.010 \pm 0.002$, $P < 0.001$; on $SGR_{WT} = -0.044 \pm 0.011$, $P < 0.001$). High early-life density (ELD_{p1}) significantly reduced growth rates in part 1 (effect of high ELD_{p1} on $SGR_{SL} = -0.005 \pm 0.002$, $P < 0.001$; on $SGR_{WT} = -0.020 \pm 0.011$, $P < 0.001$), while treatment effects on growth in part 2 were also found (indicated by the significant $(ELD*LLD)_{p2}$ interaction terms in Table 5.1). Comparison among levels of the $(ELD*LLD)_{p2}$ term showed this latter result to be driven by a large decline in growth rate for fish swapping from low to high density treatments (Fig 5.3a, b).

CCF showed a non-linear decline with age (Fig. 5.2) and was subject to significant fixed effects of stack and sex (with CCF lower in males), as well as declining with increasing gmCOMP (Table 5.1). High early life density (ELD_{p1}) reduced CCF in part 1 (effect on CCF = -0.001 ± 0.001, P = 0.037), while a significant $(ELD*LLD)_{p2}$ interaction in part 2 was again driven by a reduction in CCF for fish swapping from low density to high density (Fig 5.3c).

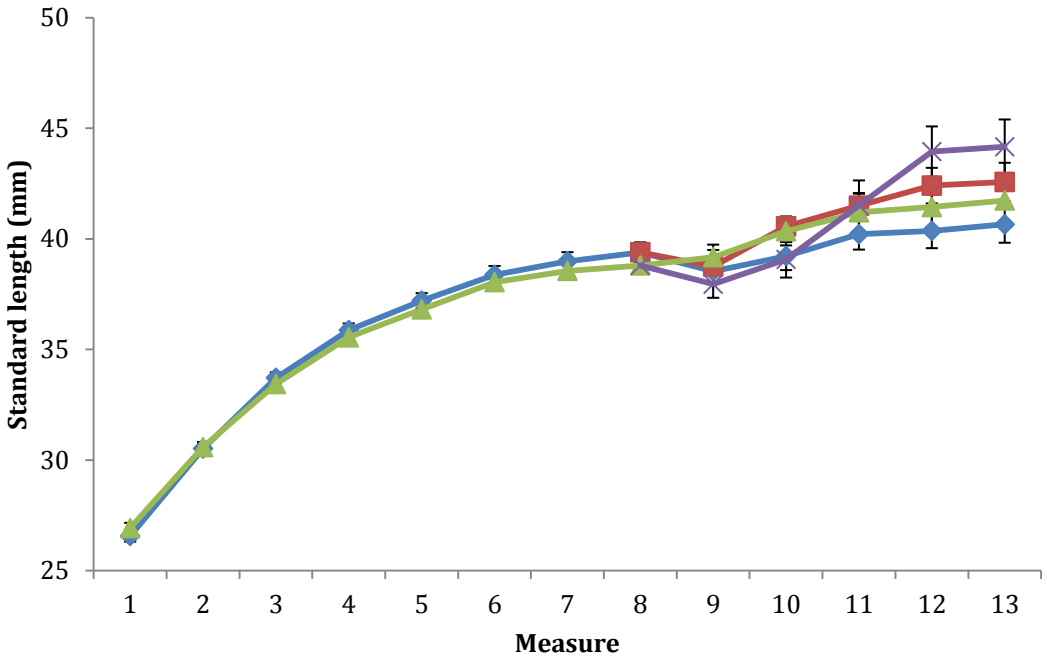
5.4.1.2 Environmental effects on life history

There was no evidence of significant ELD treatment effects on male maturation age, size (WT, SL) or condition (Table 5.1). However, some environmental effects were found for male maturation traits. For example, male maturation age was reduced by higher numbers of tank mates (gmCOMP effect of -102.10 ± 13.76 days/fish, P <0.001) but increased by a more male-biased sex ratio (sex ratio effect corresponding to an average delay of 96 ± 15 days, P <0.001, with all male tank mates relative to all female; Table 5.1). Sex ratio, but not gmCOMP, also influenced size at maturity, with males being longer and heavier at maturation in more male biased groups (sex ratio effect on $SLmat_M$ = 5.529 ± 1.024 mm, P <0.001; on $WTmat_M$ = 0.624 ± 0.129 g, P <0.001), (Table 5.1). Maturation age, mass and condition factor also differed significantly among stacks.

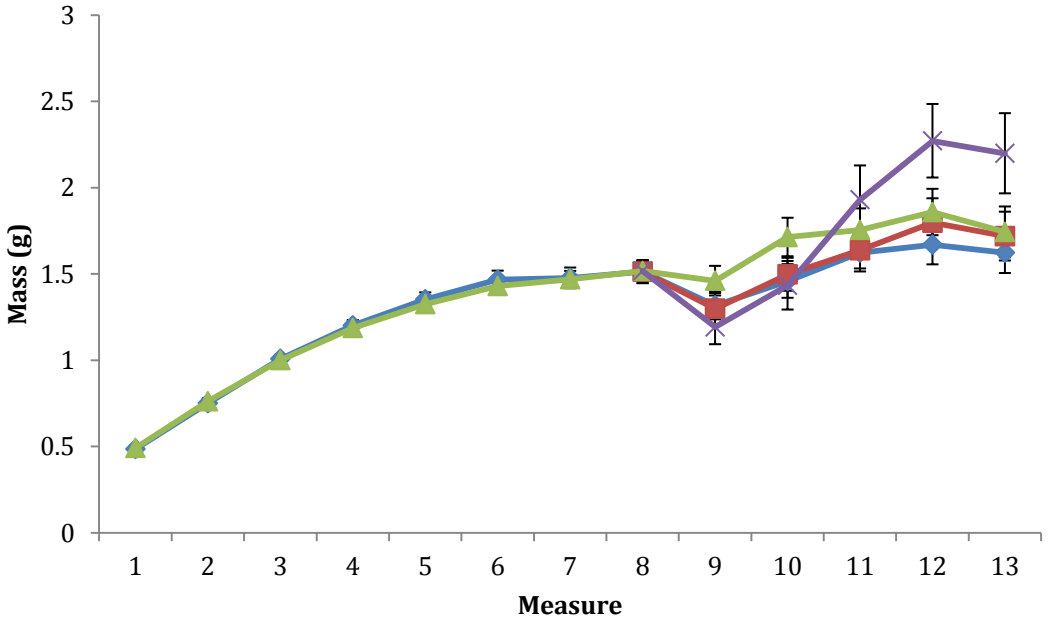
Our univariate models of longevity also showed significant stack effects, while males and individuals of unknown sex tended to have shorter lives on average than females (effect of being male relative to female, -8.219 ± 6.201 days, and effect of being of unknown sex relative to being female = -58.55 ± 17.02 days). When analysed as a three level factor, sex effects are significant (P = 0.002), however, this is clearly driven by the individuals of unknown sex. While male biased sex ratios did not adversely affect longevity, an increased number of tank mates did (gmCOMP effect = -41.65 ± 3.47 days, P <0.001). Of the density treatment effects, we found that LLD but not ELD was significant, with high LLD reducing longevity by a mean of 18.39 ± 7.68 days (P = 0.005, Table 5.1). Note that the ELD*LLD term was not significant when modelled simultaneously with the main effects of ELD and LLD, but was if modelled without LLD included (P = 0.012, effect sizes presented in Fig. 5.3d). This confirms that there were differences in longevity among the four treatment regimes, but that these can be explained as an effect of high LLD reducing longevity regardless of ELD experienced.

Fig. 5.2 Observed means for morphological and life history traits by measure for the four density treatments (Low/Low (♦); Low/High (■); High/High (▲); High/Low (X)) at each of the measures: a) Standard Length (SL, mm); b) Mass (g); condition factor; d) percentage increase in standard length per unit time (LN, SGR_{SL}); e) percentage increase in mass per unit time (LN, SGR_{WT}); f) change in condition factor per unit time (CCF); g) percentage living males matured; h) percentage survival. Error bars represent standard errors around the means.

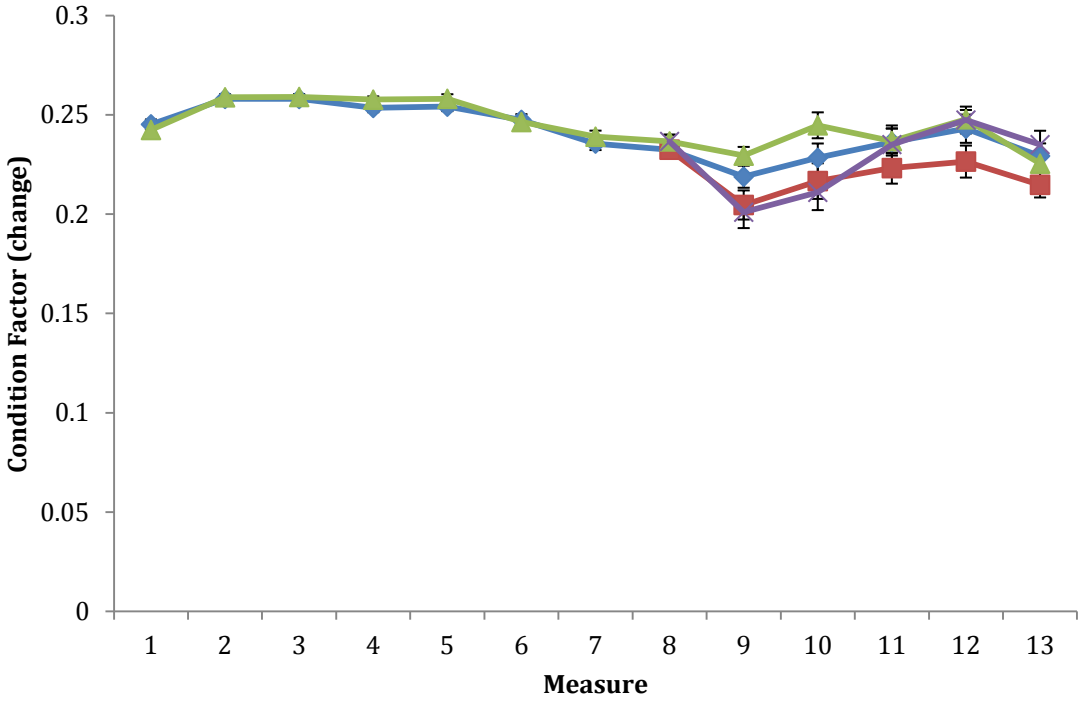
a)



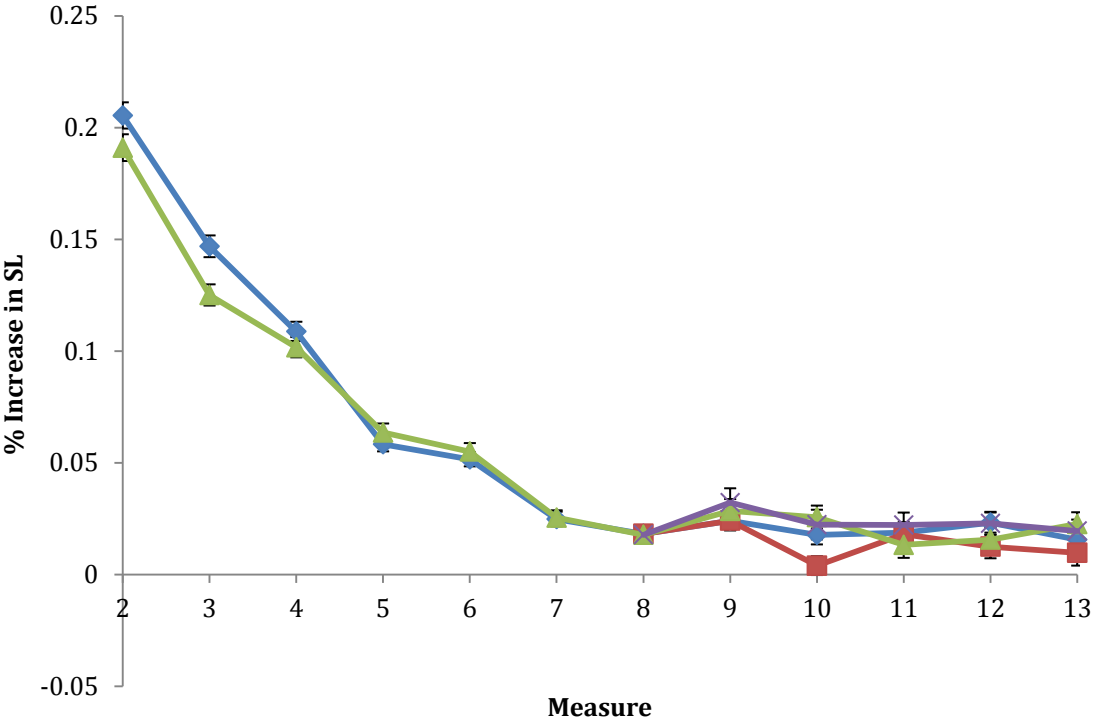
b)



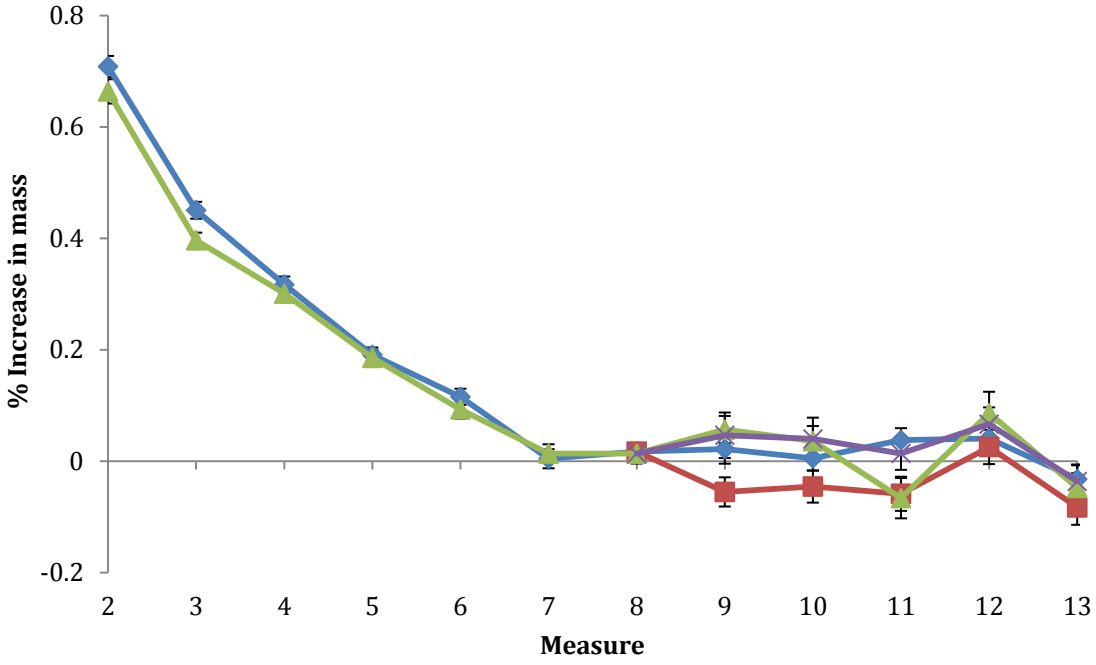
c)


0.3
0.25
0.2
0.15
0.1
0.05
0
Condition Factor (change)
1
2
3
4
5
6
7
8
9
10
11
12
13
Measure

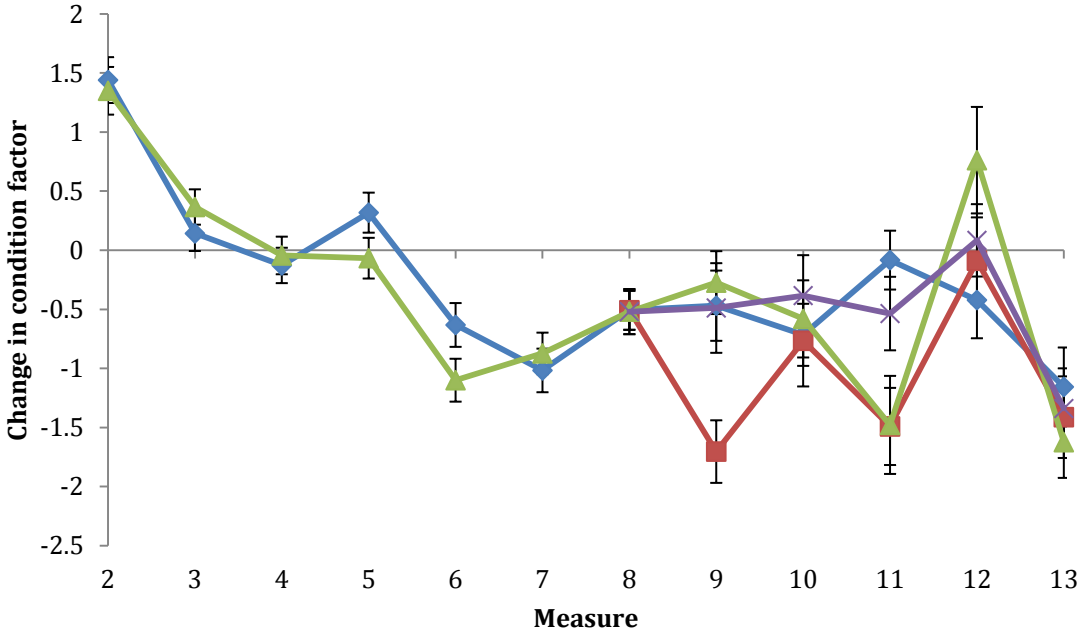
d)


0.25
0.2
0.15
0.1
0.05
0
-0.05
% Increase in SL
2
3
4
5
6
7
8
9
10
11
12
13
Measure

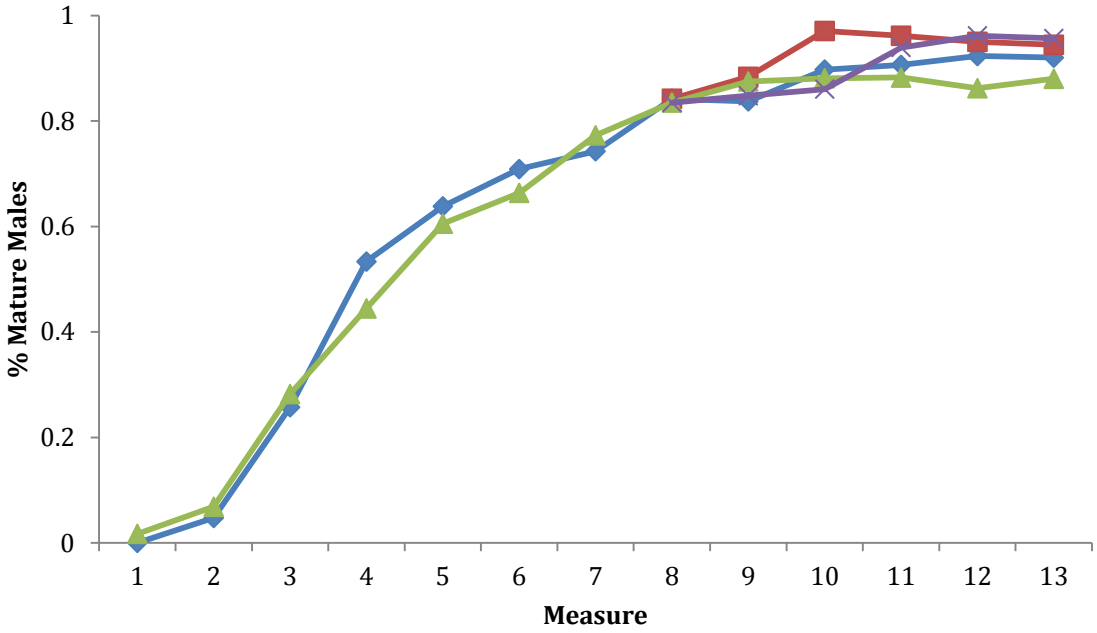
e)


0.8
0.6
0.4
0.2
0
-0.2
% Increase in mass
2
3
4
5
6
7
8
9
10
11
12
13
Measure

f)


2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
Change in condition factor
2
3
4
5
6
7
8
9
10
11
12
13
Measure

g)


% Mature Males
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13
Measure

h)

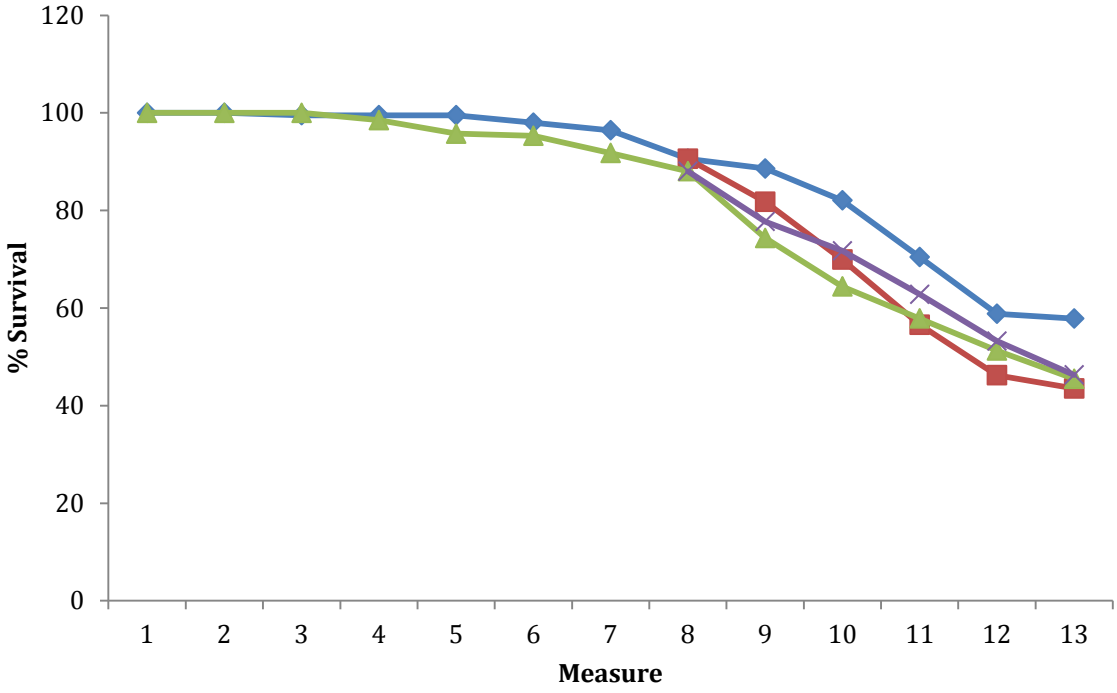

% Survival
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13
Measure

Table 5.1 Estimated size of fixed effects (coefficients) with standard errors (SE) from univariate animal models for each response variable (measured or derived). Conditional F statistics were used to assess significance (P). Since stack, trial and the early-life/late-life density $(ELD*LLD)_{p2}$ interactions are multi-level factors, coefficients are not presented ("-"). However, coefficients are given for individuals of unknown sex, illustrating where these are driving significance values (i.e. SGR_{SL} and SGR_{WT} are significantly lower in males and individuals of unknown sex; CCF is significantly lower in males than females, but is probably not significant in sex unknown individuals; significant sex effects for longevity appear to be driven by shorter living individuals of unknown sex).

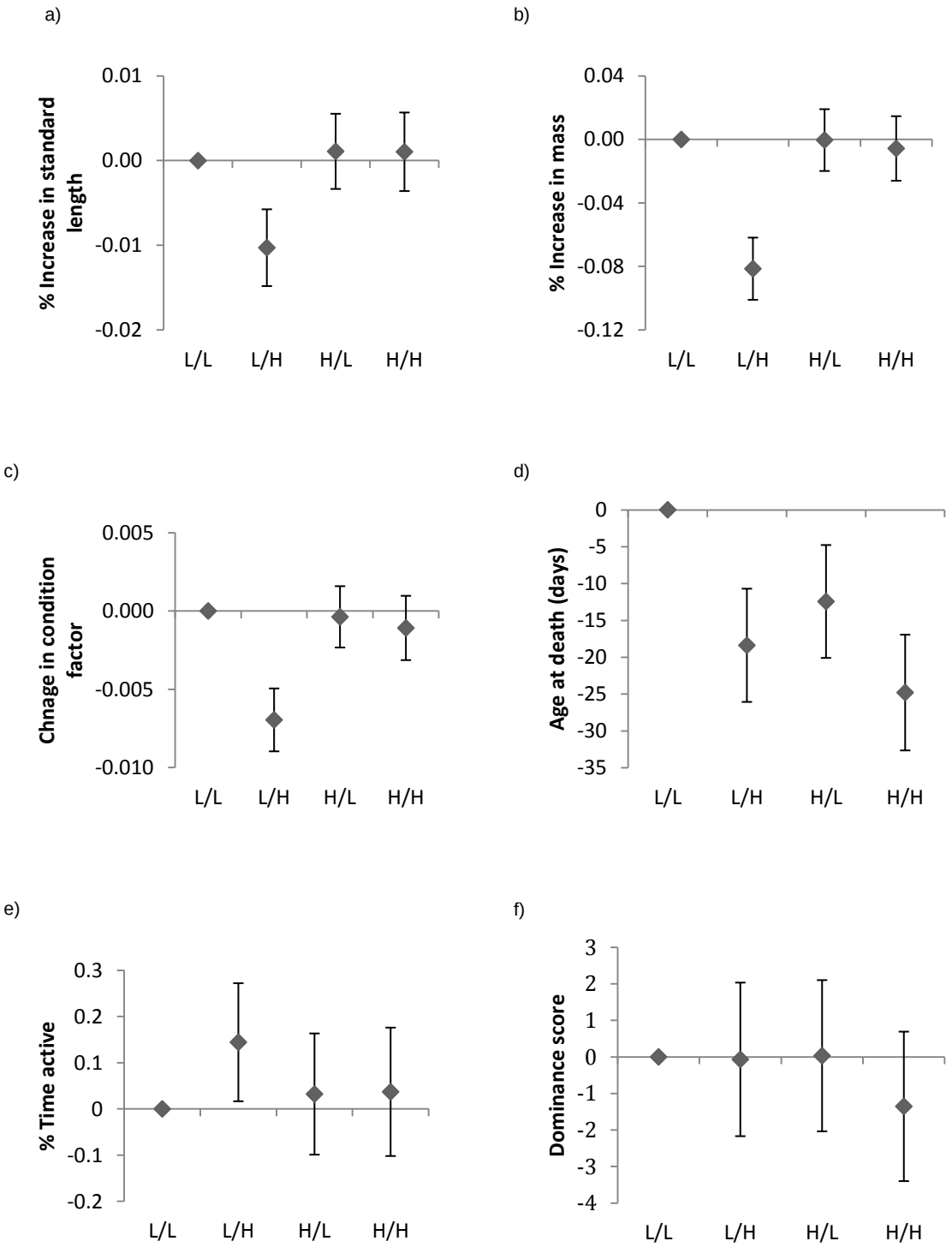
Trait	Effect	Coefficient (SE)	DF	F	P
SGR_{SL}	Mean	0.110 (0.006)	1,17.1	124.98	<0.001
	Stack	-	5,96.7	30.47	<0.001
	Measure age	-6.61×10^{-4} (2.99×10^{-5})	1,3682.1	487.34	<0.001
	Measure age^2	2.39×10^{-6} (1.31×10^{-7})	1,3740.4	330.53	<0.001
	Measure age^3	-6.09×10^{-9} (1.03×10^{-9})	1,3546.8	34.89	<0.001
	Sex (male)	-0.012 (0.002)	2,368.3	16.34	<0.001
	Sex (unknown)	-0.016 (0.007)			
	gmCOMP	-0.010 (0.002)	1,1617.5	17.08	<0.001
	Sex ratio	-0.008 (0.005)	1,1182.9	2.93	0.09
	ELD_{p1} (high)	-0.005 (0.002)	1,1640.3	32.51	<0.001
	$(ELD*LLD)_{p2}$	-	3,1753.7	2.77	0.04
SGR_{WT}	Mean	0.382 (0.026)	1,15.8	23.21	<0.001
	Stack	-	5,96.6	31.4	<0.001
	Measure age	-0.002 (0.0001)	1,3663.1	380.39	<0.001
	Measure age^2	8.81×10^{-6} (5.50×10^{-9})	1,3723.6	256.75	<0.001
	Measure age^3	-2.91×10^{-8} (4.29×10^{-9})	1,3519.4	46.17	<0.001
	Sex (male)	-0.060 (0.010)	2,375.1	17.9	<0.001
	Sex (unknown)	-0.082 (0.033)			
	gmCOMP	-0.044 (0.011)	1,2151.5	16.97	<0.001
	Sex ratio	-0.010 (0.022)	1,1518.2	0.2	0.648
	ELD_{p1} (high)	-0.020 (0.011)	1,1858.1	26.41	<0.001
	$(ELD*LLD)_{p2}$	-	3,1962.7	7.71	<0.001
CCF	Mean	0.008 (0.002)	1,3771	23.03	<0.001
	Stack		5,3771	10.98	<0.001
	Measure age	-6.17×10^{-5} (1.38×10^{-5})	1,3771	20.04	<0.001
	Measure age^2	2.59×10^{-7} (6.06×10^{-8})	1,3771	18.29	<0.001
	Measure age^3	-1.65×10^{-9} (4.79×10^{-10})	1,3771	11.81	<0.001
	Sex (male)	-0.003 (0.001)	2,3771	6.1	0.002
	Sex (unknown)	-0.004 (0.003)			
	gmCOMP	-0.002 (0.001)	1,3771	5.05	0.026
	Sex ratio	0.002 (0.002)	1,3771	1.21	0.274
	ELD_{p1} (high)	-0.001 (0.001)	1,3771	3.31	0.037
	$(ELD*LLD)_{p2}$	-	3,3771	5.03	0.002

Cont...

Table 5.1 cont...

Trait	Effect	Coefficient (SE)	DF	F	P
$AGEmat_M$	Mean	266 (9.50)	1,11.9	2812	<0.001
	Stack	-	5,53	6.09	<0.001
	gmCOMP	-102 (13.8)	1,179.8	55.1	<0.001
	Sex ratio	96.0 (15.0)	1,175.2	40.8	<0.001
	ELD (high)	-7.47 (5.67)	1,171.5	1.74	0.192
$SLmat_M$	Mean	37.7 (0.668)	1,13.6	12084	<0.001
	Stack	-	5,137.3	1.25	0.29
	gmCOMP	-0.408 (0.939)	1,177.3	0.19	0.659
	Sex ratio	5.53 (1.024)	1,176.3	29.2	<0.001
	ELD (high)	-0.517 (0.386)	1,171.3	1.79	0.185
$WTmat_M$	Mean	1.46 (0.082)	1,13.5	1196	0.306
	Stack	-	5,133.1	3.60	0.004
	gmCOMP	0.023 (0.119)	1,179.8	0.05	0.812
	Sex ratio	0.624 (0.129)	1,178.9	23.4	<0.001
	ELD (high)	-0.066 (0.049)	1,173.4	1.84	0.18
$CFmat_M$	Mean	2.59 (0.061)	1,11.4	9690	<0.001
	Stack	-	5,59.6	6.48	<0.001
	gmCOMP	0.168 (0.088)	1,181.6	3.59	0.061
	Sex ratio	-0.026 (0.097)	1,179.8	0.07	0.784
	ELD (high)	-0.020 (0.037)	1,171.7	0.30	0.582
Longevity	Mean	473 (9.55)	1,13.5	12263	<0.001
	Stack	-	5,181.5	15.0	<0.002
	Sex (male)	-8.22 (6.20)	2,368.6	6.16	0.002
	Sex (unknown)	-58.6 (17.0)			
	gmCOMP	-41.7 (3.47)	1,370.4	144	<0.001
	Sex ratio	-9.72 (11.1)	1,366.3	0.076	0.384
	ELD (high)	-12.4 (7.66)	1,357.4	2.96	0.089
	LLD (high)	-18.4 (7.68)	1,359.3	8.02	0.005
	ELD*LLD	6.03 (11.0)	1,361.9	0.30	0.579
Activity	Mean	7.47 (0.136)	1,11.8	11580	<0.001
	Stack	-	5,192.7	8.05	<0.001
	Sex (male)	-0.206 (0.077)	1,351.7	7.01	0.009
	gmCOMP	0.022 (0.111)	1,1156.6	0.04	0.834
	Sex ratio	-0.133 (0.185)	1,1072	0.52	0.47
	Day order	-0.008 (0.003)	1,1197.6	9.88	0.002
	Trial	-	3,1025.5	90.2	<0.001
	ELD_{p1} (high)	-0.048 (0.083)	1,1069.7	0.51	0.472
	$(ELD*LLD)_{p2}$		3,1056.7	0.47	0.702
$Dominance_M$	Mean	3.54 (2.00)	1,8	4.00	0.081
	Stack	-	5,54.6	0.43	0.824
	gmCOMP	0.979 (1.36)	1,706.3	0.52	0.469
	Sex ratio	-0.446 (2.09)	1,508.5	0.05	0.825
	Day order	-0.037 (0.993)	1,798.6	0.59	0.44
	Trial	-	4,691.9	0.77	0.547
	ELD_{p1} (high)	-0.366 (1.11)	1,512.4	0.04	0.832
	$(ELD*LLD)_{p2}$	-	3,718.3	0.23	0.874

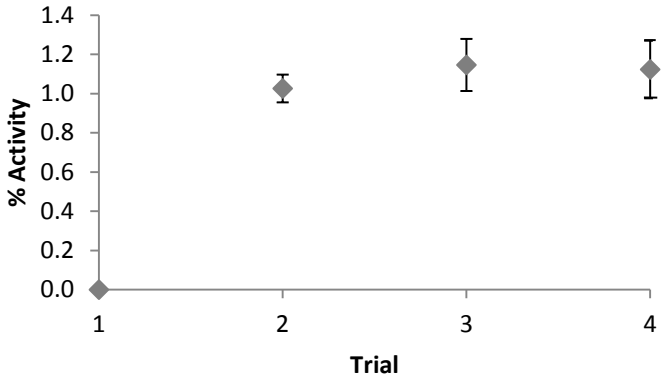
Fig. 5.3 Estimated effects of the density treatment regimes (Low/Low (L/L); Low/High (L/H); High/Low (H/L); High/High (H/H)) with bars depicting standard error on a) specific growth rate for standard length (SL) per day (SGR_{SL}); b) specific growth rate (mass) per day (SGR_{WT}); c) change in condition factor per unit time (CCF); d) longevity ; e) activity ; f) male dominance. Effect sizes are shown relative to that of the reference treatment level mean (L/L).



5.4.1.3 Environmental effects on behaviour

We found no evidence of density treatment effects on either activity in the OFT or male dominance score determined from the in-tank observations. Indeed, for male dominance none of the fixed effects included in our univariate models explained significant variance (Table 5.1). We did find stack effects on activity, and a significant effect of trial number. The latter was driven by relatively higher activity levels at trials 2-4 than at trial 1 (Fig. 5.4). This could reflect habituation although we also note that trial number is confounded with age for this trait. Our analysis also highlighted that males are less active than females (sex effect = -0.206 ± 0.078, P = 0.009) with individuals of unknown sex not differing from females. Fish also tended to be less active earlier in the day (day order effect = -0.008 ± 0.003, P = 0.002).

Fig. 5.4 Estimated effect size (with error bars) of trial number on Activity relative to the reference factor level of Trial 1. Trials 1-2 were performed prior to density swap, and trials 3-4 after density swap.



5.4.2 Repeatabilities, heritabilities and tests of GxE

After conditioning on fixed effects, we found small but significant among-individual variance for growth traits, (SGR_{SL} repeatability = 0.058 ± 0.012, X^2_1 = 37.58, P <0.001; SGR_{WT} repeatability = 0.089 ± 0.013, X^2_1 = 78.24, P <0.001) but not for CCF. Note that repeatabilities for growth are much lower than those obtained for the corresponding size traits when conditioned on the same fixed effects, (R_{SL} = 0.564 ± 0.020, X^2_1 = 2401.28, P <0.001; R_{WT} 0,447 ± 0.021, X^2_1 = 1685.51, P <0.001, full results not presented). The two personality traits also had significant among-individual variance (Table 5.2) with repeatability estimates for activity of 0.240 ± 0.032, X^2_1 = 75.11, P

<0.001 and male dominance score of 0.267 ± 0.039, X^2_1 = 71.38, P <0.001, (Table 5.2). Note also that repeatability for activity differs marginally from our previously published estimate obtained using the same data (Boulton et al., 2014 – Chapter 2 because it is conditioned on a slightly different set of fixed effects.

Table 5.2 Univariate estimates of among-individual (V_I) and residual variance (V_R) and repeatability (R) with standard error (SE) and significance value (P) for traits with repeated measures: specific growth rate for length and mass (SGR_{SL}, SGR_{WT}); change in condition factor per unit time (CCF); activity (ACT, % time active in open field trial); male dominance score (DOM_M from in-tank observations). The among- (V_I) and within-individual (residual) variance (V_R) estimates are presented for each trait along with repeatability (R). X^2_1 and P-values relate to likelihood ratio tests of the significance of V_I. Note that for univariate models only we assume the test statistic to be asymptotically distributed as a 50:50 mix of X^2_0 and X^2_1 (following Visscher 2006). Where a variance is bound to zero, the standard error and significance cannot be estimated (NE).

Trait	V_I	V_R	R	X^2_1	P
SGR_{SL}	1.69 x^{10-4} (3.46 x^{10-5})	0.003 (6.70 x^{10-5})	0.058 (0.012)	37.58	<0.001
SGR_{WT}	4.64 x^{10-3} (7.29 x^{10-4})	0.047 (0.001)	0.089 (0.013)	78.24	<0.001
CCF	0.000 (NE)	0.000 (NE)	0.000 (NE)	0.00	0.500
DOM_M	32.8 (5.78)	90.1 (5.16)	0.267 (0.039)	75.11	<0.001
ACT	0.247 (0.038)	0.783 (0.038)	0.240 (0.032)	71.38	<0.001

Further partitioning the variance in our animal model analyses revealed statistical support for additive genetic and brood tank effects on some, but not all, traits. Brood tank effects were significant for SGR_{SL}, and SGR_{WT} only. These growth traits also had significant (additive) genetic variance (Table 5.3) with low but significant heritability estimates (h_2 for SGR_{SL} = 0.043 ± 0.024, X^2_1 = 9.36, P = 0.001; h_2 for SGR_{WT}, = 0.024 ± 0.019, X^2_1 = 3.78, P = 0.026). Again we note that heritabilities of growth are much lower than those for the size traits that they were derived from (h^2_{SL} = 0.132 ± 0.079, X^2_1 = 5.12, P = 0.012; h^2_{WT} = 0.132 ± 0.084, X^2_1 = 3.98, P = 0.023, models not presented). Unsurprisingly, given the lack of significant repeatability for CCF, this trait was not heritable though condition factor itself is (h^2_{CF} = 0.067 ± 0.039, X^2_1 = 7.38, P = 0.003). Among the life history traits we found evidence of moderately high heritabilities for size at maturity in males (h_2 for $SLmat_M$ = 0.298 ± 0.164, X^2_1 = 6.72, P = 0.009; h_2 for $WTmat_M$ = 0.213 ± 0.139, X^2_1 = 3.92, P = 0.024). Heritability of male maturation age was marginally non-significant, as was that for longevity (Table 5.3). Of the personality traits, we found evidence for significant heritability of activity, our proxy measure of boldness (h^2 = 0.088 ± 0.054, X^2_1 = 5.50, P = 0.020) but not male dominance (Table 5.3).

Finally, we note that we found no evidence for genotype by environment (ELD treatment) effects (GxE) on any trait (results not shown). Some variance was partitioned to the GxE term but was non-significant for behavioural and male size at maturity traits. For all other traits, V_{GxE} was bound to zero.

Table 5.3 Univariate analyses of observed traits. Estimated variance with standard errors (SE) for brood tank (V_{BT}), permanent environment (V_{PE}), additive genetic (V_A), and residual (V_R) effects with estimated brood tank (bt^2) permanent environment (pe^2), and heritability (h^2). X^2_1 and P-values relate to likelihood ratio tests for the significance of V_{BT}, V_{PE} and V_A. Traits measured: specific growth rate for length and mass (SGR_{SL}, SGR_{WT}); change in condition factor (CF) per unit time (CCF); activity (ACT, % time active in open field trial); male dominance score (DOM_M from in-tank observations); male age at maturation ($AGEmat_M$), male length at maturation ($SLmat_M$); male mass at maturation ($WTmat_M$); male CF at maturation ($CFmat_M$); longevity in days (LONG, measured from date of entry to study). Where a variance is bound to zero, the standard error cannot be estimated (NE). NF indicates term not fitted in model.

Trait	V_R (SE)	V_{BT} (SE)	V_{PE} (SE)	V_A (SE)	bt^2 (SE)	P	pe^2 (SE)	P	h^2 (SE)	P
SGR_{SL}	2.76×10^{-03} (6.70×10^{-05})	4.40×10^{-06} (2.52×10^{-05})	8.44×10^{-06} (4.66×10^{-05})	1.27×10^{-4} (7.11×10^{-5})	0.015 (0.009)	0.007	0.003 (0.016)	0.421	0.043 (0.024)	0.001
SGR_{WT}	0.048 (0.001)	9.22×10^{-04} (5.34×10^{-04})	2.60×10^{-03} (8.44×10^{-04})	1.25×10^{-3} (9.96×10^{-4})	0.018 (0.010)	0.011	0.050 (0.016)	0.460	0.024 (0.019)	0.026
CCF	0.000 (NE)	0.000 (NE)	0.000 (NE)	0.000 (NE)	0.000 (NE)	0.500	0.000 (NE)	0.500	0.000 (NE)	0.500
ACT	0.784 (0.038)	0.000 (NE)	0.170 (0.010)	0.092 (0.056)	0.000 (NE)	0.500	0.162 (0.049)	0.016	0.088 (0.054)	0.010
DOM_M	90.220 (5.170)	6.87 (5.16)	23.3 (7.22)	3.56 (7.98)	0.055 (0.041)	0.056	0.188 (0.057)	0.015	0.029 (0.032)	0.320
$AGEmat_M$	1254.200 (213.220)	23.6 (113)	NF	302 (258)	0.015 (0.072)	0.415	NF	NF	0.191 (0.155)	0.055
$SLmat_M$	5.350 (1.110)	2.62×10^{-06} (5.44×10^{-07})	NF	2.36 (1.44)	NE	0.500	NF	NF	0.306 (0.166)	0.008
$WTmat_M$	0.092 (0.016)	1.02×10^{-08} (1.76×10^{-09})	NF	0.003 (0.002)	NE	0.500	NF	NF	0.220 (0.140)	0.021
$CFmat_M$	0.054 (0.008)	0.004 (0.005)	NF	0.008 (0.009)	0.058 (0.080)	0.207	NF	NF	0.117 (0.138)	0.160
LONG	2720.800 (231.950)	0.000 (NE)	NF	147 (149)	NE	0.500	NF	NF	0.051 (0.051)	0.068

Table 5.4 Estimated within – (**R**) and among- individual (**PE**) phenotypic variance-covariance matrices obtained from multivariate analysis of the phenotypic traits (in standard deviation units): specific growth rate for length (SGR_{SL}); male dominance score (DOM_M from in-tank observations), activity (ACT, % time active in open field trial); male age at maturation ($AGEmat_M$), male length at maturation ($SLmat_M$); male condition factor at maturation ($CFmat_M$); longevity in days (LONG, measured from date of entry to study). Variances are presented on the diagonal (shaded in grey), between-trait covariances are below and between-trait correlations are above the diagonals, with standard errors in parentheses. "NF" indicates variance not fitted, while "-" indicates parameter not estimated.

R	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.437 (0.011)	-	-	-	-	-	-
DOM_M	-	0.743 (0.042)	-	-	-	-	-
ACT	-	-	0.793 (0.038)	-	-	-	-
$AGEmat_M$	-	-	-	NF	-	-	-
$SLmat_M$	-	-	-	-	NF	-	-
$CFmat_M$	-	-	-	-	-	NF	-
LONG	-	-	-	-	-	-	NF
PE	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.034 (0.006)	0.601 (0.120)	-0.258 (0.114)	0.326 (0.105)	0.633 (0.091)	0.543 (0.098)	0.678 (0.067)
DOM_M	0.058 (0.013)	0.269 (0.047)	-0.423 (0.119)	-0.074 (0.097)	0.444 (0.083)	0.173 (0.094)	0.296 (0.096)
ACT	-0.023 (0.011)	-0.108 (0.032)	0.241 (0.038)	0.363 (0.094)	-0.252 (0.096)	-0.227 (0.099)	0.074 (0.083)
$MATage_M$	0.045 (0.015)	-0.029 (0.038)	0.133 (0.037)	0.558 (0.058)	0.295 (0.067)	-0.285 (0.068)	0.311 (0.070)
$AGEmat_M$	0.107 (0.019)	0.210 (0.047)	-0.113 (0.045)	0.201 (0.052)	0.834 (0.086)	0.161 (0.071)	0.200 (0.073)
$SLmat_M$	0.094 (0.020)	0.084 (0.047)	-0.104 (0.047)	-0.199 (0.053)	0.137 (0.063)	0.868 (0.091)	0.202 (0.075)
$CFmat_M$	0.098 (0.014)	0.120 (0.042)	0.028 (0.032)	0.182 (0.046)	0.143 (0.055)	0.148 (0.057)	0.614 (0.045)

Table 5.5 Estimated residual (**R**) permanent environment (**PE**) and additive genetic (**G**) variance-covariance matrices obtained from multivariate animal model analyses of the phenotypic traits (in standard deviation units) specific growth rate for length (SGR_{SL}); male dominance score (DOM_M from in-tank observations), activity (ACT, % time active in open field trial); male age at maturation ($AGEmat_M$), male length at maturation ($SLmat_M$); male condition factor at maturation ($CFmat_M$); longevity in days (LONG, measured from date of entry to study). Variances are presented on the diagonal (shaded in grey), between-trait covariances are below and between-trait correlations are above the diagonals, with standard errors in parentheses. Note that models were formulated such that **R** contains within-individual, and **PE** contains among-individual sources of environmental variance. The former are only identifiable for traits with repeat measures (see text for details). A brood tank effect (**BT**) was fitted for SGR_{SL} only. It was not possible to obtain model convergence with all seven traits modelled simultaneously, therefore the matrices presented here are compiled from a series of smaller models containing up to six traits (see text for full details). NF indicates variance not fitted, while "-" indicates parameter not estimated. [1] these estimated genetic correlations were further tested using bivariate models (Table 5.6, see text for details).

R	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.433 (0.010)	-	-	-	-	-	-
DOM_M	-	0.741 (0.042)	-	-	-	-	-
ACT	-	-	0.793 (0.038)	-	-	-	-
$AGEmat_M$	-	-	-	NF	-	-	-
$SLmat_M$	-	-	-	-	NF	-	-
$CFmat_M$	-	-	-	-	-	NF	-
LONG	-	-	-	-	-	-	NF
PE	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.015 (0.007)	0.622 (0.279)	-0.337 (0.251)	0.277 (0.230)	0.809 (0.241)	0.492 (0.020)	0.727 (0.142)
DOM_M	0.034 (0.016)	0.198 (0.066)	-0.268 (0.223)	-0.005 (0.173)	0.286 (0.166)	0.122 (0.161)	0.396 (0.139)
ACT	-0.017 (0.013)	-0.052 (0.045)	0.165 (0.050)	0.239 (0.171)	-0.131 (0.183)	-0.147 (0.170)	0.106 (0.129)
$AGEmat_M$	0.022 (0.018)	-0.001 (0.050)	0.062 (0.050)	0.413 (0.076)	0.375 (0.126)	-0.341 (0.109)	0.325 (0.091)
$SLmat_M$	0.077 (0.023)	0.098 (0.066)	-0.041 (0.059)	0.185 (0.070)	0.592 (0.121)	0.204 (0.125)	0.251 (0.104)
$CFmat_M$	0.051 (0.023)	0.046 (0.062)	-0.050 (0.059)	-0.183 (0.067)	0.131 (0.082)	0.700 (0.111)	0.132 (0.096)
LONG	0.071 (0.015)	0.134 (0.046)	0.031 (0.038)	0.160 (0.050)	0.146 (0.061)	0.083 (0.063)	0.569 (0.050)
G	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.011 (0.008)	0.421 (0.502)	-0.214 (0.428)	-0.067 (0.457)	0.272 (0.418)	[1]0.750 (0.387)	[1]0.843 (0.429)
DOM_M	0.012 (0.017)	0.073 (0.074)	[1]-0.804 (0.442)	-0.502 (0.496)	[1]0.788 (0.434)	0.293 (0.622)	-0.341 (0.624)
ACT	-0.007 (0.014)	-0.064 (0.050)	0.087 (0.056)	[1]0.751 (0.275)	-0.520 (0.385)	-0.420 (0.495)	0.177 (0.522)
$AGEmat_M$	-0.003 (0.019)	-0.052 (0.056)	0.085 (0.056)	0.148 (0.093)	-0.029 (0.438)	-0.204 (0.466)	0.280 (0.506)
$SLmat_M$	0.015 (0.025)	0.109 (0.080)	-0.078 (0.068)	-0.006 (0.085)	0.259 (0.155)	-0.066 (0.483)	-0.197 (0.502)
$CFmat_M$	0.032 (0.023)	0.031 (0.065)	-0.049 (0.062)	-0.031 (0.077)	-0.013 (0.097)	0.156 (0.117)	0.555 (0.443)
LONG	0.020 (0.014)	-0.020 (0.037)	0.012 (0.036)	0.021 (0.043)	-0.021 (0.055)	0.048 (0.050)	0.040 (0.036)
BT	SGR_{SL}	DOM_M	ACT	$AGEmat_M$	$SLmat_M$	$CFmat_M$	LONG
SGR_{SL}	0.008 (0.004)	-	-	-	-	-	-

5.4.3 *MULTIVARIATE ANALYSES*

Comparing models (A) and (B) for the set of seven traits retained provided strong support for the presence of phenotypic covariance between traits at the among-individual level (X^2_{21} = 255.16, P <0.001). Almost all pairwise correlation estimates in the **PE** matrix estimated under model (B) are nominally significant at α=0.05 based on |r|≥ 2SE (Table 5.4). Only the correlations between DOM_M and $AGEmat_M$, DOM_M and $CFmat_M$ and between ACT and LONG appear to be non-significant based on this assumption. Individuals with consistently high growth (SGR_{SL}) tend (if male) to have greater dominance scores (r_{PE} = 0.601 ± 0.120). They also tend to mature at a larger size and in better condition (r_{PE} of SGR_{SL} and $SLmat_M$ = 0.633 ± 0.091; r_{PE} of SGR_{SL} and $CFmat_M$ = 0.543 ± 0.098) but at a later age (r_{PE} of SGR_{SL} and $AGEmat_M$ = 0.326 ± 0.105) and live longer lives (r_{PE} of SGR_{SL} and LONG = 0.678 ± 0.067). Contrary to our expectation, activity (ACT, our proxy for boldness) is actually negatively correlated with social dominance (r_{PE} = -0.423 ± 0.119). More active individuals also have lower maturation size and condition, but higher maturation age (r_{PE} between ACT and $AGEmat_M$ = 0.363 ± 0.094) (Table 5.4). All traits were positively phenotypically correlated with LONG at the among-individual level (Table 5.4) though the correlation was weak (and not significant) for ACT. Thus, to the extent that longevity is a valid proxy for fitness, selection through mortality (under these experimental conditions) favours fast growing individuals, with high dominance scores that mature late at large size.

Despite reducing the number of traits from ten to seven, we were still unable to obtain a stable model convergence with the full multivariate animal model. However, convergence was achieved for a model of all remaining traits excluding longevity. For this set of six traits, LRT between the model with no **G** fitted (model (B)) and that with a diagonal **G** matrix (i.e. V_A for each trait but no genetic covariances, model (C)) showed the improvement in fit to be marginally non-significant X^2_6 = 12.48, P = 0.052). LRT between models (B) and (D), (no **G** versus full **G**) presented a significant improvement in model fit (X^2_{21} = 34.58, P = 0.031), but model (C) was not significantly better than model (D) (X^2_{15} = 22.1, P = 0.105). Taken together we interpret these results as supporting the presence of variance in **G** for this set of traits but as not providing robust statistical support for genetic contributions to the phenotypic covariance that is present among traits. To complete our estimation of the full **G** and **PE** structures between traits under model (C) we fitted additional models to obtain parameters

relating to longevity. Specifically variance for longevity, and covariance and correlations between longevity and all other traits apart from SGR_{SL} was obtained using a five trait model, (DOM_M, ACT, $AGEmat_M$, $SLmat_M$, LONG), while covariance between SGR_{SL} and LONG was achieved via a trivariate model including $AGEmat_M$. Covariance and correlations between longevity and male maturation condition factor were obtained from a model fitting longevity with the three male maturation traits. Estimates of **G** and **PE** among all seven traits were compiled in this way and are presented in Table 5.5. Since traits were scaled to standard deviation units for multivariate analysis the estimates of V_A on the diagonal of **G** can actually be interpreted as heritabilities, but are not conditional on fixed effects (and therefore differ somewhat from those presented in Table 5.3). Note also that the **PE** matrix in Table 5.5 has a different interpretation to that in Table 5.4. Specifically, it now characterises the portion of between-trait among-individual phenotypic covariance that is not due to (additive) genetic effects (or brood tank effects in the case of SGR_{SL}).

We reiterate that, at least for the set of six traits that could be analysed simultaneously, multivariate comparisons do not provide robust statistical support for significant covariance in the **G** matrix overall. However, we note that while large standard errors suggest that power is generally limiting (Table 5.5) in five cases, r_G estimates are at or approaching $|r_G|>2SE$ (Table 5.5). Interpretation is complicated because in all such cases, one trait involved in the genetic correlations had a non-significant estimate of V_A in the univariate analysis. To further explore these cases, we fitted (for each pair of traits) additional bivariate animal models: (i) additive genetic variance was included for trait 1 only (where trait 1 was significantly heritable according to univariate model results, (ii) additive genetic variance was included for both traits (assuming $COV_A = 0$) and (iii) COV_A was also fitted (and the estimate rescaled to r_G).

The results of these *post hoc* tests are presented in Table 5.6. For three of the five trait pairs tested, model (iii) was a significantly better fit to the data than model (i). Therefore we conclude that, despite lack of significant V_A in univariate analyses, there is actually support for genetic effects on $AGEmat_M$, $CFmat_M$ and LONG. Furthermore, significant positive genetic correlations are found between growth rate (SGR_{SL},) and both $CFmat_M$, and LONG, (model (iii) versus (ii) comparisons, Table 5.6), while r_G between ACT and $CFmat_M$ was marginally non-significant (bivariate model estimate, r_G = 0.751 (0.261), model (iii) versus (ii) comparison $X^2_1 = 3.406$, P=0.065). Bivariate analyses involving DOM_M yielded less clear cut results since, while the strong positive

genetic correlation with $SLmat_M$ was significant (model (iii) versus (ii) comparison $X^2_1 = 5.034$, $P = 0.025$), model (iii) was not a significantly better fit than model (i) where genetic effects were included on $SLmat_M$ only. We therefore conclude that significant genetic effects on DOM_M are not supported by this analysis (in agreement with univariate models). However, we note that if male dominance is truly heritable, then genotypes associated with high DOM_M appear also to be associated with larger size at maturity and lower ACT (although the latter relationship is marginally non-significant; model (iii) versus (ii) comparison $X^2_1 = 3.44$, $P = 0.064$, Table 5.6).

Table 5.6 Estimates of between-trait correlations, r_G with standard errors (SE) for those traits with marginal significance from the **G** matrix in Table 5.5. For trait 1, heritability was supported in the univariate analyses (see Table 5.1), while for Trait 2 it was not. Estimates of chi squared (X^2_{DF}) are from likelihood ratio tests between the models (i, ii, iii) as annotated in the text.

		Log likelihood			(ii vs i)		(iii vs ii)		(iii vs i)		
Trait 1	Trait 2	(i)	(ii)	(iii)	X^2_1	P	X^2_1	P	X^2_2	P	r_G (SE)
ACT	DOM_M	-1030.840	-1030.630	-1028.910	0.420	0.517	3.44	0.064	3.86	0.145	-0.926 (0.442)
$SLmat_M$	DOM_M	-483.392	-483.181	-480.664	0.422	0.516	5.03	0.025	5.46	0.065	0.889 (0.343)
ACT	$AGEmat_M$	-674.564	-672.002	-670.299	5.12	0.024	3.41	0.065	8.53	0.014	0.751 (0.261)
SGR_{SL}	$CFmat_M$	-567.201	-565.995	-563.708	2.41	0.120	4.57	0.032	6.99	0.030	0.812 (0.243)
SGR_{SL}	LONG	-583.473	-582.787	-578.746	1.37	0.241	8.08	0.004	9.45	0.009	0.863 (0.270)

5.5 DISCUSSION

The aims of this study were to ascertain the effects of competition on growth, life history and personality traits and to investigate the genetic covariance between traits related to social dominance. With these objectives in mind we exposed a pedigreed population of a small tropical fish, *Xiphophorus birchmanni*, (64 groups of eight, n = 384) to different levels of competition by manipulating housing density. Data on morphology, behaviour, life history and fitness (longevity) were collected over the 50 week period of the study. These were analysed to test the effects of competitive regime on mean phenotype and to determine the extent that phenotypic (co)variation was attributable to genetic effects.

5.5.1 THE EFFECTS OF INCREASED COMPETITION ON PHENOTYPE AND FITNESS

As predicted, we found evidence that density (i.e. level of competition for space) influenced phenotypes and fitness. For example, as measured by changes in both weight and standard length, individual growth rates during early life (i.e. part 1 of the study) were reduced by experiencing high density (as was the rate of condition factor increase). This is consistent with the widespread reporting of density dependent growth rates in fishes (e.g. Rothschild, 1986; Lorenzen and Enberg, 2002; Hixon et al., 2012). We note that in addition to the effects of our experimental density treatment, further evidence for the reduction of growth and condition from competition was provided by significant negative effects of the (geometric) mean number of competitors. This effect was included to account for changes in density experienced due to mortality within groups. Significant density treatment effects on later life growth were also found, and were driven in particular by reduced growth rates in fish that experienced the LH regime. Thus, it seems that switching from a low to a high competition environment part way through development may impose greater challenges to growth and condition factor than consistently experiencing high density. Conversely, individuals experiencing the HL regime actually had the greatest mean size at the end of the experiment (as seen in Figure 5.2a and b for standard length and mass respectively). This pattern is broadly consistent with some form of compensatory growth, a widely reported phenomenon in fishes entailing a phase of accelerated growth following a period of growth depression, usually when favourable conditions are restored (e.g. Metcalfe and Monaghan, 2001; Ali et al., 2003; Royle et al., 2005).

There was less evidence that our density treatment had major effects on personality or life history. We found no effect on boldness or maturation, results that contrast with a

number of other studies showing effects of early environment on personality (Niemelä et al., 2012; Patrick et al., 2013) and life history (Rowe and Thorpe, 1990). Neither were there effects on dominance score but this was not unexpected since dominance was assessed within groups (i.e. among males experiencing the same treatment regime). Given the expected close links between growth and maturation in fishes (Sohn and Crews, 1977; Snelson, 1984; Godø and Moksness, 1987; Rowe and Thorpe, 1990; Adams and Huntingford, 1997; Morita and Fukuwaka, 2006), it was somewhat surprising that there were no significant effects of density treatment on male maturation, especially as growth was affected and among-individual correlations were found for these traits (discussed below). However, male maturation age was negatively impacted by increasing (geometric) mean number of competitors. This suggests that the number of interacting competitors (rather than density treatment *per se*) may be important. Previous studies on male maturation in *Xiphophorus sp*. have also noted strong effects arising from the composition of social group such as sex ratio or the presence of dominant males; (e.g. Borowsky, 1978; Borowsky, 1987; Campton, 1992; Walling et al., 2007). Here we find that males experiencing a more male biased group of competitors matured on average later and at larger size. This type of plasticity may well be adaptive if males need to be bigger to compete successfully with rivals (e.g. Ryan et al., 1990; Morris et al., 1992; Preston et al., 2003). In swordtails, male growth declines dramatically at maturation (Basolo, 1988; Walling et al., 2007) so it is expected that increased male-male competition will result in sexual selection that favours larger maturing males, even if this comes at the cost of a delayed maturation time (e.g. Basolo, 1988; Alcock and Houston, 1996; Beaugrand et al., 1996; Benson and Basolo, 2006). Therefore, while our results do not demonstrate significant density treatment effects, there is some support for competitive effects on life history arising from the number and sex ratio of competitors experienced.

We also found that fitness (longevity) was directly influenced by the competitive environment. Treatment effects showed that longevity was reduced by experiencing high density in later life, and although non-significant, swapping from low density in early-life was more detrimental than *vice versa*. The effect of the mean number of competitors experienced also favoured lower numbers for increased survival. These density effects demonstrate that competition reduces average (absolute) fitness, a pattern that is found across animal taxa, with recent work also noting this effect in humans (Mariani et al., 2009).

5.5.2 *Among-individual correlations between traits and fitness*

After controlling for all fixed effects, our mixed model analyses provided strong evidence of among-individual variance in those traits with repeated measures (i.e. growth, activity, dominance score). We also found evidence of significant correlations between phenotypic traits (at the among-individual level) and between traits and fitness. However, not all of the relationships found were as we had predicted. Our re-analysis of activity confirmed that this trait is repeatable over the full time course of the experiment (Boulton et al., 2014). We also found among-individual variance for male dominance score confirming that, at least within a given social context, (i.e. each group in this study) dominance is a repeatable trait of the individual. However, we had predicted a positive correlation between individual boldness (where we use activity as a proxy) and dominance in line with results from other studies (Dingemanse and de Goede, 2004; Sundstrom et al., 2004; Webster et al., 2007; Dahlbom et al., 2011). In fact, we find a significant negative among-individual correlation between activity and dominance score, a result that is difficult to explain. Although somewhat speculative, it is possible that the negative association between boldness and dominance reflects alternate male strategies for obtaining resources (food and/or mating opportunities) that have been reported in some *Xiphophorus* species (Ryan and Causey, 1989; Zimmerer and Kallman, 1989; Ryan et al., 1992; Cummings and Gelineau-Kattner, 2009). For instance, socially dominant males may be able to hold territories in the natural environment, with subordinates having to use more mobile and exploratory (i.e. bold) behaviours to find undefended resources.

Other correlations with dominance score were more in line with our predictions. Thus we found that more dominant males tended to grow faster (as measured by percentage changes in standard length and weight). This agrees with previous work on this population (e.g. Wilson et al., 2013) and supports the hypothesis that dominance may determine size via effects on growth as well as *vice versa*. While we have focussed our analyses on growth rates rather than absolute size, it is perhaps worth noting that there are also strong among-individual correlations between dominance score and size after conditioning on the fixed effects described earlier (e.g. r_{PE} between SL and DOM_M = 0.650 ± 0.074, P <0.001). We also found, as predicted, that dominant males had greater longevity, while a number of other traits were also correlated with this measure of fitness. Thus, under laboratory conditions, selection through differential survival tends to favour fish that grow fast. Faster-than-average growing males tend to be the socially

dominant individuals that also mature later, at larger sizes and in better condition. We note that although survival is crucial for fitness, fecundity is equally important; however, it was not possible to monitor reproduction in our fish due to logistical constraints, such as assigning offspring to particular parents, and low reproductive rates during the period of the experiment. In other words, if selection through survival appears to favour later-maturing males in our study, this may not be the case for selection through reproductive success, or overall measures of fitness.

5.5.3 EVIDENCE FOR GENETIC EFFECTS

Our animal model analyses confirmed the presence of significant additive genetic effects contributing to observed phenotypic (co)variance but provided no evidence of genotype-by-environment interactions (although we acknowledge that sample sizes here were insufficient to provide powerful tests of GxE). Therefore, although statistical support for heritability varied across traits and in some cases between univariate and multivariate analyses, we conclude that there is evidence for genetic variance in boldness (activity), as well as in growth, life history in males (age, size and condition at maturity) and fitness (longevity) under laboratory conditions. The presence of genetic variance means that there is scope for adaptive evolution (Falconer and Mackay, 1996) although the extent that the traits involved can respond independently to selection on them will depend on the genetic covariance/correlation structure in **G** (Walsh and Blows, 2009). While our estimates of genetic correlations between traits were characterised by high levels of uncertainty, they were significant (or nearly so) in several cases. In most instances the sign of the genetic correlation matched that of the phenotypic correlation as discussed above. Two results from our genetic analysis are worth highlighting in particular.

Firstly, our estimate for the heritability of male dominance itself was very low (approximately 3%) and not statistically significant. Taking the lack of additive genetic variance for dominance score at face value implies that variation in competitive ability will not be a major driver of genetic variance for resource dependent traits and that the phenotypic relationships between dominance and other traits must be due to environmental not genetic effects. Potentially important implications for phenotypic evolution are raised by this result since it has recently been argued that if genetic variance in life history traits does come from genetic differences in competitive ability, it will not necessarily facilitate a selection response (Wilson 2014). This is because, if dominance is heritable, winning resources in competition will depend on a focal

individual's genotype and the genotypes of its competitors giving rise to indirect or social genetic effects (IGEs; Moore et al., 1997; Moore et al., 2002). When IGEs are present, selection on life history traits will cause a correlated evolution towards a more competitive social environment that offsets the phenotypic change otherwise expected (Hadfield et al., 2010).

Here we did not attempt to explicitly model IGEs of growth, life history or fitness, but can infer their likely absence from a lack of genetic variation for dominance. Nonetheless, non-genetic indirect effects, arising from competitor phenotypes, may well be playing an important role (Wilson et al., 2013). For instance, reduced growth rates in behaviourally subordinate fish could be an indirect consequence of experiencing harassment and bullying from fish with dominant phenotypes (as opposed to a direct consequence of obtaining less resource, e.g. food). More generally it is well known that physiological effects of chronic social stressors such as bullying can impact behaviour, health, life history and survival in animal populations (Pickering and Pottinger, 1989; Blanchard et al., 1998; Wingfield et al., 1998; Gregory and Wood, 1999; Boonstra et al., 2001; Barton, 2002). Individual fitness may depend therefore not only on the ability to win resources (and thus the phenotypic traits that promote resource winning) but also on the ability to cope with the social stress imposed by socially dominant conspecifics.

Although not included in the multivariate modelling for reasons of parsimony, brood effects on dominance score were marginally non-significant in our univariate analysis. Thus investigating whether early life environmental effects on social dominance are really present, and if so how they arise, may be an interesting area for future exploration. Our experimental aquaria set-up controlled for water quality differences between brood tanks, but is possible that the social environment provided by siblings may have affected phenotypes causal to dominance prior to their entry to the density-treatment study. Maternal effects represent another possible source of brood tank effects. In fact maternal effects on growth, a correlate of social dominance in our study, are known to occur in live bearing Poeciliid fishes (Lindholm et al., 2006). While they are widely assumed to arise from differential nutritional provisioning of eggs (and/or embryos in matrotrophic species), evidence from other fish taxa shows that deposition of maternal hormones (e.g. corticosteroids) and other substances into eggs can also alter offspring phenotype and fitness (Wilson et al., 2010b; Giesing et al., 2011).

A second important finding to emerge from our quantitative genetic analyses is that the among-individual variation in boldness previously reported (Boulton et al., 2014) is underpinned by significant heritable variation. Although it has long been known that genes influence personality in humans (e.g. Horn et al., 1976; Jang et al., 1996; Bouchard and McGue, 2003; Pilia et al., 2006) comparable studies on animals, particularly wild ones, are still quite rare (but see: Drent et al., 2003; Dingemanse et al., 2004; van Oers et al., 2005; Dingemanse et al., 2009). Thus our results add to a slowly emerging picture of genetic differences among individuals being important determinants of animal personality.

5.5.4 CONCLUSION

In summary, this study sought to investigate the direct effects of social competition on phenotype and fitness, test for among-individual variation in competitive ability (i.e. dominance) and investigate the multivariate genetic architecture linking traits causal and consequent to dominance. We found that higher levels of competition caused reductions in growth and fitness (longevity), while there was also some evidence for effects on male life history. Dominance score was repeatable in males, and positively correlated with growth and longevity at the among-individual level as predicted. However, while we found a correlation between personality (boldness) and dominance, the sign of this relationship was counter to our predictions. Thus, fish that were bolder actually tended to be less dominant. This result is something of an anomaly when set against the wider context of empirical studies of boldness. Investigations into the extent, and functional significance, of personality variation in wild *Xiphophorus* would be useful to tease out the biological significance of this result. We also found evidence of genetic (co)variance underpinning observed phenotypic variation. However, the estimated heritability of dominance itself was low and not statistically significant. We therefore conclude that correlations between dominance and other aspects of phenotype and fitness considered here are driven primarily by environmental rather than genetic effects.

CHAPTER 6

DISCUSSION

6.1 OVERVIEW

This thesis has sought to investigate the complex and intricate association of traits both causal and consequent to social dominance, a set of interactions referred to in the introduction as the *dynamic of dominance*. The simple pictorial model depicted in Figure 1.2 highlights an expectation that complex relationships between social dominance, size and growth, life history, stress physiology and personality will arise when individuals compete for limited resources. These relationships arise at the among-individual level due to differences in genes carried and / or environmental effects experienced, and are expected to ultimately affect individual fitness. Understanding how variation in, and covariation between traits is distributed at the among-individual level is core to understanding this dynamic, while elucidating the genetic basis of trait (co)variation is necessary if we wish to unravel its evolutionary implications. For this reason the empirical approach taken has necessarily been multivariate. Throughout this work, the strategy adopted for analysing the data with linear mixed effect models is largely borrowed from the field of quantitative genetics. This follows the recommendations of researchers working in the relatively new but rapidly expanding field of animal personality, who have advocated wider application of this approach in behavioural studies (Dingemanse et al., 2010; Dingemanse and Wolf, 2010; Dingemanse and Dochtermann, 2013). In this final chapter the main conclusions of the thesis are summarised. What has been learned regarding each component of the *dynamic of dominance*, and the validity of the overall concept from empirical studies of the swordtail fishes *Xiphophorus birchmanni* and *X. helleri* is outlined. Some limitations and omissions of the current work are then highlighted together with some suggested directions for future research that could usefully address these.

6.2 PERSONALITY IN SWORDTAILS

For the purposes of this thesis, personality was defined as among-individual variation in behaviour that is repeatable across time and context. Several personality traits have been identified, including those of boldness, aggression, exploration general activity and sociability (e.g. Dingemanse and de Goede, 2004; Svartberg et al., 2005). That

personality traits, notably boldness and aggression, should be positively linked to resource acquisition has been widely hypothesised (e.g. Sih et al., 2004a; Sih et al., 2004b; Biro and Stamps, 2008; Dzieweczynski and Crovo, 2011) and empirical studies to date have tended to confirm this pattern (e.g. Dingemanse et al., 2004; Brown et al., 2005a; Bell and Sih, 2007; Stamps, 2007; Biro and Stamps, 2008; Cote et al., 2010; Ariyomo and Watt, 2012; Rudin and Briffa, 2012; Mutzel et al., 2013). However, how important personality traits are in the determination of long term social dominance will depend on just how stable individual personality traits are. In Chapter 2 of this thesis, this question was addressed. To test for among-individual variation in (multivariate) behaviour across context, open field trials were used alongside exploration and emergence trials. Both experimental set ups involved quantifying individual behaviour in novel arenas and have been widely used to study boldness in fishes. To evaluate stability across time the study incorporated repeated observations of behaviour from two temporally separate sampling periods. The first of these covered a long time period effectively representative of individual lifetimes (50 weeks), while the second study period (four weeks) was more typical of the predominately short–term personality studies published to date.

Modelling a suite of behavioural traits observed on individuals across the two trial types revealed a very strong axis of among-individual variance. Furthermore, this axis of variation was broadly similar when estimated from long- and short-term periods and could be biologically interpreted as a shy-bold continuum. Repeatability for individual behaviours tended to be somewhat higher when estimated from data with short inter-observation interval. Since behavioural repeatability is often the statistical signature that personality is inferred from (Réale et al., 2007), it follows that conclusions about the importance of personality (e.g. for determining lifetime fitness) may generally be anticonservative if behaviour is observed over short time periods only. Nonetheless, the use of an overlapping set of individuals in the long- and short-term periods of our study allowed us to conclude that multivariate behaviour is (relatively) stable among individuals across lifetimes.

6.3 LINKING PHYSIOLOGICAL STRESS WITH CONTEST BEHAVIOUR AND OUTCOME

Competition is likely to be an important source of stress in animal populations, particularly where agonistic behaviours are used to acquire resources and assert dominance (Blanchard et al., 1998; Blanchard et al., 2001). For instance, the presence of socially dominant individuals can be a source of chronic stress in subordinates, with

the former tending to bully, harass or even cause physical injury to the latter. This will tend to exacerbate the already negative consequences of reduced resource acquisition in subordinates. Therefore, while the ability to cope with chronic stress may be important for determining fitness in subordinate animals (Ling et al., 2009), the acute stress response is also expected to influence contest behaviour and outcome. Under laboratory conditions, where dyadic contests are commonly staged to investigate social dominance, individuals may be stressed by the experimental protocols themselves, potentially giving rise to false results or conclusions.

In Chapter 3, behavioural and endocrine data from a study of male-male contests in the green swordtail (*X. helleri,* a widely used model for aggression and dominance studies) were obtained from collaborators at the University of Alabama. Because *X. helleri* is a very small fish and the collection of blood plasma for analysis would likely be fatal, a novel (non-invasive) water-borne hormone assay was used to measure HPI (hypothalamic-pituitary-interrenal) axis activity in the form of circulating cortisol levels. These data were used to test hypothesised links between physiological stress response and contest behaviour and outcome. Based on the stress coping style (SCS) model (Koolhaas et al., 1999) explored further in Chapter 4 (see below) we predicted that expression of low pre-contest cortisol levels would be associated with rapid behavioural recovery from disturbances caused by experimental protocol (i.e. capture, confinement and lifting a dividing partition). We also predicted that individuals showing these characteristics (commonly termed *proactive* stress coping style) would tend to be both aggressive and dominant. Thus they would have short latencies to initiate contests and would more likely be the eventual contest winners. Finally, we predicted that post-contest levels of cortisol would be lower in winners than losers. This follows the simple expectation that, all else being equal, losing will be more stressful than winning contests.

However, while the data did support associations between contest behaviour and stress physiology these were largely counter to our original predictions. Thus individuals did not readily conform to the usual *reactive* versus *proactive* coping style model. For instance, it was contest initiators that had the higher baseline cortisol levels and highest physiological stress response (measured as the change between pre- and post- contest cortisol expression). However, what we were unable to test for here was post-contest timescale for recovery of baseline (pre-contest) cortisol levels. It would, for example, be interesting to know if contest winners recovered baseline levels more

quickly than losers as has been reported in some other studies (Netherton et al., 2004). It is also possible that the single observation per fish was not sufficient to reveal a pattern of underlying among-individual variation in stress coping styles.

6.4 Is personality part of an integrated stress coping style?

As noted above, a potentially important limitation of the data available for analysis in Chapter 3 was that it comprised only a single observation per fish. Thus while the relationships found between contest related behaviours and physiological stress could have been driven by differences among fish, they might equally reflect trial specific processes. In fact, while the stress coping style model posits among-individual variation, the vast majority of studies testing it have used only a single measure approach. Therefore, in Chapter 4 the link between stress and behaviour was explored further using a repeated measures approach. By investigating if, and how, personality (i.e. repeatable behavioural characteristics) is associated with repeatable stress response physiology, the aim was to test the stress coping style model itself. If the model is valid then both behavioural and physiological stress response traits should not only be repeatable but should also change in an integrated manner along a major axis of among-individual variation. In other words there should be among-individual correlation structure between behaviour and physiology.

For this study we used the non-invasive endocrine sampling method employed in Chapter 3, but focussed on the personality trait of boldness (as identified in Chapter 2), rather than aggression. Our behavioural experiment combined open field trials (OFT, considered to be a mild stressor in small fishes; Walsh and Cummins, 1976; Archard et al., 2012) with a more severe acute stress stimulus in the form of a simulated predator attack. Multivariate behavioural responses to the mildly stressful OFT were repeatable and consistent with a shy-bold axis of variation, confirming the result of our earlier study (Chapter 2). However, boldness did not clearly predict the behavioural response to the more severe stressor, and in fact evidence of a repeatable behavioural response to the simulated predator attack was relatively weak.

Using the water-borne hormone collection method, physiological measures, i.e. cortisol and 11-ketotestosterone (11KT, an androgen that has been linked to both boldness and dominance in fishes; Borg and Mayer, 1995; Desjardins et al., 2008; Archard et al., 2012) were collected before and after the behavioural trial to test for repeatable among-individual variation in baseline and stress response hormone levels. Our data

suggested that baseline levels of 11KT, but not cortisol, were repeatable among individuals over the time period of our study (4 weeks). In fact *post hoc* analysis did indicate significant among-individual variance in baseline cortisol, but only over shorter time periods. Thus it appears that individuals can differ consistently in this important stress-related trait, but that these differences may have low temporal stability, thus rather limiting the utility of the SCS model. Additionally, to the extent that baseline cortisol did differ among individuals, it was positively correlated with boldness (i.e. proactive type behaviours) rather than negatively as hypothesized.

The finding from Chapter 3, that stress coping styles are not easy to validate, was therefore endorsed by this study. Together these results raise doubts about the general applicability of the SCS model that was originally developed in rodent studies(Koolhaas et al., 1999), but has also found support in some fish studies (Øverli et al., 2004). In *Xiphophorus sp.* there is certainly evidence of repeatable behavioural variation along axes of boldness (this study) and aggression (Wilson et al., 2013). However, it is less clear that there is stable among-individual variation in stress physiology. Furthermore, we found a general tendency for individuals viewed as behaviourally proactive (i.e. more bold, more aggressive) to have higher cortisol levels, not lower as expected under the SCS model.

6.5 TOWARDS THE *DYNAMIC OF DOMINANCE*

The final study of this thesis (Chapter 5) sought to bring together ideas from previous chapters by investigating the extent that personality contributes to social dominance, and by quantifying the relationships between dominance and growth, life history and longevity (used as a proxy for fitness). It also sought to determine the extent that components of the *dynamic of dominance* are shaped by both genetic and environmental effects, the latter including plastic responses to the level of competition experienced. To these ends we combined quantitative genetic modelling with experimental manipulation of housing density (where higher density was assumed to equal increased competition) applied to a captive bred generation of *X. birchmanni* produced from known parental crosses. All fish were phenotyped for size and growth traits, while because the same individuals were used in this study as in Chapter 2, behavioural data on boldness were already available. Additionally, for males only, we were able to allocate scores for within-group social dominance based on behavioural observations of fish in their home environments and obtain data on life history traits (age, size and condition at maturity).

As expected, competition from high density environments directly affected some traits, notably reducing growth during the first 28 weeks of the study and longevity during the latter 22 weeks. Dominance score was repeatable in males, consistent with a stable (within-group) dominance hierarchy, and dominant males tended to grow faster and mature later at a larger size. Importantly, dominant individuals also tended to live longer, confirming the premise that, in competition, some individuals win fitness at the expense of others (Brockelman, 1975). While these findings are consistent with our *a priori* expectations, not all results fit the *dynamic of dominance* model perfectly.

For example, while we had predicted a positive correlation between boldness (where we used activity as a proxy based on the findings of Chapter 2) and dominance, the relationship we found was actually negative. One possible explanation for this result is that activity in the OFT is a poor measure of boldness in *X. birchmanni*. Certainly, a number of authors have argued that activity and boldness are better considered as separate (if potentially correlated) personality traits (e.g. Burns, 2008; Brown and Irving, 2014). However, analyses presented in Chapters 2 and 4 of this thesis showed that activity is strongly correlated among individuals with other behaviours that are widely used as indicators of bold personality (e.g. exploring a large area of the novel arena, spending a higher proportion of time exposed in the centre of the tank). A second possibility is that the negative relationship between boldness and dominance reflects the territorial nature of male *Xiphophorus sp*. In simple terms, it may be that dominant individuals do not need to go looking for trouble (or indeed opportunities to acquire resources), rather, they wait for trouble to come and find them.

Behavioural data on *Xiphophorus sp*. in the wild are limited; however, males do set up territories and defend them (Franck and Ribowski, 1993; Franck et al., 1998). Certainly based on personal observations, fish that are dominant in their home tanks do not so much actively pursue male competitors as defend the immediate territory around their resource (e.g. female(s)). Consequently, it may be that subordinate individuals are required to adopt a more mobile strategy, moving through the environment in search of unguarded resources (food, mates). Note that if so, in this system social dominance may be driving differences in personality rather than *vice versa*. Undoubtedly, more behavioural data from *Xiphophorus* studied *in situ* would provide useful insights as they have for other fish taxa. This is because personality differences in the wild are likely to impact fitness via multiple routes. For example, bold fish may be better at finding resources but are also easier targets for predators (Brown and Braithwaite, 2004).

Better understanding of the functional significance of the boldness variation discovered here in the ecological context of a natural environment is therefore important.

6.6 GENETIC AND ENVIRONMENTAL EFFECTS ON THE *DYNAMIC OF DOMINANCE*

The quantitative genetic modelling employed in the Chapter 5 study allowed us to begin untangling the genetic and environmental influences on traits contained within the *dynamic of dominance* concept. Certainly, genetic variance for personality (i.e. boldness, as we have interpreted it) was present in the captive bred population of *X. birchmanni*. Genetic (co)variance also seems to be present within and between a number of traits that are influenced by competition (i.e. size, growth, male maturation). To the extent that genetic (co)variance is present in directions of phenotypic change favoured by selection under natural conditions, it will provide scope for adaptive evolution. However, there was limited support for social dominance itself being heritable. Thus while it was clear that individuals do differ in competitive ability, and that this has downstream consequences (e.g. for growth and longevity) we cannot reject the null hypothesis that this is due to environmental effects alone.

We acknowledge that statistical power may be a limiting factor here, particularly since dominance could only be assessed in one sex, and it is possible that a larger study would detect genetic variance. Moreover, while our heritability estimate for dominance was low, even a small amount of genetic variance could have important evolutionary consequences. For example, if individuals grow quickly because they are dominant, then, given the strong positive genetic correlation between dominance and growth, it is possible that a large proportion of the genetic variance for growth is being driven by dominance. Because genetic variance arising from differences in competitive ability may not fully allow a selection response (Hadfield, 2010; Wilson, 2014), selection for faster growth (seen under laboratory conditions, but also widely reported in the wild, e.g. Sogard, 1997; Brown and Braithwaite, 2004) may result in even less phenotypic change than predicted given the (already low) heritability of growth.

6.7 LIMITATIONS AND SUGGESTIONS FOR FUTURE STUDIES

In this thesis the *dynamic of dominance,* a conceptual model that highlights the need for multivariate and multidisciplinary studies of competition was explored. In using this idea as a template, extensive empirical studies in *Xiphophorus* fishes have answered some questions but also raised a number of others.

For instance, what is the actual ecological significance of the personality variation uncovered? Personality variation is clearly present and, at least under the artificial laboratory conditions used here, is related to growth, male maturation and longevity. Personality was also demonstrated to be stable across time periods representative of likely life spans in the wild. However, if the statistical signature of personality is clear, a biological interpretation for the trait we have labelled as boldness remains ambiguous. Determining the difference between boldness and other personality traits such as curiosity, exploration, fearfulness or anxiety is very difficult (e.g. Boissy, 1995; Burns, 2008; Carter et al., 2012; Carter et al., 2013). In fact an argument could be made that attempts to do so offer comparatively little new biological insight for those researchers with a primary focus on ecological and evolutionary (as opposed to psychological) questions. In contrast, more efforts to observe animals in natural environments may shed light on the functional importance of personality traits without the need for labels from standardised (psychological) terminology. For instance, if a particular behavioural phenotype is known to forage more widely, resulting in higher resource acquisition but greater risk of predation, then it is of little practical importance whether such individuals are labelled as bold or exploratory.

Within the laboratory setting used here, incorporation of further behavioural phenotypes into the *X. birchmanni* work may have been useful. For instance, while agonistic behaviours were examined in the *X. helleri* work in Chapter 2, whether individual aggressiveness and boldness were correlated within a behavioural syndrome as has been hypothesised was not formally tested. In fact two rounds of controlled dyadic contests (n = 870 trials) between pairs of *X. birchmanni* males were carried out in a neutral arena following protocols previously used successfully in their wild-caught parents (Wilson et al., 2013). However, due to a lack of agonistic behaviours observed, these data were not useful to characterise aggression or assign contest winners. Possible explanations for the lack of interaction include inexperience (or young age) of males at the time trials were conducted, and / or lack of motivation (e.g. no female or food resource to defend). It also seems possible that capture and transfer of fish to the experimental arena may have proved stressful enough to disrupt male-male aggression. This hypothesis is supported by the fact that aggressive behaviours were observed between males within undisturbed rearing groups (and indeed these were used to calculate dominance scores).

A second key question to emerge concerns the generality of the widely applied stress coping style model (SCS; Koolhaas et al., 1999). At least for swordtails it seems clear that SCS presents an overly simplistic description of the acute stress response that did not stand up well to rigorous empirical scrutiny. Where relationships were found between personality and endocrine traits, they were in the opposite direction to predictions based on our interpretation of a broad equivalence between *bold* and *proactive* behavioural types as presented in the literature. Had collection of physiological traits been feasible for the entire pedigreed population, a better understanding of stress coping styles in *Xiphophorus* species and their integration with the *dynamic of dominance* may have been reached. For example, measures of baseline cortisol levels at regular intervals throughout the long term study may have exposed signals of chronic stress, and allowed us to test whether this was associated with the slower growth and reduced longevity found in less dominant individuals. Such an expansion of the current study would also have allowed genetic parameters for stress-related endocrine traits to have been thoroughly investigated.

A final point to note was that while sample sizes used here were large in comparison to those typically used in studies of animal personality and stress response, they were relatively small for the purposes of quantitative genetic analysis. In fact a second generation of breeding had been planned to provide greater numbers of study participants but surviving adults failed to breed reliably under laboratory conditions and thus this was not feasible. Although swordtails were chosen in large part due to their known formation of stable dominance hierarchies, a more prolific Poeciliid (e.g. guppy) might be a better choice of study species going forward. Regardless of species, a larger study would clearly provide more precise estimates of quantitative genetic parameters. This in turn might allow for stronger conclusions to be drawn regarding the presence (or lack thereof) of genetic variance for social dominance. Moreover, a larger sample size would have provided more power to test for GxE across competition environments (i.e. differences in **G** between traits expressed in high- versus low-density treatments). Similarly we could compare phenotype-fitness associations across treatments to ask if selection (through longevity) changed with increasing levels of social competition. Finally, given a larger study it would also be interesting to formally test for indirect genetic effects (IGEs) on resource-dependent traits and fitness using the analytical extensions to the animal model outlined by Bijma (2010). This would allow testing of the prediction that, in the case that social dominance is truly heritable,

IGEs on resource dependent traits such as growth will arise from competition and act to constrain phenotypic responses to selection.

BIBLIOGRAPHY

Abbott, D.H., Keverne, E.B., Bercovitch, F.B., Shively, C.A., Medoza, S.P., Saltzman, W., Snowdon, C.T., Ziegler, T.E., Banjevic, M., Garland, T., Sapolsky, R.M., 2003. Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. Hormones and Behavior 43, 67-82.

Adams, C.E., Huntingford, F.A., 1997. Growth, maturation and reproductive investment in Arctic charr. Journal of Fish Biology 51, 750-759.

Alcock, J., Houston, T.F., 1996. Mating systems and male size in Australian hylaeine bees (Hymenoptera: Colletidae). Ethology 102, 591-610.

Ali, M., Nicieza, A., Wootton, R.J., 2003. Compensatory growth in fishes: a response to growth depression. Fish and Fisheries 4, 147-190.

Amouriq, L., 1964. L'activite et le pheromone social cez *Lebistes reticulatus* (Poeciliidea, Cyprinodontiformes). Comptes Rendus de l'Académie des Sciences Paris 259, 2701-2702.

Amouriq, L., 1967. Sensibilite des *Lebistes reticulatus* male a la substance dynamogene emise par des femelles de poeciliidie et Gasterosteidae. Revue de Comportement Animal 4, 83-86.

Andrade, O., Orihuela, A., Solano, J., Galina, C.S., 2001. Some effects of repeated handling and the use of a mask on stress responses in Zebu cattle during restraint. Applied Animal Behaviour Science 71, 175-181.

Araya-Ajoy, Y.G., Dingemanse, N.J., 2014. Characterizing behavioural 'characters': an evolutionary framework. Proceedings of the Royal Society B-Biological Sciences 281, 26-45.

Archard, G.A., Braithwaite, V.A., 2011. Increased exposure to predators increases both exploration and activity level in *Brachyrhaphis episcopi*. Journal of Fish Biology 78, 593-601.

Archard, G.A., Earley, R.L., Hanninen, A.F., Braithwaite, V.A., 2012. Correlated behaviour and stress physiology in fish exposed to different levels of predation pressure. Functional Ecology 26, 637-645.

Ariyomo, T.O., Watt, P.J., 2012. The effect of variation in boldness and aggressiveness on the reproductive success of zebrafish. Animal Behaviour 83, 41-46.

Ariyomo, T.O., Carter, M., Watt, P.J., 2013a. Heritability of boldness and aggressiveness in the zebrafish. Behaviour Genetics 43, 161-167.

Ariyomo, T.O., Carter, M., Watt, P.J., 2013b. Heritability of Boldness and Aggressiveness in the Zebrafish. Behaviour genetics 43:161-167.

Arnott, G., Elwood, R.W., 2008. Information gathering and decision making about resource value in animal contests. Animal Behaviour 76, 529-542.

Arnott, G., Elwood, R., 2009a. Probing aggressive motivation in a cichlid fish. Biology Letters 5, 762-764.

Arnott, G., Elwood, R.W., 2009b. Assessment of fighting ability in animal contests. Animal Behaviour 77, 991-1004.

Arnott, G., Elwood, R.W., 2010. Startle durations reveal visual assessment abilities during contests between convict cichlids. Behavioural Processes 84, 750-756.

Ashley, P.J., 2007. Fish welfare: Current issues in aquaculture. Applied Animal Behaviour Science 104, 199-235.

Aubin-Horth, N., Deschenes, M., Cloutier, S., 2012. Natural variation in the molecular stress network correlates with a behavioural syndrome. Hormones and Behavior 61, 140-146.

Bakker, T.C.M., 1986. Aggressiveness in sticklebacks (*Gasterosteus-aculeatus L*) - a behaviour-genetic study. Behaviour 98, 1-144.

Balm, P.H.M., 1997. Immune-Endocrine reactions. In: Iwama, G.K., A.D.Pickering, Sumpter, J.P., C.B.Schreck (Eds.), Fish Stress and Health in Aquaculture. Cambridge University Press, Cambridge, pp. 195-221.

Bao, I.Y., Kallman, K.D., 1978. Genetic and endocrine control of sexual-matiration in hybrids between *Xiphophorus-helleri* and *Xiphophorous-maculatus*. American Zoologist 18, 669-669.

Barber, I., Walker, P., Svensson, P.A., 2004. Behavioural responses to simulated avian predation in female three spined sticklebacks: the effect of experimental *Schistocephalus solidus* infections. Behaviour 141, 1425-1440.

Barton, B.A., 2002. Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. Integrative and Comparitive Biology 42, 517-525.

Barton, B.A., Iwama, G.K., 1991. Physiological changes in fish from stress in aquaculture with emphasis on the response and effects of corticosteroids Reviews of Fish Diseases 1, 3-26.

Basolo, A.L., 1988. Female preference for body size, sword length and sword colour in the swordtail, *Xiphophorus helleri*. American Zoologist 28, A152-A152.

Basolo, A.L., Trainor, B.C., 2002. The conformation of a female preference for a composite male trait in green swordtails. Animal Behaviour 63, 469-474.

Baur, A., Baur, B., 1992. Response in growth, reproduction and life-span to reduced competition pressure in the land snail *Balea-perversa*. Oikos 63, 298-304.

Beaugrand, J.P., 1997. Resolution of agonistic conflicts in dyads of acquainted green swordtails (*Xiphophorus helleri*): a game with perfect information. Behavioural Processes 41, 293-310.

Beaugrand, J.P., Zayan, R., 1985. An experimental-model of aggressive dominance in *Xiphophorus helleri* (Pisces, Poeciliidae). Behavioural Processes 10, 1-52.

Beaugrand, J.P., Cotnoir, P.A., 1996. The role of individual differences in the formation of triadic dominance orders of male green swordtail fish (*Xiphophorus helleri*). Behavioural Processes 38, 287-296.

Beaugrand, J.P., Goulet, C., 2000. Distinguishing kinds of prior dominance and subordination experiences in males of green swordtail fish (*Xiphophorus helleri*). Behavioural Processes 50, 131-142.

Beaugrand, J.P., Payette, D., Goulet, C., 1996. Conflict outcome in male green swordtail fish dyads (*Xiphophorus helleri*): interaction of body size, prior dominance/subordination experience, and prior residency. Behaviour 133, 303-319.

Bell, A., Sih, A., 2007. Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). Ecology Letters 10, 828-834.

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

Bell, M.B.V., Nichols, H.J., Gilchrist, J.S., Cant, M.A., Hodge, S.J., 2012. The cost of dominance: suppressing subordinate reproduction affects the reproductive success of dominant female banded mongooses. Proceedings of the Royal Society B-Biological Sciences 279, 619-624.

Benson, K.E., Basolo, A.L., 2006. Male-male competition and the sword in male swordtails, *Xiphophorus helleri*. Animal Behaviour 71, 129-134.

Benus, R.F., Bohus, B., Koolhaas, J.M., Vanoortmerssen, G.A., 1991. Heritable Variation for Aggression as a Reflection of Individual Coping Strategies. Experientia 47, 1008-1019.

Bergsma, R., Kanis, E., Knol, E.F., Bijma, P., 2008. The contribution of social effects to heritable variation in finishing traits of domestic pigs (*Sus scrofa*). Genetics 178, 1559-1570.

Bernier, N.J., Alderman, S.L., Bristow, E.N., 2008. Heads or tails? Stressor-specific expression of corticotropin-eleasing factor and urotensin I in the preoptic area and caudal neurosecretory system of rainbow trout. Journal of Endocrinology 196, 637-648.

Bernstein, I.S., 1976. Dominance, agression and reproduction in primate societies. Journal of Theoretical Biology 60, 459-472.

Bijma, P., 2010. Estimating indirect genetic effects: precision of estimates and optimum designs. Genetics 186, 1013-1028.

Biro, P.A., Stamps, J.A., 2008. Are animal personality traits linked to life-history productivity? Trends in Ecology & Evolution 23, 361-368.

Blanchard, R.J., McKittrick, C.R., Blanchard, D.C., 2001. Animal models of social stress: effects on behavior and brain neurochemical systems. Physiology & Behavior 73, 261-271.

Blanchard, R.J., Nikulina, J.N., Sakai, R.R., McKittrick, C., McEwen, B., Blanchard, D.C., 1998. Behavioral and endocrine change following chronic predatory stress. Physiology and Behavior 63, 561-569.

Blumstein, D.T., Daniel, J.C., 2007. Quantifying Behavior the JWatcher Way Sinauer Associates Inc., Sunderland, MA, 211 pp.

Boissy, A., 1995. Fear and fearfulness in animals. Quarterly Review of Biology 70, 165-191.
Boonstra, R., McColl, C.J., Karels, T.J., 2001. Reproduction at all costs: the adaptive stress response of male Arctic ground squirrels. Ecology 82, 1930-1946.
Borg, B., Mayer, I., 1995. Androgens and behaviour in the three-spined stickleback. Behaviour 132, 1025-1035.
Borowsky, R., 1978. Social inhibition of maturation in natural populations of *Xiphophorous variatus* (Pisces Poecilidae). Science 201, 933-935.
Borowsky, R.L., 1973. Social control of adult size in males of *Xiphophorus variatus* Nature 245, 332-335.
Borowsky, R.L., 1987. Agonistic behaviour and social inhibition of maturation in fishes of the genus *Xiphophorous* (Poecilidae). Copeia 3, 792-796.
Borowsky, R.L., McClelland, M., Cheng, R., Welsh, J., 1995. Arbritrarily primed DNA-fingerprinting for phylogenetic reconstruction in vertebrates - the *Xiphophorus* model. Molecular Biology and Evolution 12, 1022-1032.
Bouchard, T.J., McGue, M., 2003. Genetic and environmental influences on human psychological differences. Journal of Neurobiology 54, 4-45.
Boulton, K., Sinderman, B., Pearce, M., Earley, R., Wilson, A., 2012. He who dares only wins sometimes: physiological stress and contest behaviour in *Xiphophorus helleri*. Behaviour 149, 977-1002.
Boulton, K., Grimmer, A.J., Rosenthal, G.G., Walling, C.A., Wilson, A.J., 2014. How stable are personalities? A multivariate view of behavioural variation over long and short timescales in the sheepshead swordtail, *Xiphophorus birchmanni*. Behavioural Ecology and Sociobiology 68, 791-803.
Boulton, K., Massault, C., Koning, D.J.D., Houston, R., Haley, C., Batargias, C., Bovenhuis, H., Canario, A., Kotoulas, G., Tsigenopoulos, C., 2011. QTL for growth and stress response in the gilthead seabream (*Sparus aruata*). Aquaculture 319 58-66.
Brelin, D., Petersson, E., Dannewitz, J., Dahl, J., Winberg, S., 2008. Frequency distribution of coping strategies in four populations of brown trout (*Salmo trutta*). Hormones and Behavior 53, 546-556.
Briffa, M., Sneddon, L.U., 2007. Physiological constraints on contest behaviour. Functional Ecology 21, 627-637.
Briffa, M., Rundle, S.D., Fryer, A., 2008. Comparing the strength of behavioural plasticity and consistency across situations: animal personalities in the hermit crab *Pagurus bernhardus*. Proceedings of the Royal Society B-Biological Sciences 275, 1305-1311.
Brockelman, W.Y., 1975. Competition, fitness of offspring and optimal clutch size. American Naturalist 109, 677-699.
Brodin, T., Lind, M.I., Wiberg, M.K., Johansson, F., 2013. Personality trait differences between mainland and island populations in the common frog (*Rana temporaria*). Behavioral Ecology and Sociobiology 67, 135-143.
Brommer, J.E., 2013. On between-individual and residual (co)variances in the study of animal personality: are you willing to take the "individual gambit"? Behavioral Ecology and Sociobiology 67, 1027-1032.
Bronson, F.H., 1973. Establishment of social rank among grouped male mice - relative effects on circulating FSH, LH and corticosterone. Physiology & Behavior 10, 947-951.
Brown, C., Braithwaite, V.A., 2004. Size matters: a test of boldness in eight populations of the poeciliid *Brachyraphis episcopi*. Animal Behaviour 68, 1325-1329.
Brown, C., Irving, E., 2014. Individual personality traits influence group exploration in a feral guppy population. Behavioral Ecology 25, 95-101.
Brown, C., Jones, F., Braithwaite, V., 2005a. In situ examination of boldness-shyness traits in the tropical poeciliid, *Brachyraphis episcopi*. Animal Behaviour 70, 1003-1009.
Brown, C., Gardner, C., Braithwaite, V.A., 2005b. Differential stress responses in fish from areas of high- and low-predation pressure. Journal of Comparative Physiology B 175, 305-312.
Brown, C., Burgess, F., Braithwaite, V.A., 2007. Heritable and experiential effects on boldness in a tropical poeciliid. Behavioural Ecology and Sociobiology 62, 237-243.

Budaev, S.V., 1997a. "Personality" in the guppy (*Poecilia reticulata*): a correlational study of exploratory behavior and social tendency. Journal of Comparative Psychology 111, 399-411.
Budaev, S.V., 1997b. Alternative styles in the European wrasse, *Symphodus ocellatus*: Boldness-related schooling tendency. Environmental Biology of Fishes 49, 71-78.
Budaev, S.V., 1999. Sex differences in the Big Five personality factors: testing an evolutionary hypothesis. Personality and Individual Differences 26, 801-813.
Budaev, S.V., 2010. Using principal components and factor analysis in animal behaviour research: caveats and guidelines. Ethology 116, 472-480.
Budaev, S.V., Zworykin, D.D., 2002. Individuality in fish behavior: ecology and comparative psychology. Journal of Ichthyology 42, S189-S195.
Budaev, S.V., Zworykin, D.D., Mochek, A.D., 1999. Consistency of individual differences in behaviour of the lion-headed cichlid, *Steatocranus casuarius*. Behavioural Processes 48, 49-55.
Burns, J.G., 2008. The validity of three tests of temperament in guppies (*Poecilia reticulata*). Journal of Comparative Psychology 122, 344-356.
Bushuev, A.V., Kerimov, A.B., Ivankina, E.V., 2010. Estimation of heritability and repeatability of resting metabolic rate in birds, with free-living pied flycatchers *Ficedula hypoleuca* (Ayes: Passeriformes) as an example. Zhurnal Obshchei Biologii 71, 402-424.
Campton, D.E., 1992. Heritability of body size of green swordtails, *Xiphophorous helleri.* 1. Sib analysis of males reared individually and in groups. Journal of Heredity 83, 43-48.
Carere, C., Caramaschi, D., Fawcett, T.W., 2010. Covariation between personalities and individual differences in coping with stress: converging evidence and hypotheses. Current Zoology 56, 728-740.
Carere, C., Groothuis, T.G.G., Mostl, E., Daan, S., Koolhaas, J.M., 2003. Fecal corticosteroids in a territorial bird selected for different personalities: daily rhythm and the response to social stress. Hormones and Behavior 43, 540-548.
Carere, C., Drent, P.J., Privitera, L., Koolhaas, J.M., Groothuis, T.G.G., 2005. Personalities in great tits, *Parus major*: stability and consistency. Animal Behaviour 70, 795-805.
Carlson, A.A., Young, A.J., Russell, A.F., Bennett, N.C., McNeilly, A.S., Clutton-Brock, T., 2004. Hormonal correlates of dominance in meerkats (*Suricata suricatta*). Hormones and Behavior 46, 141-150.
Carter, A.J., Marshall, H.H., Heinsohn, R., Cowlishaw, G., 2012. How not to measure boldness: novel object and antipredator responses are not the same in wild baboons. Animal Behaviour 84, 603-609.
Carter, A.J., Feeney, W.E., Marshall, H.H., Cowlishaw, G., Heinsohn, R., 2013. Animal personality: what are behavioural ecologists measuring? Biological Reviews 88, 465-475.
Carvalho, C.F., Leitão, A.V., Funghi, C., Batalha, H.R., Reis, S., Mota, P.G., Lopes, R.J., Cardoso, G.C., 2013. Personality traits are related to ecology across a biological invasion. Behavioral Ecology 24, 1081-1091.
Chase, J.M., Abrams, P.A., Grover, J.P., Diehl, S., Chesson, P., Holt, R.D., Richards, S.A., Nisbet, R.M., Case, T.J., 2002. The interaction between predation and competition: a review and synthesis. Ecology Letters 5, 302-315.
Chervet, N., Zöttl, M., Schürch, R., Taborsky, M., Heg, D., 2011. Repeatability and heritability of behavioural types in a social cichlid. International Journal of Evolutionary Biology 2011: 321729.
Chrousos, G.P., 1998. Stressors, stress, and neuroendocrine integration of the adaptive response - The 1997 Hans Selye Memorial Lecture. Stress of Life 851, 311-335.
Cobb, J.S., Tamm, G.R., 1975. Dominance status and molt order in lobsters *Homarus americanus* Marine Behaviour and Physiology 3, 119-124.
Conrad, J.L., Sih, A., 2009. Behavioural type in newly emerged steelhead *Oncorhynchus mykiss* does not predict growth rate in a conventional hatchery rearing environment. Journal of Fish Biology 75, 1410-1426.
Conrad, J.L., Weinersmith, K.L., Brodin, T., Saltz, J.B., Sih, A., 2011. Behavioural syndromes in fishes: a review with implications for ecology and fisheries management. Journal of Fish Biology 78, 395-435.

Constantz, G.D., 1989. Reproductive biology of Poeciliid fishes. In: Meffe, G.K., Snelson, F.F. (Eds.), Ecology and Evolution of Livebearing Fishes (Poeciliidae). Prentice Hall, Englewood Cliffs, New Jersey, pp. 33-50.

Constanz, G.D., 1984. Sperm competition in Poeciliid fishes. In: Smith, R.L. (Ed.), Sperm Competition and the Evolution of Animal Mating Systems. Academic Press, New York, pp. 465-475.

Cote, J., Clobert, J., Brodin, T., Fogarty, S., Sih, A., 2010. Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. Philosophical Transactions of the Royal Society B-Biological Sciences 365, 4065-4076.

Creel, S., 2001. Social dominance and stress hormones. Trends in Ecology & Evolution 16, 491-497.

Cummings, M.E., Gelineau-Kattner, R., 2009. The energetic costs of alternative male reproductive strategies in *Xiphophorus nigrensis*. Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology 195, 935-946.

Dahlbom, S.J., Lagman, D., Lundstedt-Enkel, K., Sundstrom, L.F., Winberg, S., 2011. Boldness predicts social status in Zebrafish (*Danio rerio*). Plos One 6, e23565.

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

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

David, M., Auclair, Y., Cezilly, F., 2012. Assessing short- and long-term repeatability and stability of personality in captive Zebra finches using longitudinal data. Ethology 118, 932-942.

Depeche, J., 1976. Aquisition and limits of embryonic trophic autonomy during development of the viviparous teleostean fish *Poecilia reticulata*. Bulletin Biologique de la France et de la Belgique 110, 45-97.

Desjardins, J.K., Stiver, K.A., Fitzpatrick, J.L., Milligan, N., Van Der Kraak, G.J., Balshine, S., 2008. Sex and status in a cooperative breeding fish: behavior and androgens. Behavioral Ecology and Sociobiology 62, 785-794.

DeVries, A.C., 2002. Interaction among social environment, the hypothalamic-pituitary-adrenal axis, and behavior. Hormones and Behavior 41, 405-413.

DiBattista, J.D., Anisman, H., Whitehead, M., Gilmour, K.M., 2005. The effects of cortisol administration on social status and brain monoaminergic activity in rainbow trout *Oncorhynchus mykiss*. Journal of Experimental Biology 208, 2707-2718.

Dingemanse, N.J., de Goede, P., 2004. The relation between dominance and exploratory behavior is context-dependent in wild great tits. Behavioral Ecology 15, 1023-1030.

Dingemanse, N.J., Wolf, M., 2010. Recent models for adaptive personality differences: a review. Philosophical Transactions of the Royal Society B-Biological Sciences 365, 3947-3958.

Dingemanse, N.J., Dochtermann, N.A., 2013. Quantifying individual variation in behaviour: mixed-effect modelling approaches. Journal of Animal Ecology 82, 39-54.

Dingemanse, N.J., Dochtermann, N.A., Wright, J., 2010. A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets. Animal Behaviour 79, 439-450.

Dingemanse, N.J., Dochtermann, N.A., Nakagawa, S., 2012a. Defining behavioural syndromes and the role of 'syndrome deviation' in understanding their evolution. Behavioral Ecology and Sociobiology 66, 1543-1548.

Dingemanse, N.J., Both, C., Drent, P.J., Tinbergen, J.M., 2004. Fitness consequences of avian personalities in a fluctuating environment. Proceedings of the Royal Society B-Biological Sciences 271, 847-852.

Dingemanse, N.J., Both, C., van Noordwijk, A.J., Rutten, A.L., Drent, P.J., 2003. Natal dispersal and personalities in great tits (*Parus major*). Proceedings of the Royal Society B-Biological Sciences 270, 741-747.

Dingemanse, N.J., Wright, J., Kazem, A.J.N., Thomas, D.K., Hickling, R., Dawnay, N., 2007. Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. Journal of Animal Ecology 76, 1128-1138.

Dingemanse, N.J., Bouwman, K.M., van de Pol, M., van Overveld, T., Patrick, S.C., Matthysen, E., Quinn, J.L., 2012b. Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. Journal of Animal Ecology 81, 116-126.

Dingemanse, N.J., Van der Plas, F., Wright, J., Reale, D., Schrama, M., Roff, D.A., Van der Zee, E., Barber, I., 2009. Individual experience and evolutionary history of predation affect expression of heritable variation in fish personality and morphology. Proceedings of the Royal Society B-Biological Sciences 276, 1285-1293.

Dochtermann, N.A., Jenkins, S.H., 2007. Behavioural syndromes in Merriam's kangaroo rats (*Dipodomys merriami*): a test of competing hypotheses. Proceedings of the Royal Society B-Biological Sciences 274, 2343-2349.

Dochtermann, N.A., Roff, D.A., 2010. Applying a quantitative genetics framework to behavioural syndrome research. Proceedings of the Royal Society B-Biological Sciences 365, 4013-4020.

Drent, P.J., van Oers, K., van Noordwijk, A.J., 2003. Realized heritability of personalities in the great tit (*Parus major*). Proceedings of the Royal Society B-Biological Sciences 270, 45-51.

Dufty, A.M., Clobert, J., Moller, A.P., 2002. Hormones, developmental plasticity and adaptation. Trends in Ecology & Evolution 17, 190-196.

Dzieweczynski, T.L., Crovo, J.A., 2011. Shyness and boldness differences across contexts in juvenile three-spined stickleback *Gasterosteus aculeatus* from an anadromous population. Journal of Fish Biology 79, 776-788.

Earley, R., 2006. *Xiphophorus*: carving a niche towards a broader understanding of aggression and dominance. Zebrafish 3, 283-293.

Earley, R.L., Dugatkin, L.A., 2002. Eavesdropping on visual cues in green swordtail (Xiphophorus helleri) fights: a case for networking. Proceedings of the Royal Society of London Series B-Biological Sciences 269, 943-952.

Earley, R.L., Hsu, Y., 2008. Reciprocity between endocrine state and contest behavior in the killifish, *Kryptolebias marmoratus*. Hormones and Behavior 53, 442-451.

Earley, R.L., Tinsley, M., Dugatkin, L.A., 2003. To see or not to see: does previewing a future opponent affect the contest behavior of green swordtail males (*Xiphophorus helleri*)? Naturwissenschaften 90, 226-230.

Earley, R.L., Edwards, J.T., Aseem, O., Felton, K., Blumer, L.S., Karom, M., Grober, M.S., 2006. Social interactions tune aggression and stress responsiveness in a territorial cichlid fish (*Archocentrus nigrofasciatus*). Physiology & Behavior 88, 353-363.

Ellis, T., James, J.D., Stewart, C., Scott, A.P., 2004. A non-invasive stress assay based upon measurement of free cortisol released into the water by rainbow trout. Journal of Fish Biology 65, 1233-1252.

Ellis, T., James, J.D., Sundh, H., Fridell, F., Sundell, K., Scott, A.P., 2007. Non-invasive measurement of cortisol and melatonin in tanks stocked with seawater Atlantic salmon. Aquaculture 272, 698-706.

Evans, J.P., Gasparini, C., Pilastro, A., 2007. Female guppies shorten brood retention in response to predator cues. Behavioral Ecology and Sociobiology 61, 719-727.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. Pearson Education Ltd, Harlow, Essex, England.

Ferrari, C., Pasquaretta, C., Carere, C., Cavallone, E., von Hardenberg, A., Reale, D., 2013. Testing for the presence of coping styles in a wild mammal. Animal Behaviour 85, 1385-1396.

Fevolden, S.E., Refstie, T., Gjerde, B., 1993. Genetic and phenotypic parameters for cortisol and glucose stress response in Atlantic salmon and Rainbow trout. Aquaculture 118, 205-216.

Fischer, E.K., Harris, R.M., Hofmann, H.A., Hoke, K.L., 2014. Predator exposure alters stress physiology in guppies across timescales. Hormones and Behavior 65, 165-172.

Fletcher, T.C., 1997. Dietary effects on stress and health. In: Iwama, G.K., A.D.Pickering, Sumpter, J.P., C.B.Schreck (Eds.), Fish Stress and Health in Aquaculture. Cambridge University Press, Cambridge, pp. 223-246.

Fox, H.E., White, S.A., Kao, M.H.F., Fernald, R.D., 1997. Stress and dominance in a social fish. Journal of Neuroscience 17, 6463-6469.

Fraisse, F., Cockrem, J.F., 2006. Corticosterone and fear behaviour in white and brown caged laying hens. British Poultry Science 47, 110-119.

Francis, R.C., 1988. On the relationship between aggression and social dominance. Ethology 78, 223-237.

Franck, D., Ribowski, A., 1989. Escalating fights for rank-order position between male swordtails *(Xiphophorus helleri):* effects of prior rank-order experience and information transfer. Behavioral Ecology and Sociobiology 24, 133-143.
Franck, D., Ribowski, A., 1993. Dominance hierarchies of male green swordtails (*Xiphophorous helleri*) in nature. Journal of Fish Biology 43, 497-499.
Franck, D., Dikomey, M., Schartl, M., 2001. Selection and the maintenance of a colour pattern polymorphism in the green swordtail (*Xiphophorus helleri*). Behaviour 138, 467-486.
Franck, D., Hannes, R.-P., Lannfermann, H., Ribowski, A., 1985. Effects of social isolation on aggressiveness in fish with special reference to the swordtail (*Xiphophorous helleri*) Behavioural Processes 10 415-421.
Franck, D., Klamroth, B., Taebel-Hellwig, A., Schartl, M., 1998. Home ranges and satellite tactics of male green swordtails (*Xiphophorus helleri*) in nature. Behavioural Processes 43, 115-123.
Gabor, C.R., Contreras, A., 2012. Measuring water-borne cortisol in *Poecilia latipinna*:is the process stressful, can stress be minimized and is cortisol correlated with sex steroid release rates? Journal of Fish Biology 81, 1327-1339.
Giesing, E.R., Suski, C.D., Warner, R.E., Bell, A.M., 2011. Female sticklebacks transfer information via eggs: effects of maternal experience with predators on offspring. Proceedings of the Royal Society B-Biological Sciences 278, 1753-1759.
Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2009. ASReml user guide release 3.0. VSNi, Hemel Hempstead, UK.
Glenn, A.L., Raine, A., Schug, R.A., Gao, Y., Granger, D.A., 2011. Increased testosterone-to-cortisol ratio in psychopathy. Journal of Abnormal Psychology 120, 389-399.
Godø, O.R., Moksness, E., 1987. Growth and maturation of Norwegian Coastal Cod and Northeast Arctic Cod under different conditions. Fisheries Research 5, 235-242.
Gosling, S.D., 2001. From mice to men: what can we learn about personality from animal research? Psychological Bulletin 127, 45-86.
Gosling, S.D., John, O.P., 1999. Personality dimensions in nonhuman animals: A cross-species review. Current Directions in Psychological Science 8, 69-75.
Goymann, W., Wingfield, J.C., 2004. Allostatic load, social status and stress hormones: the costs of social status matter. Animal Behaviour 67, 591-602.
Gregory, T.R., Wood, C.M., 1999. The effects of chronic plasma cortisol elevation on the feeding behaviour, growth, competitive ability, and swimming performance of juvenile rainbow trout. Physiological and Biochemical Zoology 72, 286-295.
Guhl, A.M., Craig, J.V., Mueller, C.D., 1960. Selective breeding for aggressiveness in chickens. Poultry Science 39, 970-980.
Gutiérrez-Rodríguez, C., Shearer, A.E., Morris, M.R., De Queiroz, K., 2008. Phylogeography and monophyly of the swordtail fish species *Xiphophorus birchmanni* (Cyprinodontiformes, Poeciliidae). Zoologica Scripta 37, 129-139.
Haas-Andela, H., 1976. Invitro culture of embryos of Xiphophorine fish and their raising to adulthood. Zoologischer Anzeiger 197, 1-5.
Hadfield, J.D., 2010a. MCMCglmm course notes. http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf, accessed December 2013.
Hadfield, J.D., 2010b. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software 33, 1-22.
Hadfield, J.D., Wilson, A.J., Garant, D., Sheldon, B.C., Kruuk, L.E.B., 2010. The misuse of BLUP in ecology and evolution. American Naturalist 175, 116-125.
Hall, C.S., 1934. Emotional behavior in the rat 1. Defecation and urination as measures of individual differences in emotionality. Journal of Comparative Psychology 18, 385-403.
Hand, J.L., 1986. Resolution of social conflicts - dominance, egalitarianism, spheres of dominance, and game-theory. Quarterly Review of Biology 61, 201-220.
Hannes, R.P., 1984. Fighting changes levels of androgens and corticoids in winner and loser swordtails differently. Aggressive Behavior 10, 156-156.
Hessing, M.J.C., Hagelso, A.M., Schouten, W.G.P., Wiepkema, P.R., Vanbeek, J.A.M., 1994. Individual behavioural and physiological strategies in pigs. Physiology & Behavior 55, 39-46.

Hinde, R.A., Datta, S., 1981. Dominance - an intervening variable. Behavioral and Brain Sciences 4, 442-442.
Hixon, M.A., Anderson, T.W., Buch, K.L., Johnson, D.W., McLeod, J.B., Stallings, C.D., 2012. Density dependence and population regulation in marine fish: a large-scale, long-term field manipulation. Ecological Monographs 82, 467-489.
Horn, J.M., Plomin, R., Rosenman, R., 1976. Heritability of personality-traits in adult male twins. Behavior Genetics 6, 17-30.
Hrbek, T., Seekinger, J., Meyer, A., 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Molecular Phylogenetics and Evolution 43, 986-998.
Hsu, Y., Lee, I.H., Lu, C.-K., 2009. Prior contest information: mechanisms underlying winner and loser effects. Behavioral Ecology and Sociobiology 63, 1247-1257.
Hsu, Y.Y., Wolf, L.L., 2001. The winner and loser effect: What fighting behaviours are influenced? Animal Behaviour 61, 777-786.
Hsu, Y.Y., Earley, R.L., Wolf, L.L., 2006. Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. Biological Reviews 81, 33-74.
Huntingford, F.A., 1976. Relationship between anti-predator behavior and aggression among conspecifics in 3-spined stickleback, *Gasterosteus aculeatus*. Animal Behaviour 24, 245-260.
Huntingford, F.A., Metcalfe, N.B., Thorpe, J.E., Graham, W.D., Adams, C.E., 1990. Social dominance and body size in Atlantic salmon parr, *Salmo salar L*. Journal of Fish Biology 36, 877-881.
Huntingford, F.A., Andrew, G., Mackenzie, S., Morera, D., Coyle, S.M., Pilarczyk, M., Kadri, S., 2010. Coping strategies in a strongly schooling fish, the common carp *Cyprinus carpio*. Journal of Fish Biology 76, 1576-1591.
Jackson, W.M., 1991. Why do winners keep winning? Behavioral Ecology and Sociobiology 28, 271-276.
Jang, K.L., Livesley, W.J., Vernon, P.A., Jackson, D.N., 1996. Heritability of personality disorder traits: a twin study. Acta Psychiatrica Scandinavica 94, 438-444.
Johnson, E.O., Kamilaris, T.C., Chrousos, G.P., Gold, P.W., 1992. Mechanisms of stress - a dynamic overview of hormonal and behavioral homeostasis. Neuroscience and Biobehavioural Reviews 16, 115-130.
Johnson, J.C., Sih, A., 2005. Precopulatory sexual cannibalism in fishing spiders *(Dolomedes triton*): a role for behavioral syndromes. Behavioral Ecology and Sociobiology 58, 390-396.
Jones, K.A., Godin, J.-G.J., 2010. Are fast explorers slow reactors? Linking personality type and anti-predator behaviour. Proceedings of the Royal Society B-Biological Sciences 277, 625-632.
Kanda, L.L., Louon, L., Straley, K., 2012. Stability in activity and boldness across time and context in captive Siberian dwarf hamsters. Ethology 118, 518-533.
Kime, D.E., Manning, N.J., 1982. Seasonal patterns of free and conjugated androgens in the brown trout *Salmo trutta*. General and Comparative Endocrinology 48, 222-231.
Koolhaas, J.M., deBoer, S.F., Bohus, B., 1997. Motivational systems or motivational states: behavioural and physiological evidence. Applied Animal Behaviour Science 53, 131-143.
Koolhaas, J.M., de Boer, S.F., Coppens, C.M., Buwalda, B., 2010. Neuroendocrinology of coping styles: towards understanding the biology of individual variation. Frontiers in Neuroendocrinology 31, 307-321.
Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C., Ruis, M.A.W., Blokhuis, H.J., 1999. Coping styles in animals: current status in behavior and stress-physiology. Neuroscience and Biobehavioural Reviews 23, 925-935.
Korte, S.M., Koolhaas, J.M., Wingfield, J.C., McEwen, B.S., 2005. The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. Neuroscience and Biobehavioural Reviews 29, 3-38.
Koyama, N., 1970. Changes in dominance rank and division of a wild Japanese monkey troop in Arashiyama. Primates 11, 335-390.
Kruk, M.R., Halasz, J., Meelis, W., Haller, J., 2004. Fast positive feedback between the adrenocortical stress response and a brain mechanism involved in aggressive behavior. Behavioral Neuroscience 118, 1062-1070.

Kruuk, L.E.B., Slate, J., Wilson, A.J., 2008. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. Annual Review of Ecology Evolution and Systematics 39, 525-548.
Kruuk, L.E.B., Slate, J., Pemberton, J.M., Brotherstone, S., Guinness, F., Clutton-Brock, T., 2002. Antler size in red deer: heritability and selection but no evolution. Evolution 56, 1683-1695.
Kurvers, R.H.J.M., Nolet, B.A., Prins, H.H.T., Ydenberg, R.C., van Oers, K., 2012. Boldness affects foraging decisions in barnacle geese: an experimental approach. Behavioral Ecology 23, 1155-1161.
Lee, J.S.F., Berejikian, B.A., 2008. Effects of the rearing environment on average behaviour and behavioural variation in steelhead. Journal of Fish Biology 72, 1736-1749.
Liker, A., Szekely, T., 2005. Mortality costs of sexual selection and parental care in natural populations of birds. Evolution 59, 890-897.
Lincoln, G.A., 1972. Role of antlers in red deer. Journal of Experimental Zoology 182, 233-&.
Lindholm, A.K., Hunt, J., Brooks, R., 2006. Where do all the maternal effects go? Variation in offspring body size through ontogeny in the live-bearing fish *Poecilia parae*. Biology Letters 2, 586-589.
Ling, T.J., Forster, G.L., Watt, M.J., Korzan, W.J., Renner, K.J., Summers, C.H., 2009. Social status differentiates rapid neuroendocrine responses to restraint stress. Physiology and Behaviour 96, 218-232.
Lorenzen, K., Enberg, K., 2002. Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. Proceedings of the Royal Society of London Series B-Biological Sciences 269, 49-54.
Magellan, K., Kaiser, H., 2010a. The function of aggression in the swordtail, *Xiphophorus helleri*: resource defence. Journal of Ethology 28, 239-244.
Magellan, K., Kaiser, H., 2010b. Male aggression and mating opportunity in a poeciliid fish. African Zoology 45, 18-23.
Mariani, F., Pérez-Barahona, A., Raffi, N., Nakagawa, S., 2009. Life Expectancy and the Environment. Institute for the Study of Labor (IZA), Bonn, Germany.
Massault, C., Hellemans, B., Louro, B., Batargias, C., Van Houdt, J.K.J., Canario, A., Volckaert, F.A.M., Bovenhuis, H., Haley, C., de Koning, D.J., 2010. QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. Animal Genetics 41, 337-345.
Mayer, I., Borg, B., Schulz, R., 1990. Seasonal changes in and effect of castration/androgen replacement on the plasma-levels of five androgens in the male three-spined stickleback, *Gasterosteus aculeatus* L. General and Comparative Endocrinology 79, 23-30.
Meffe, G.K., Snelson, F.F., 1989. An ecological overview of poeciliid fishes. In: Meffe, G.K., Snelson, F.F. (Eds.), Ecology and Evolution of Livebearing Fishes (Poeciliidae). Prentice Hall, Englewood Cliffs, New Jersey, pp. 13-31.
Metcalfe, N.B., Monaghan, P., 2001. Compensation for a bad start: grow now, pay later? Trends in Ecology & Evolution 16, 254-260.
Meyer, A., Morrissey, J.M., Schartl, M., 1994. Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. Nature 368, 539-542.
Mommsen, T.P., Vijayan, M.M., Moon, T.W., 1999. Cortisol in teleosts: dynamics, mechanisms of action, and metabolic regulation. Reviews in Fish Biology and Fisheries 9, 211-268.
Moore, A.J., Brodie, E.D., Wolf, J.B., 1997. Interacting phenotypes and the evolutionary process. 1. Direct and indirect genetic effects of social interactions. Evolution 51, 1352-1362.
Moore, A.J., Haynes, K.F., Preziosi, R.F., Moore, P.J., 2002. The evolution of interacting phenotypes: genetics and evolution of social dominance. American Naturalist 160, S186-S197.
Moretz, J.A., 2003. Aggression and RHP in the northern swordtail fish, *Xiphophorus cortezi*: the relationship between size and contest dynamics in male-male competition. Ethology 109, 995-1008.
Morita, K., Fukuwaka, M.A., 2006. Does size matter most? The effect of growth history on probabilistic reaction norm for salmon maturation. Evolution 60, 1516-1521.
Morris, M.R., Batra, P., Ryan, M.J., 1992. Male-male competition and access to females in the swordtail *Xiphophorous nigrensis*. Copeia, 980-986.

Morris, M.R., Mussel, M., Ryan, M.J., 1995. Vertical bars on male *Xiphophorous multilineatus* - a signal that deters rival mates and attracts females. Behavioral Ecology 6, 274-279.
Morris, M.R., De Queiroz, K., Morizot, D.C., 2001. Phylogenetic relationships among populations of northern swordtails (*Xiphophorus*) as inferred from allozyme data. Copeia, 65-81.
Morrissey, M.B., Parker, D.J., Korsten, P., Pemberton, J.M., Kruuk, L.E.B., Wilson, A.J., 2012. The prediction of adaptive evolution: empirical application of the second theorem of selection and comparison to the breeder's equation. Evolution 66, 2399-2410.
Muller, H., 2012. Individual consistency in foraging behaviour and response to predator threat in the bumblebee *Bombus terrestris* (Hymenoptera: Apidae). Entomologia Generalis 34, 9-22.
Mutzel, A., Dingemanse, N.J., Araya-Ajoy, Y.G., Kempenaers, B., 2013. Parental provisioning behaviour plays a key role in linking personality with reproductive success. Proceedings of the Royal Society B-Biological Sciences 280.
Nakagawa, S., Schielzeth, H., 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biological Reviews 85, 935-956.
Netherton, J.D., Grober, M.S., Earley, R.L., 2004. Temporal decay of cortisol in green swordtail fish (*Xiphophorus helleri*) following aggressive encounters: differences between winners and losers? Hormones and Behavior 46, 117-118.
Niemelä, P.T., Vainikka, A., Lahdenpera, S., Kortet, R., 2012. Nymphal density, behavioral development, and life history in a field cricket. Behavioral Ecology and Sociobiology 66, 645-652.
Nilsson, J.A., Svensson, M., 1996. Sibling competition affects nestling growth strategies in marsh tits. Journal of Animal Ecology 65, 825-836.
Oliveira, R.F., Silva, J.F., Simoes, J.M., 2011. Fighting zebrafish: characterization of aggressive behavior and winner-loser effects. Zebrafish 8, 73-81.
Oliveira, R.F., Hirschenhauser, K., Carneiro, L.A., Canario, A.V.M., 2002. Social modulation of androgen levels in male teleost fish. Comparitive Biochemistry and Physiology Part-B: Biochemistry & Molecular Biology 132, 203-215.
Ostner, J., Heistermann, M., Schülke, O., 2008. Dominance, aggression and physiological stress in wild male Assamese macaques (*Macaca assamensis*). Hormones and Behavior 54, 613-619.
Øverli, Ø., Winberg, S., Pottinger, T.G., 2005. Behavioral and neuroendocrine correlates of selection for stress responsiveness in rainbow trout - a review. Integrative and Comparative Biology 45, 463-474.
Øverli, Ø., Pottinger, T.G., Carrick, T.R., Øverli, E., Winberg, S., 2002. Differences in behaviour between rainbow trout selected for high- and low-stress responsiveness. Journal of Experimental Biology 205, 391-395.
Øverli, Ø., Sorensen, C., Pulman, K.G.T., Pottinger, T.G., Korzan, W.J., Summers, C.H., Nilsson, G.E., 2007. Evolutionary background for stress-coping styles: relationships between physiological, behavioral, and cognitive traits in non-mammalian vertebrates. Neuroscience and Biobehavioural Reviews 31, 396-412.
Øverli, Ø., Korzan, W.J., Hoglund, E., Winberg, S., Bollig, H., Watt, M., Forster, G.L., Barton, B.A., Øverli, E., Renner, K.J., Summers, C.H., 2004. Stress coping style predicts aggression and social dominance in rainbow trout. Hormones and Behavior 45, 235-241.
Parker, G.A., 1974. Assessment strategy and evolution of fighting behavior. Journal of Theoretical Biology 47, 223-243.
Parzefall, J., 1973. Attraction and sexual cycle of poeciliids. In: Schroeder, J.H. (Ed.), Genetics and mutagenesis of fish. Springer-Verlag, Berlin, pp. 177-183.
Patrick, S.C., Charmantier, A., Weimerskirch, H., 2013. Differences in boldness are repeatable and heritable in a long-lived marine predator. Ecology and Evolution 3, 4291-4299.
Payne, R.W., Murray, D.M., Harding, S.A., Baird, D.B., Soutar, D.M., 2005. GenStat for Windows, Introduction. VSN International Ltd. , Hemel Hempstead, UK.
Pellis, S.M., McKenna, M.M., 1992. Intrinsic and extrinsic influences on play fighting in rats - effects of dominance, partners playfulness, temperament and neonatal exposure to testosterone propionate. Behavioral Brain Research 50, 135-145.
Pervin, L., John, O.P., 1997. Personality: theory and research. Wiley, New York.
Pickering, A.D., 1993. Growth and stress in fish production. Aquaculture 111, 51-63.

Pickering, A.D., Pottinger, T.G., 1989. Stress responses and disease resistance in salmonid fish - effects of chronic elevation of plasma-cortisol. Fish Physiology and Biochemistry 7, 253-258.
Pilia, G., Chen, W.-M., Scuteri, A., Orru, M., Albai, G., Dei, M., Lai, S., Usala, G., Lai, M., Loi, P., Mameli, C., Vacca, L., Deiana, M., Olla, N., Masala, M., Cao, A., Najjar, S.S., Terracciano, A., Nedorezov, T., Sharov, A., Zonderman, A.B., Abecasis, G.R., Costa, P., Lakatta, E., Schlessinger, D., 2006. Heritability of cardiovascular and personality traits in 6,148 sardinians. Plos Genetics 2, 1207-1223.
Pintor, L.M., Sih, A., Bauer, M.L., 2008. Differences in aggression, activity and boldness between native and introduced populations of an invasive crayfish. Oikos 117, 1629-1636.
Pollux, B.J.A., Pires, M.N., Banet, A.I., Reznick, D.N., 2009. Evolution of placentas in the fish family Poeciliidae: an empirical study of macroevolution. Annual Review of Ecology Evolution and Systematics 40, 271-289.
Pottinger, T.G., 2008. The stress response in fish - mechanisms, effects and measurement. In: Branson, E.J. (Ed.), Fish Welfare. Blackwell publishing.
Pottinger, T.G., 2010. A multivariate comparison of the stress response in three salmonid and three cyprinid species: evidence for inter-family differences. Journal of Fish Biology 76, 601-621.
Pottinger, T.G., Carrick, T.R., 1999. Modification of the plasma cortisol response to stress in rainbow trout by selective breeding. General and Comparative Endocrinology 116, 122-132.
Pottinger, T.G., Carrick, T.R., 2001. Stress responsiveness affects dominant-subordinate relationships in rainbow trout. Hormones and Behavior 40, 419-427.
Prenter, J., Taylor, P.W., Elwood, R.W., 2008. Large body size for winning and large swords for winning quickly in swordtail males, *Xiphophorus helleri*. Animal Behaviour 75, 1981-1987.
Preston, B.T., Stevenson, I.R., Pemberton, J.M., Coltman, D.W., Wilson, K., 2003. Overt and covert competition in a promiscuous mammal: the importance of weaponry and testes size to male reproductive success. Proceedings of the Royal Society B-Biological Sciences 270, 633-640.
Price, J., Sloman, L., Gardner, R., Gilbert, P., Rohde, P., 1994. The social competition hypothesis of depression. British Journal of Psychiatry 164, 309-315.
Raoult, V., Brown, C., Zuberi, A., Williamson, J.E., 2012. Blood cortisol concentrations predict boldness in juvenile mulloway (*Argyosomus japonicus*). Journal of Ethology 30, 225-232.
Rauchenberger, M., Kallman, K.D., Morizot, D.C., 1990. Monophyly and geography of the Rio Panuco Basin Mexico swordtails genus *Xiphophorous* with descriptions of four new species. American Museum Novitates, 1-41.
Reale, D., Dingemanse, N.J., Kazem, A.J.N., Wright, J., 2010. Evolutionary and ecological approaches to the study of personality. Philosophical Transactions of the Royal Society B-Biological Sciences 365, 3937-3946.
Réale, D., Festa-Bianchet, M., 2000. Quantitative genetics of life-history traits in a long-lived wild mammal. Heredity 85, 593-603.
Réale, D., Gallant, B.Y., Leblanc, M., Festa-Bianchet, M., 2000. Consistency of temperament in bighorn ewes and correlates with behaviour and life history. Animal Behaviour 60, 589-597.
Réale, D., Reader, S.M., Sol, D., McDougall, P.T., Dingemanse, N.J., 2007. Integrating animal temperament within ecology and evolution. Biological Reviews 82, 291-318.
Relyea, R.A., 2004. Fine-tuned phenotypes: Tadpole plasticity under 16 combinations of predators and competitors. Ecology 85, 172-179.
Ribowski, A., Franck, D., 1993. Demonstration of strength and concealment of weakness in escalating fights of male swordtails (*Xiphophorous helleri*). Ethology 93, 265-274.
Roberts, B.W., DelVecchio, W.F., 2000. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. Psychological Bulletin 126, 3-25.
Roff, D.A., 2000. Trade-offs between growth and reproduction: an analysis of the quantitative genetic evidence. Journal of Evolutionary Biology 13, 434-445.
Ronning, B., Moe, B., Bech, C., 2005. Long-term repeatability makes basal metabolic rate a likely heritable trait in the zebra finch *Taeniopygia guttata*. Journal of Experimental Biology 208, 4663-4669.

Rosen, D.E., 1979. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. Bulletin of the American Museum of Natural History 162, 269-375.
Rosen, D.E., Gordon, M., 1953. Functional anatomy and evolution of male genitalia in poeciliid fishes. Zoologica [New York] 38, 1-47.
Rosen, D.E., Bailey, R.M., 1963. The Poeciliid fishes (Cyprinodontiformes), their structure, zoogeography and systematics. Bulletin of American Museum of Natural History 126, 1-176.
Rosenthal, G.G., Evans, C.S., 1998. Female preference for swords in *Xiphophorus helleri* reflects a bias for large apparent size. Proceedings of the National Academy of Sciences of the United States of America 95, 4431-4436.
Rosenthal, G.G., Evans, C.S., Miller, W.L., 1996. Female preference for dynamic traits in the green swordtail, *Xiphophorus helleri*. Animal Behaviour 51, 811-820.
Rosenthal, G.G., de la Rosa Reyna, X.F., Kazianis, S., Stephens, M.J., Morizot, D.C., Ryan, M.J., Garcia de Leon, F.J., 2003. Dissolution of sexual signal complexes in a hybrid zone between the swordtails *Xiphophorus birchmanni* and *Xiphophorus malinche* (Poeciliidae). Copeia, 299-307.
Rosenthal, H.L., 1952. Observations on reproduction of the Poecilid *Lebistes reticulatus* (Peters). Biological Bulletin 102, 30-38.
Rothschild, M.J., 1986. Dynamics of marine fish populations. Harvard University Press, Cambridge, MA.
Rowe, D.K., Thorpe, J.E., 1990. Suppression of maturation in male atlantic slamon (*Salmo-salar L*) parr by reduction in feeding and growth during spring months. Aquaculture 86, 291-313.
Royle, N.J., Lindstrom, J., Metcalfe, N.B., 2005. A poor start in life negatively affects dominance status in adulthood independent of body size in green swordtails *Xiphophorus helleri*. Proceedings of the Royal Society B-Biological Sciences 272, 1917-1922.
Rudin, F.S., Briffa, M., 2012. Is boldness a resource-holding potential trait? Fighting prowess and changes in startle response in the sea anemone, *Actinia equina*. Proceedings of the Royal Society B-Biological Sciences 279, 1904-1910.
Ryan, M.J., Causey, B.A., 1989. Alternative mating-behaviour in the swordtails *Xiphophorus-nigrensis* and *Xiphophorus-pygmaeus* (Pisces, Poeciliidae). Behavioral Ecology and Sociobiology 24, 341-348.
Ryan, M.J., Keddyhector, A., 1992. Directional patterns of female mate choice and the role of sensory biases. American Naturalist 139, S4-S35.
Ryan, M.J., Hews, D.K., Wagner, W.E., 1990. Sexual selection on alleles that determine body size in the swordtail *Xiphophorus-nigrensis*. Behavioral Ecology and Sociobiology 26, 231-237.
Ryan, M.J., Pease, C.M., Morris, M.R., 1992. A genetic polymorphism in the swordtail *Xiphophorous nigrensis* - testing the prediction of equal fitnesses. American Naturalist 139, 21-31.
Scheiner, S.M., 1993. Genetics and evolution of phenotypic plasticity. Annual Review of Ecology and Systematics 24, 35-68.
Schjolden, J., Stoskhus, A., Winberg, S., 2005. Does individual variation in stress responses and agonistic behavior reflect divergent stress coping strategies in juvenile rainbow trout? Physiological and Biochemical Zoology 78, 715-723.
Schoener, T.W., 1983. Field experiments on interspecific competition. American Naturalist 122, 240-285.
Scott, A.P., Liley, N.R., 1994. Dynamics of excretion of 17-alpha,20-beta-dihydroxy-4-pregnen-3-one 20-sulfate, and of the glucuronides of testosterone and 17-beta-estradiol, by urine of reproductively mature male and female rainbow-trout (*Oncorhynchus-mykiss*). Journal of Fish Biology 44, 117-129.
Scott, A.P., Ellis, T., 2007. Measurement of fish steroids in water - a review. General and Comparative Endocrinology 153, 392-400.
Scott, A.P., Hirschenhauser, K., Bender, N., Oliveira, R., Earley, R.L., Sebire, M., Ellis, T., Pavlidis, M., Hubbard, P.C., Huertas, M., Canario, A., 2008. Non-invasive measurement of steroids in fish-holding water: important considerations when applying the procedure to behaviour studies. Behaviour 145, 1307-1328.
Scrimshaw, N.S., 1945. Embryonic development in Poecilid fishes. Biological Bulletin 88, 233-246.

Sebire, M., Katsiadaki, I., Scott, A.P., 2007. Non-invasive measurement of 11-ketotestosterone, cortisol and androstenedione in male three-spined stickleback (*Gasterosteus aculeatus*). General and Comparative Endocrinology 152, 30-38.

Selye, H., 1973. Evolution of stress concept. American Scientist 61, 692-699.

Sgoifo, A., De Boer, S.F., Buwalda, B., Korte-Bouws, G., Tuma, J., Bohus, B., Zaagsma, J., Koolhaas, J.M., 1998. Vulnerability to arrhythmias during social stress in rats with different sympathovagal balance. American Journal of Physiology-Heart Circulatory Physiology 275, H460-H466.

Sih, A., Bell, A., Johnson, J.C., 2004a. Behavioral syndromes: an ecological and evolutionary overview. Trends in Ecology & Evolution 19, 372-378.

Sih, A., Bell, A.M., Johnson, J.C., Ziemba, R.E., 2004b. Behavioral syndromes: An integrative overview. Quarterly Review of Biology 79, 241-277.

Sih, A., Crowley, P., McPeek, M., Petranka, J., Strohmeier, K., 1985. Predation, competition, and prey communities - a review of field experiments. Annual Review of Ecology and Systematics 16, 269-311.

Sih, A., Cote, J., Evans, M., Fogarty, S., Pruitt, J., 2012. Ecological implications of behavioural syndromes. Ecology Letters 15, 278-289.

Silberg, J., Pickles, A., Rutter, M., Hewitt, J., Simonoff, E., Maes, H., Carbonneau, R., Murrelle, L., Foley, D., Eaves, L., 1999. The influence of genetic factors and life stress on depression among adolescent girls. Archives of General Psychiatry 56, 225-232.

Sinn, D.L., Apiolaza, L.A., Moltschaniwskyj, N.A., 2006. Heritability and fitness-related consequences of squid personality traits. Journal of Evolutionary Biology 19, 1437-1447.

Sloman, K.A., Metcalfe, N.B., Taylor, A.C., Gilmour, K.M., 2001. Plasma cortisol concentrations before and after social stress in rainbow trout and brown trout. Physiological and Biochemical Zoology 74, 383-389.

Smiseth, P.T., Wright, J., Kolliker, M., 2008. Parent-offspring conflict and co-adaptation: behavioural ecology meets quantitative genetics. Proceedings of the Royal Society B-Biological Sciences 275, 1823-1830.

Smith, B.R., Blumstein, D.T., 2008. Fitness consequences of personality: a meta-analysis. Behavioral Ecology 19, 448-455.

Snelson, F.F., 1984. Seasonal maturation and growth of males in a natural population of *Poecilia latipinna*. Copeia, 252-255.

Sogard, S.M., 1997. Size-selective mortality in the juvenile stage of teleost fishes: a review. Bulletin of Marine Science 60, 1129-1157.

Sohn, J.J., Crews, D., 1977. Size-mediated onset of genetically determined maturation in the platyfish, *Xiphophorus maculatus* Proceedings of the National Academy of Sciences 74 4547-4548.

Stamps, J., Groothuis, T.G.G., 2010. The development of animal personality: relevance, concepts and perspectives. Biological Reviews 85, 301-325.

Stamps, J.A., 2007. Growth-mortality tradeoffs and 'personality traits' in animals. Ecology Letters 10, 355-363.

Stearns, S.C., Koella, J.C., 1986. The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. Evolution 40, 893-913.

Stram, D.O., Lee, J.W., 1994. Variance-components testing in the longitudinal mixed effects model. Biometrics 50, 1171-1177.

Summers, C.H., 2002. Social interaction over time, implications for stress responsiveness. Integrative and Comparative Biology 42, 591-599.

Sundstrom, L.F., Petersson, E., Hojesjo, J., Johnsson, J.I., Jarvi, T., 2004. Hatchery selection promotes boldness in newly hatched brown trout (*Salmo trutta*): implications for dominance. Behavioral Ecology 15, 192-198.

Sutherland, M.A., Huddart, F.J., 2012. The effect of training first-lactation heifers to the milking parlor on the behavioral reactivity to humans and the physiological and behavioral responses to milking and productivity. Journal of Dairy Science 95, 6983-6993.

Svartberg, K., Tapper, I., Temrin, H., Radesater, T., Thorman, S., 2005. Consistency of personality traits in dogs. Animal Behaviour 69, 283-291.

Taves, M.D., Desjardins, J.K., Mishra, S., Balshine, S., 2009. Androgens and dominance: sex-specific patterns in a highly social fish (*Neolamprologus pulcher*). General and Comparative Endocrinology 161, 202-207.
Thaker, M., Lima, S.L., Hews, D.K., 2009. Acute corticosterone elevation enhances antipredator behaviors in male tree lizard morphs. Hormones and Behavior 56, 51-57.
Thibault, R.E., Schultz, R.J., 1978. Reproductive adaptations among viviparous fishes (*Cyprinodontiformes Poeciliidea*). Evolution 32, 320-333.
Thomson, J.S., Watts, P.C., Pottinger, T.G., Sneddon, L.U., 2011. Physiological and genetic correlates of boldness: characterising the mechanisms of behavioural variation in rainbow trout, *Oncorhynchus mykiss*. Hormones and Behavior 59, 67-74.
Thomson, J.S., Watts, P.C., Pottinger, T.G., Sneddon, L.U., 2012. Plasticity of boldness in rainbow trout, *Oncorhynchus mykiss*: do hunger and predation influence risk-taking behaviour? Hormones and Behavior 61, 750-757.
Toms, C.N., Echevarria, D.J., Jouandot, D.J., 2010. A methodological review of personality-related studies in fish: focus on the shy-bold axis of behavior. International Journal of Comparative Psychology 23, 1-25.
Turner, C.L., 1937. Reproductive cycles and superfetation in poeciliid fishes. Biological Bulletin 72, 145-164.
van Oers, K., de Jong, G., van Noordwijk, A.J., Kempenaers, B., Drent, P.J., 2005. Contribution of genetics to the study of animal personalities: a review of case studies, pp. 1185-1206.
Van Reenen, C.G., O'Connell, N.E., Van der Werf, J.T.N., Korte, S.M., Hopster, H., Jones, R.B., Blokhuis, H.J., 2005. Responses of calves to acute stress: individual consistency and relations between behavioral and physiological measures. Physiology & Behavior 85, 557-570.
Vaz-Serrano, J., Ruiz-Gomez, M.L., Gjoen, H.M., Skov, P.V., Huntingford, F.A., Øverli, O., Hoglund, E., 2011. Consistent boldness behaviour in early emerging fry of domesticated Atlantic salmon (*Salmo salar*): decoupling of behavioural and physiological traits of the proactive stress coping style. Physiology & Behavior 103, 359-364.
Veenema, A.H., 2009. Early life stress, the development of aggression and neuroendocrine and neurobiological correlates: what can we learn from animal models? Frontiers in Neuroendocrinology 30, 497-518.
Verbeek, M.E.M., Boon, A., Drent, P.J., 1996. Exploration, aggressive behavior and dominance in pair-wise confrontations of juvenile male great tits. Behaviour 133, 945-963.
Verbeek, P., Iwamoto, T., Murakami, N., 2008. Variable stress-responsiveness in wild type and domesticated fighting fish. Physiology & Behavior 93, 83-88.
Visscher, P.M., 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. Twin Research and Human Genetics 9, 490-495.
Walling, C.A., Royle, N.J., Metcalfe, N.B., Lindstrom, J., 2007. Green swordtails alter their age at maturation in response to the population level of male ornamentation. Biology Letters 3, 144-146.
Walsh, B., Blows, M.W., 2009. Abundant genetic variation plus strong selection = multivariate genetic constraints: a geometric view of adaptation, Annual Review of Ecology Evolution and Systematics, pp. 41-59.
Walsh, R.N., Cummins, R.A., 1976. Open-field test - critical review. Psychological Bulletin 83, 482-504.
Warren, E.W., Callaghan, S., 1975. Individual differences in response to and open field test by guppy *Poecilia reticulata* (Peters). Journal of Fish Biology 7, 105-113.
Wcislo, W.T., 1989. Behavioural environments and evolutionary change. Annual Review of Ecology and Systematics 20, 137-169.
Webster, M.M., Ward, A.J.W., Hart, P.J.B., 2007. Boldness is influenced by social context in threespine sticklebacks (*Gasterosteus aculeatus*). Behaviour 144, 351-371.
Wechsler, B., 1995. Coping and coping strategies - a behavioural review. Applied Animal Behaviour Science 43, 123-134.
Wedemeyer, G.A., Barton, B.A., McLeay, D.J., 1990. Stress and acclimation. In: Schreck, C.B., Moyle, P.B. (Eds.), Methods for Fish Biology. American Fisheries Society, Bethesda MD, pp. 451-489.

Wendelaar Bonga, S.E., 1997. The stress response in fish. Physiological Reviews 77, 591-625.
Wesley, R.L., Cibils, A.F., Mulliniks, J.T., Pollak, E.R., Petersen, M.K., Fredrickson, E.L., 2012. An assessment of behavioural syndromes in rangeland-raised beef cattle. Applied Animal Behaviour Science 139, 183-194.
West-Eberhard, M.J., 1979. Sexual selection, social competition, and evolution. Proceedings of the American Philosophical Society 51, 222-234.
Wilson, A.J., 2014. Competition as a source of constraint on life history evolution in natural populations. Heredity 112, 70-78.
Wilson, A.J., Grimmer, A., Rosenthal, G.G., 2013. Causes and consequences of contest outcome: aggressiveness, dominance and growth in the sheepshead swordtail, *Xiphophorus birchmanni*. Behavioral Ecology and Sociobiology 67, 1151-1161.
Wilson, A.J., Boer, M.d., Arnott, G., A.Grimmer, 2011a. Integrating personality research and animal contest theory: aggressiveness in green swordtail *Xiphophorous helleri*. Plos One 6, e28024.
Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., Kruuk, L.E.B., Nussey, D.H., 2010a. An ecologist's guide to the animal model. Journal of Animal Ecology 79, 13-26.
Wilson, A.J., Morrissey, M.B., Adams, M.J., Walling, C.A., Guinness, F.E., Pemberton, J.M., Clutton-Brock, T.H., Kruuk, L.E.B., 2011b. Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. Journal of Evolutionary Biology 24, 772-783.
Wilson, D.S., 1998. Adaptive individual differences within single populations. Philosophical Transactions of the Royal Society B-Biological Sciences 353, 199-205.
Wilson, D.S., Clark, A.B., Coleman, K., Dearstyne, T., 1994. Shyness and boldness in humans and other animals. Trends in Ecology & Evolution 9, 442-446.
Wilson, K.S., C.Tucker, Denvir, M., Kenyon, C., Del-Pozo, J., 2010b. Zebrafish (*Danio rerio*) as a model of investigation of corticosteroid manipulation during early vertebrate development. Scottish Society for Experimental Proceedings (Abstract).
Winberg, S., Schjolden, J., Øverli, Ø., Pottinger, T., 2007. Stress and stress coping in fish, behavioural correlates and neuroendocrine mechanisms. Comparitive Biochemistry and Physiology Part-A: Molecular & Integrative Physiology 146, S77-S77.
Winge, O., 1937. Succession of broods in Lebistes. Nature 140, 467-467.
Wingfield, J.C., Maney, D.L., Breuner, C.W., Jacobs, J.D., Lynn, S., Ramenofsky, M., Richardson, R.D., 1998. Ecological bases of hormone-behavior interactions: the "emergency life history stage". American Zoologist 38, 191-206.
Wong, B.B.M., Rosenthal, G.G., 2006. Female disdain for swords in a swordtail fish. American Naturalist 167, 136-140.
Wong, S.C., Dykstra, M., Campbell, J.M., Earley, R.L., 2008. Measuring water-borne cortisol in convict cichlids (*Amatitlania nigrofasciata*): is the procedure a stressor? Behaviour 145, 1283-1305.
Wourms, J.P., 1981. Viviparity - the maternal -fetal relationship in fishes. American Zoologist 21, 473-515.
Young, A.J., Carlson, A.A., Monfort, S.L., Russell, A.F., Bennett, N.C., Clutton-Brock, T., 2006. Stress and the suppression of subordinate reproduction in cooperatively breeding meerkats. Proceedings of the National Academy of Sciences of the United States of America 103, 12005-12010.
Zar, J.H., 1996. *Biostatistical analysis*. Prentice-Hall, Upper Saddle River, NJ.
Zimmerer, E.J., Kallman, K.D., 1989. Genetic-basis for alternative reproductive tactics in the pygmy swordtail, *Xiphophorus-nigrensis*. Evolution 43, 1298-1307.

APPENDIX 1

SUPPLEMENTARY TABLES

Table A1.1 MCMCglmm analyses of the binary Emergence trait. Table shows a) the intraclass correlation (IC - the binary equivalent of the repeatability (see methods)) from a univariate model, and b) among-individual correlation (r_I) estimates from bivariate models of Emergence and each open field trial trait.

Model	Trait(s)		IC	r_I	95% HPD interval	
					Lower	Upper
a)	Emergence	-	0.090		0.024	0.177
b)	Emergence	Track Length		0.641	0.303	0.999
	Emergence	Activity		0.736	0.488	0.977
	Emergence	Area Covered		0.560	0.308	0.920
	Emergence	Time in Middle		0.573	0.300	0.872

Table A1.2 Univariate analyses of observed behavioural traits using the full (ALL), long- (LT) and short-term (ST) study data fitted using ASReml. The among- (V_I) and within-individual (residual) variance (V_R) estimates are presented for each trait along with repeatability (R). χ^2_1 and P-values relate to likelihood ratio tests of the significance of V_I. Note that for univariate models only we assume the test statistic to be asymptotically distributed as a mix of 50:50 χ^2_0 and χ^2_1 (following Visscher 2006). Behavioural traits studied: Track- length (TL), Activity (Act), Area Covered (AC), Time in Middle of tank (TIM), Emergence (Em). Behavioural traits studied: Track Length (TL), Activity (Act), Area Covered (AC), Time in Middle of tank (TIM), Emergence (Em).

Data	Trait	V_I (SE)	V_R (SE)	R (SE)	χ^2_1	P
ALL	TL	0.132 (0.025)	0.658 (0.029)	0.167 (0.029)	56.1	<0.001
	Act	0.159 (0.027)	0.668 (0.029)	0.193 (0.029)	75.0	<0.001
	AC	0.124 (0.026)	0.767 (0.033)	0.140 (0.027)	41.9	<0.001
	TIM	0.185 (0.029)	0.682 (0.030)	0.214 (0.029)	82.6	<0.001
	Em	0.058 (0.024)	0.889 (0.039)	0.061 (0.025)	6.88	0.005
LT	TL	0.143 (0.028)	0.689 (0.033)	0.172 (0.031)	41.5	<0.001
	Act	0.165 (0.028)	0.655 (0.031)	0.201 (0.031)	64.0	<0.001
	AC	0.141 (0.030)	0.768 (0.037)	0.155 (0.031)	31.8	<0.001
	TIM	0.206 (0.033)	0.693 (0.033)	0.229 (0.032)	69.9	<0.001
	Em	0.072 (0.028)	0.887 (0.043)	0.075 (0.029)	7.87	0.003
ST	TL	0.457 (0.154)	0.520 (0.067)	0.468 (0.093)	41.3	<0.001
	Act	0.369 (0.133)	0.571 (0.073)	0.393 (0.095)	29.2	<0.001
	AC	0.186 (0.089)	0.663 (0.085)	0.220 (0.089)	10.1	0.002
	TIM	0.248 (0.101)	0.594 (0.076)	0.295 (0.093)	17.2	<0.001
	Em	0.061 (0.069)	0.885 (0.113)	0.064 (0.071)	1.03	0.156

Table A1.3 Estimates of fixed effects (with standard errors in parentheses) from univariate mixed models of each behavioural trait for the data combined and for the long- (LT) and short-term (ST) studies. Significance was assessed using conditional *F* statistics and all models contained a random effect of individual identity. Coefficients are not presented for Stack, Treatment and Trial*Treatment due to their being multilevel factors. All individuals from ST were housed in the same stack therefore this covariate was not included in the ST analyses. Traits: Track Length (TL), Activity (Act), Area Covered (AC), Time in Middle (TIM), Emergence (Em).

Dataset	Response	Fixed Effect	Coefficient (SE)	DF	*F*	P
All	TL	Mean	1.90 (0.121)	1,332.7	1667	<0.001
		Sex	-0.046 (0.061)	1,348.8	0.56	0.454
		Day order	-0.006 (0.002)	11,376.1	8.01	0.005
		Stack		6,545.5	53.1	<0.001
		Trial	0.230 (0.028)	11,126.2	208	<0.001
		Treatment		3,339.3	1.56	0.201
		Trial*Treatment		31,375.9	2.68	0.046
	Act	Mean	3.22 (0.125)	1,347.1	3860	<0.001
		Sex	-0.145 (0.064)	1,365.2	5.07	0.026
		Day order	-0.003 (0.002)	11,373	1.96	0.164
		Stack		6,564.6	33.5	<0.001
		Trial	0.238 (0.028)	11,129.5	225	<0.001
		Treatment		3,353.4	3.86	0.01
		Trial*Treatment		31,374.5	4.75	0.003
	AC	Mean	2.80 (0.127)	1,339.8	2204	<0.001
		Sex	0.252 (0.063)	1,354.5	15.8	<0.001
		Day order	-0.006 (0.002)	11,376.9	7.57	0.006
		Stack		6,555.8	7.87	<0.001
		Trial	0.179 (0.030)	11,141.5	112	<0.001
		Treatment		3,347.2	2.37	0.071
		Trial*Treatment		31,363.8	2.24	0.083
	TIM	Mean	1.58 (0.128)	1,342.4	622	<0.001
		Sex	0.528 (0.067)	1,361.7	62.4	<0.001
		Day order	-0.009 (0.002)	11,368.7	15.6	<0.001
		Stack		6,559.80	9.52	<0.001
		Trial	0.075 (0.029)	11,119.7	10.4	0.001
		Treatment		3,348.2	0.85	0.47
		Trial*Treatment		31,367.3	6.13	<0.001
	Em	Mean	0.665 (0.130)	1,297.7	141.4	<0.001
		Sex	0.222 (0.060)	1,301.2	13.5	<0.001
		Day order	0.007 (0.003)	11,342	6.21	0.007
		Stack		6,525.1	9.53	<0.001
		Trial	-0.085 (0.032)	11,138	8.31	0.004
		Treatment		3,306.7	2.63	0.051
		Trial*Treatment		31,111.1	1.07	0.048
LT	TL	Mean	1.72 (0.144)	1,350.1	612	<0.001
		Sex	-0.043 (0.065)	1,348	0.44	0.505
		Day order	-0.007 (0.002)	11,219.6	9.80	0.002
		Stack		5,353.6	7.81	<0.001
		Trial	0.310 (0.043)	1,976.1	226	<0.001
		Treatment		3,354.8	1.34	0.263
		Trial*Treatment		3,980.2	0.26	0.853
	Act	Mean	3.07 (0.143)	1,350.7	2107	<0.001
		Sex	-0.164 (0.066)	1,349.2	6.17	0.014
		Day order	-0.004 (0.002)	11,219.5	3.16	0.078
		Stack		5,354.1	9.25	<0.001
		Trial	0.311 (0.042)	1,969.7	242	<0.001
		Treatment		3,355.2	2.89	0.036
		Trial*Treatment		3,973.6	0.67	0.571

Dataset	Response	Fixed Effect	Coefficient (SE)	DF	*F*	P
LT	AC	Mean	2.70 (0.150)	1,341.7	1466	<0.001
		Sex	0.244 (0.067)	1,339.4	13.3	<0.001
		Day order	-0.007 (0.002)	11,217.9	8.15	0.005
		Stack		5,345.4	5.21	<0.001
		Trial	0.282 (0.045)	1,973.9	99.6	<0.001
		Treatment		3,346.5	1.32	0.27
		Trial*Treatment		3,978.1	1.42	0.237
	TIM	Mean	1.67 (0.150)	1,349.4	588	<0.001
		Sex	0.540 (0.071)	1,348.4	58.4	<0.001
		Day order	-0.010 (0.002)	11,216.3	17.1	<0.001
		Stack		5,352.6	8.49	<0.001
		Trial	0.075 (0.043)	1,962.3	2.30	0.132
		Treatment		3,353.7	1.27	0.285
		Trial*Treatment		3,966.2	3.03	0.029
	Em	Mean	0.654 (0.155)	1,336.3	144	<0.001
		Sex	0.198 (0.064)	1,330.8	9.46	0.002
		Day order	0.009 (0.003)	11,179.3	7.51	0.007
		Stack		5,342.4	4.67	<0.001
		Trial	-0.085 (0.049)	1,983.7	7.90	0.005
		Treatment		3,340.8	1.77	0.153
		Trial*Treatment		3,995.8	1.39	0.244
ST	TL	Mean	2.49 (0.508)	1,26.9	33.2	<0.001
		Sex	0.064 (0.267)	1,38.2	0.06	0.81
		Day order	0.013 (0.007)	1,122.2	4.11	0.046
		Trial	-0.029 (0.065)	1,121	5.12	0.027
		Treatment		3,27.2	0.72	0.547
		Trial*Treatment		3,120.9	4.35	0.006
	Act	Mean	3.14 (0.521)	1,26.9	70.0	<0.001
		Sex	0.270 (0.253)	1,35.7	1.14	0.292
		Day order	0.011 (0.007)	1,122.6	2.78	0.1
		Trial	0.031 (0.068)	1,121.1	7.58	0.007
		Treatment		3,27.1	0.99	0.411
		Trial*Treatment		3,121	3.36	0.021
	AC	Mean	1.26 (0.542)	1,27	17.2	<0.001
		Sex	0.394 (0.212)	1,32.1	3.45	0.072
		Day order	0.005 (0.007)	1,124.4	0.51	0.474
		Trial	0.145 (0.073)	1,121.7	14.4	<0.001
		Treatment		3,27.1	2.82	0.058
		Trial*Treatment		3,121.5	0.65	0.588
	TIM	Mean	-0.027 (0.521)	1,27.1	3.99	0.056
		Sex	0.447 (0.224)	1,33.5	3.98	0.054
		Day order	0.001 (0.007)	1,123.6	0.01	0.904
		Trial	0.246 (0.069)	1,121.5	17.0	<0.001
		Treatment		3,27.2	2.87	0.056
		Trial*Treatment		3,121.4	1.46	0.231
	Em	Mean	1.53 (0.612)	1,26.9	19.3	<0.001
		Sex	0.438 (0.186)	1,28	5.53	0.026
		Day order	-0.005 (0.008)	1,135.4	0.43	0.512
		Trial	-0.062 (0.083)	1,125.7	0.73	0.395
		Treatment		3,27	1.69	0.193
		Trial*Treatment		3,123.1	0.79	0.502

Table A1.4 Tests for heterogeneity of variance structures across density treatments for each behavioural trait in the LT data sets. Presented are χ^2 statistics with associated P-values for comparing models with homogeneous and heterogeneous (i.e. treatment specific) among-individual and residual variances. Significant heterogeneity of variance components across density treatments is indicated for Track Length (TL) and Time in Middle (TIM) only. Treatment specific variance components for TL estimated under the heterogeneous model (not shown) demonstrate lower repeatability (SE) in the High/High treatment R = 0.122 (0.075), relative to other treatment classes (Low/Low R = 0.311 (0.073), Low/High R = 0.209 (0.087), High/Low R = 0.273 (0.097)). For TIM, repeatability (SE) is reduced in the Low/High treatment R = 0.063 (0.037), relative to other treatment classes (Low/Low R = 0.243 (0.050), High/High R = 0.212 (0.067), High/Low R = 0.206 (0.061)).

Trait	χ^2_6	P
Track Length (TL)	38.3	<0.001
Activity	7.18	0.30
Area Covered	1.99	0.92
Time in Middle (TIM)	21.0	0.002
Emergence	2.07	0.91

Table A1.5 Detail of morphological and physiological measurements for individual fish, where: Trial is the competition in which the individual participated; Fish is the identity assumed in the particular trial; SL is Standard Length; BD is Body Depth; SwL is sword length; LSA is Lateral Surface Area; Lat Swim is latency to swim; Lat Init is latency to initiate; Status is W, win and L, lose; PreF is pre-contest cortisol level; PostF is Post-contest cortisol level; SR is stress response; Esc denotes fight escalation (Y) per trial.

Trial	Fish	SL (mm)	BD (mm)	SwL (mm)	LSA (mm2)	Lat Swim (secs)	Lat Init (secs)	Status	Pre F (µg/dl)	Post F (µg/dl)	SR	Esc
1	A	46.5	14	23.4	674.4	10	63	L	3.8	1.223	-2.576	N
1	B	45.05	14.8	13	679.74	*	*	W	0.809	2.962	2.153	-
2	A	43.6	13.2	17.3	592.82	*	*	W	0.895	5.065	4.171	Y
2	B	43.8	13.3	13.4	595.94	67	98	L	2.258	3.807	1.549	-
4	A	40.1	11.9	12.7	489.89	7	11	L	3.049	2.63	-0.419	Y
4	B	39.8	12.2	11.9	497.46	*	*	W	1.044	1.44	0.396	-
5	A	35.05	11	14.2	399.75	3	*	L	1.337	4.554	3.218	Y
5	B	35.65	10.7	13.3	394.76	*	14	W	1.227	4.717	3.491	-
7	A	37.2	10.65	18.2	414.38	*	*	L	1.242	1.826	0.585	Y
7	B	37.2	10.75	17.75	417.65	103	109	W	3.35	1.695	-1.655	-
8	A	35.1	10.3	16.05	377.58	6	24	W	2.11	1.634	-0.476	Y
8	B	35.15	10	15.65	367.15	*	*	L	1.663	2.281	0.619	-
9	A	42.2	12.05	24.7	533.21	85	92	W	2.909	2.322	-0.587	Y
9	B	42.5	12.2	18.25	536.75	*	*	L	1.45	2.046	0.596	-
10	A	39.95	11.75	20.3	489.71	*	*	L	1.534	1.84	0.306	Y
10	B	40.35	12.1	18.5	506.74	51	196	W	1.315	2.394	1.078	-
11	A	48.8	13.9	19	697.32	242	*	W	0.889	1.369	0.48	N
11	B	46.6	14.15	20.55	679.94	*	254	L	3.292	4.259	0.967	-
12	A	32.6	9.25	9.3	310.85	127	184	W	0.159	0.805	0.646	Y
12	B	33.85	9.3	7.95	322.76	*	*	L	0.74	0.986	0.246	-
13	A	38.4	11.5	21.85	463.45	*	*	W	1.623	1.338	-0.285	Y
13	B	39.2	11.5	20.2	471	14	26	L	1.958	4.684	2.726	-
14	A	37.5	10.8	16	421	*	*	L	1.337	1.82	0.483	Y
14	B	37.4	10.9	17.6	425.26	240	465	W	1.514	0.673	-0.841	-
15	A	41.4	12.25	17.6	524.75	246	*	*	0.593	0.256	-0.337	N
15	B	40.1	12.35	17.5	512.74	*	304	*	1.06	0.321	-0.739	-
16	A	39.7	11.2	15	459.64	*	181	L	1.371	1.03	-0.34	N
16	B	38.7	11.3	15.1	452.41	101	*	W	0.38	0.889	0.508	-
17	A	34.85	10	12.05	360.55	66	1275	W	0.952	1.104	0.151	N
17	B	35.6	10	11.75	367.75	*	*	L	1.868	3.177	1.308	-
18	A	33.9	9.9	9.35	344.96	*	*	W	0.638	0.516	-0.122	Y
18	B	33.3	10	13.45	346.45	640	641	L	0.617	0.163	-0.454	-
19	A	39.85	10.9	20.05	454.42	*	1640	*	1.258	1.102	-0.156	Y
19	B	39.75	11.15	22.6	465.81	116	*	*	0.567	2.12	1.554	-
20	A	35.8	10.1	22.35	383.93	109	155	L	0.742	2.279	1.537	N
20	B	36.75	10	22.4	389.9	*	*	W	0.977	1.246	0.269	-
21	A	33.6	9.6	15.15	337.71	*	*	W	0.35	0.241	-0.108	N
21	B	33.3	9.25	16.9	324.93	27	30	L	0.205	0.892	0.687	-
22	A	35.6	10.3	14.5	381.18	278	*	L	0.102	0.327	0.224	N
22	B	34.6	10.4	13.7	373.54	*	354	W	1.132	1.151	0.019	-
23	A	41.9	12.3	11	526.37	*	*	W	0.578	0.596	0.018	Y
23	B	41.9	12.35	13.45	530.92	421	460	L	1.621	2.267	0.646	-
24	A	39.5	11.25	14.6	458.98	17	121	L	0.722	2.704	1.981	Y
24	B	40.05	10.95	16.75	455.3	*	*	W	1.993	2.179	0.186	-
25	A	33.4	9.6	18	338.64	61	96	L	2.876	2.271	-0.606	N
25	B	35	9.8	17.3	360.3	*	*	W	0.536	0.744	0.207	-
26	A	39.2	10.4	19.5	427.18	46	48	*	2.499	1.7	-0.799	N
26	B	37.65	10.5	16.3	411.63	*	*	*	1.686	1.814	0.128	-
27	A	34.15	10.05	10.8	354.01	142	*	L	0.12	0.642	0.522	N
27	B	34.85	10	10.5	359	*	189	W	0.429	0.329	-0.1	-
28	A	35	9.5	13.4	345.9	*	*	W	0.3	0.303	0.003	N
28	B	33.4	9.4	11.6	325.56	83	108	L	1.12	0.516	-0.604	-
29	A	34.4	10.12	11.3	359.43	540	*	W	0.069	0.358	0.29	N
29	B	34.42	9.8	18.85	356.17	*	586	L	0.468	2.099	1.63	-
30	A	32.85	9.26	17.5	321.69	*	*	L	0.179	0.094	-0.085	Y
30	B	33.15	10	9.3	340.8	554	3202	W	0.308	0.412	0.104	-

Table A1.6 Estimates of among-individual (V_I) and residual (V_R) variance for all traits with standard errors in parentheses. V_I can be interpreted as repeatability since (transformed) traits were scaled to standard deviation units. Models were fitted using ASReml and likelihood ratio tests used to determine the statistical significance of V_I assuming the test statistic to be asymptotically distributed as a of mix 50:50 χ^2_0 and χ^2_1 (following Visscher, 2006). Note however that the assumption of residual normality is violated for emREF, which was analysed here on the observed (i.e. 0/1) data scale. All models contained fixed effects as shown in full in supplemental Table S2.

Trait	V_I	V_R	χ^2	P
TL	0.256 (0.137)	0.737 (0.121)	8.77	0.002
ACT	0.291 (0.146)	0.708 (0.116)	11.0	<0.001
AC	0.191 (0.120)	0.792 (0.130)	5.26	0.011
TIM	0.054 (0.086)	0.914 (0.151)	0.49	0.243
TOR	0.229 (0.129)	0.750 (0.123)	7.34	0.004
emREF	0.140 (0.103)	0.736 (0.160)	3.07	0.04
F_{PRE}	0.078 (0.071)	0.604 (0.099)	1.89	0.085
$11KT_{PRE}$	0.200 (0.112)	0.590 (0.098)	7.62	0.003
F_{POST}	0.007 (0.074)	0.930 (0.153)	0.01	0.461
$11KT_{POST}$	0.000 (-)*	0.803 (0.120)	0	0.5

*With V_I constrained to positive parameter space the estimate was bound to zero such that no SE can be estimated.

Table A1.7 Estimated fixed effects from univariate mixed models of all traits. Models were fitted using ASReml including individual identity as a random effect (see Table S1 for variance component estimates). Conditional *F*-tests were used to assess significance of all fixed effects. Trial was fitted as a five level factor. The contrasts among factor levels are not shown here but are depicted in Supplemental Figure S1. *Day order* was fitted as a linear effect while *Stack* was a two level factor (Effect size indicates the difference for Stack 2 relative to Stack 1). A linear effect of fish *mass* was also included in models of endocrine traits.

Trait	Fixed effect	Effect size (SE)	DF	*F*	P
TL	*Mean*	1.43 (0.265)	1,18	71.8	<0.001
	Trial		4,74.2	1.87	0.136
	Day order	0.021 (0.016)	1,75.7	1.72	0.175
	Stack	0.180 (0.287)	1,18.7	0.39	0.539
ACT	*Mean*	1.27 (0.268)	1,18.1	61.5	<0.001
	Trial		4,74.2	1.50	0.228
	Day order	0.022 (0.015)	1,75.6	1.76	0.163
	Stack	0.265 (0.297)	1,18.7	0.80	0.383
AC	*Mean*	1.61 (0.259)	1,18	107	<0.001
	Trial		4,74.3	1.66	0.182
	Day order	0.019 (0.016)	1,76	1.10	0.24
	Stack	0.298 (0.268)	1,18.8	1.24	0.279
TIM	*Mean*	0.895 (0.247)	1,16.4	58.1	<0.001
	Trial		1,73.1	0.83	0.513
	Day order	0.023 (0.017)	1,76.4	1.86	0.179
	Stack	0.459 (0.222)	1,17.7	4.28	0.052
TOR	*Mean*	1.22 (0.261)	1,18	65.5	<0.001
	Trial		4,74.2	1.39	0.247
	Day order	0.028 (0.016)	1,75.5	3.26	0.077
	Stack	0.329 (0.278)	1,18.6	1.40	0.251
emREF	*Mean*	0.853 (0.428)	1,17.4	28.0	<0.001
	Trial		4,71.3	4.42	0.003
	Day order	-0.013 (0.016)	1,73.0	0.67	0.42
	Stack	-0.513 (0.244)	1,17.9	4.42	0.05
F_{PRE}	*Mean*	6.93 (0.216)	1,17.1	4215	<0.001
	Trial		4,74.7	3.44	0.012
	Day order	-0.023 (0.014)	1,76	2.72	0.106
	Stack	-1.09 (0.216)	1,18.9	25.7	<0.001
	Mass	-0.039 (0.332)	1,27.6	0.01	0.908
$11KT_{PRE}$	*Mean*	13.2 (0.243)	1,17	9828	<0.001
	Trial		4,74.4	7.82	<0.001
	Day order	-0.056 (0.014)	1,74.8	16.4	<0.001
	Stack	-0.390 (0.270)	1,18.7	2.09	0.164
	mass	0.466 (0.393)	1,34.8	1.40	0.245
F_{POST}	*Mean*	8.85 (0.244)	1,16.7	7766	<0.001
	Trial		4,74.4	1.89	0.121
	Day order	-0.024 (0.017)	1,77	1.94	0.17
	Stack	-0.465 (0.216)	1,18.6	4.65	0.044
	WT	0.349 (0.342)	1,23	1.04	0.319
$11KT_{POST}$	*Mean*	9.82 (0.226)	1,89	10696	<0.001
	Trial		4,89	4.88	0.001
	Day order	-0.037 (0.016)	1,89	5.08	0.028
	Stack	-0.368 (0.201)	1,89	3.36	0.072
	WT	0.003 (0.314)	1,89	0.13	0.714

Table A1.8 Eigen vector decomposition of the **I** matrix estimated among behavioural traits observed in the modified open field trial prior to the simulated predator attack.

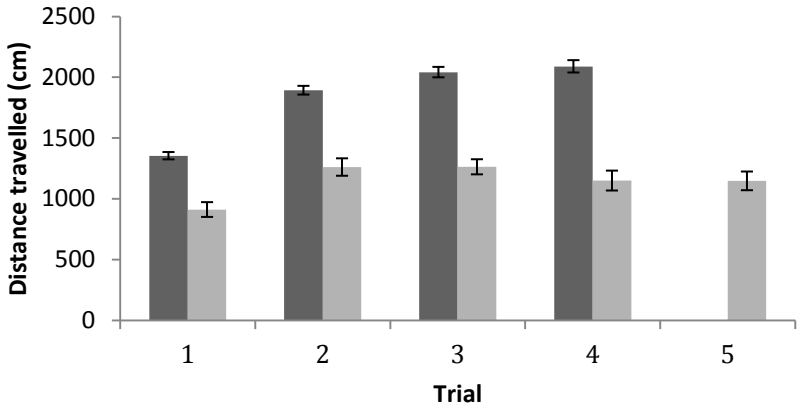
Eigen Vector	1	2	3	4	5
Eigen Value	1.11	0.029	0.013	0.002	0.000
Percentage of variance explained	96.2	2.52	1.12	0.167	0.009
Trait loadings:					
Track Length	0.491	-0.444	0.198	0.513	0.510
Activity	0.523	0.058	0.325	0.184	-0.764
Area Covered	0.430	-0.419	-0.689	-0.386	-0.123
Time in Middle	0.274	0.697	-0.505	0.412	0.125
Time Out of Refuge	0.474	0.373	0.354	-0.620	0.355

APPENDIX 2

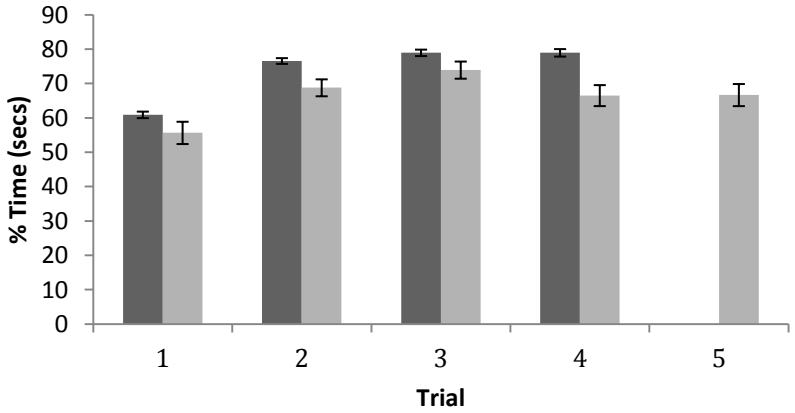
SUPPLEMENTARY FIGURES

Figure A2.1 Summary of raw behavioural data showing observed mean (± standard error) by Trial in long- (dark grey) and short- (light grey) term studies for a) Track Length, b) Activity, c) Area covered, d) Time in middle, and e) Emergence is represented as a percentage and therefore does not have an associated error. The long-term study (LT) comprised four Trials, while there were five Trials in the short-term (ST) study

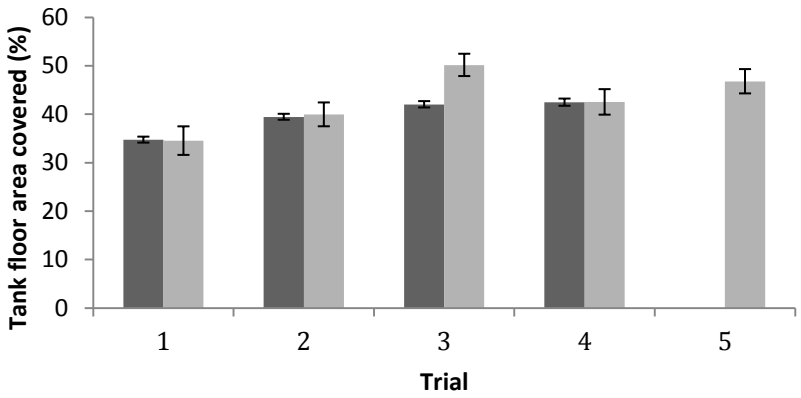
a)



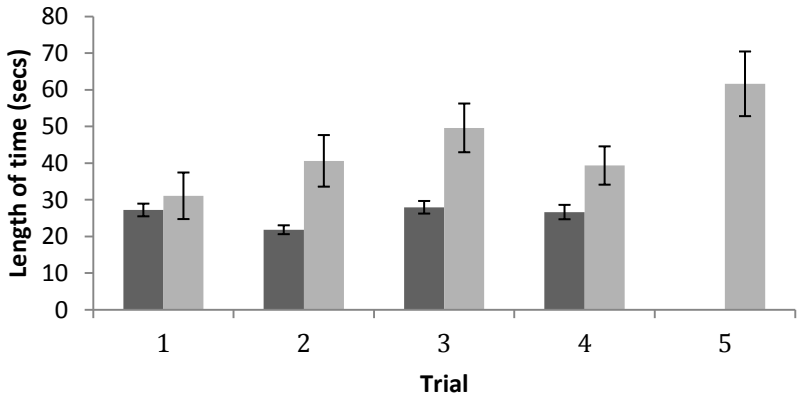
b)



c)


Tank floor area covered (%)
60
50
40
30
20
10
0
1
2
3
4
5
Trial

d)


Length of time (secs)
80
70
60
50
40
30
20
10
0
1
2
3
4
5
Trial

e)

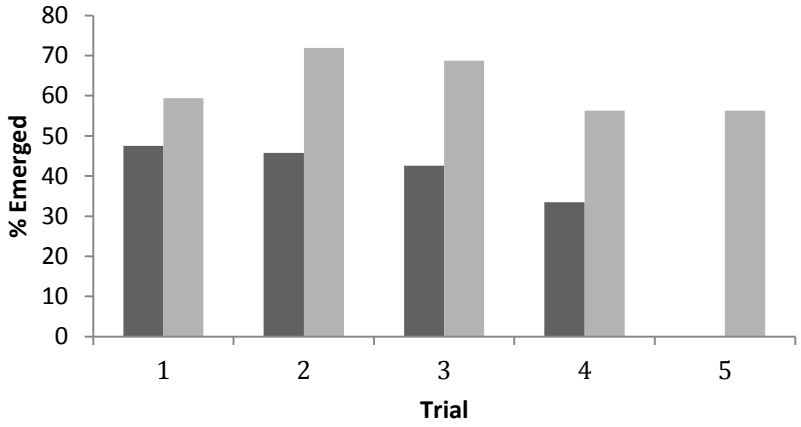

% Emerged
80
70
60
50
40
30
20
10
0
1
2
3
4
5
Trial

Figure A2.2 Posterior distribution of the intra-class correlation for the binary trait of Emergence modelled in MCMCglmm. The posterior mode for the intraclass correlation, IC = 0.109, 95% HPD interval 0.041 – 0.194.

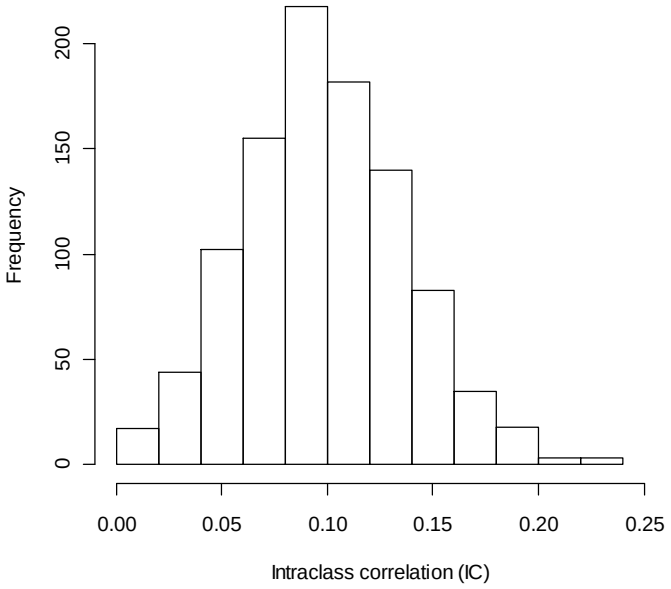


Figure A2.3 Parametric bootstrap distributions for θ, the estimated angle between $EV1_{LT}$ and $EV1_{ST}$ in the case that a) $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ are equal to their REML estimates, and b) $\mathbf{I}_{LT}$ and $\mathbf{I}_{ST}$ are equal such that the true angle between leading eigenvectors is zero. Distributions are based on 5000 pairs of simulated matrices (see main text for further details).

a) b)

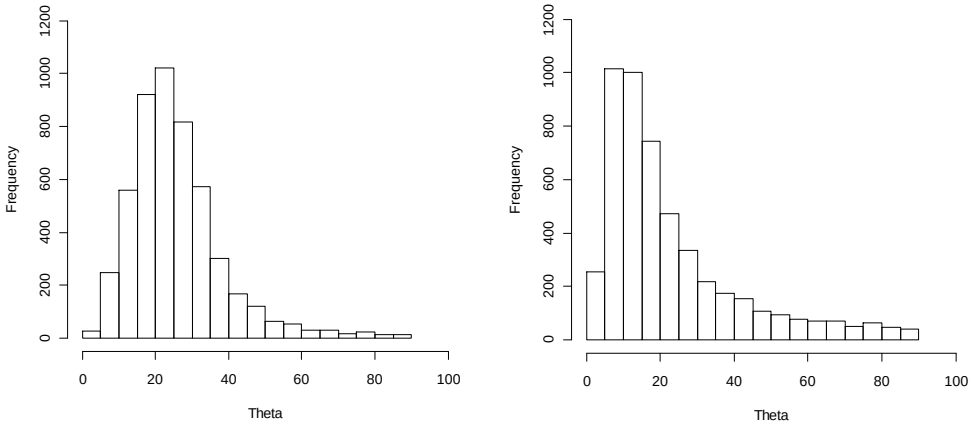
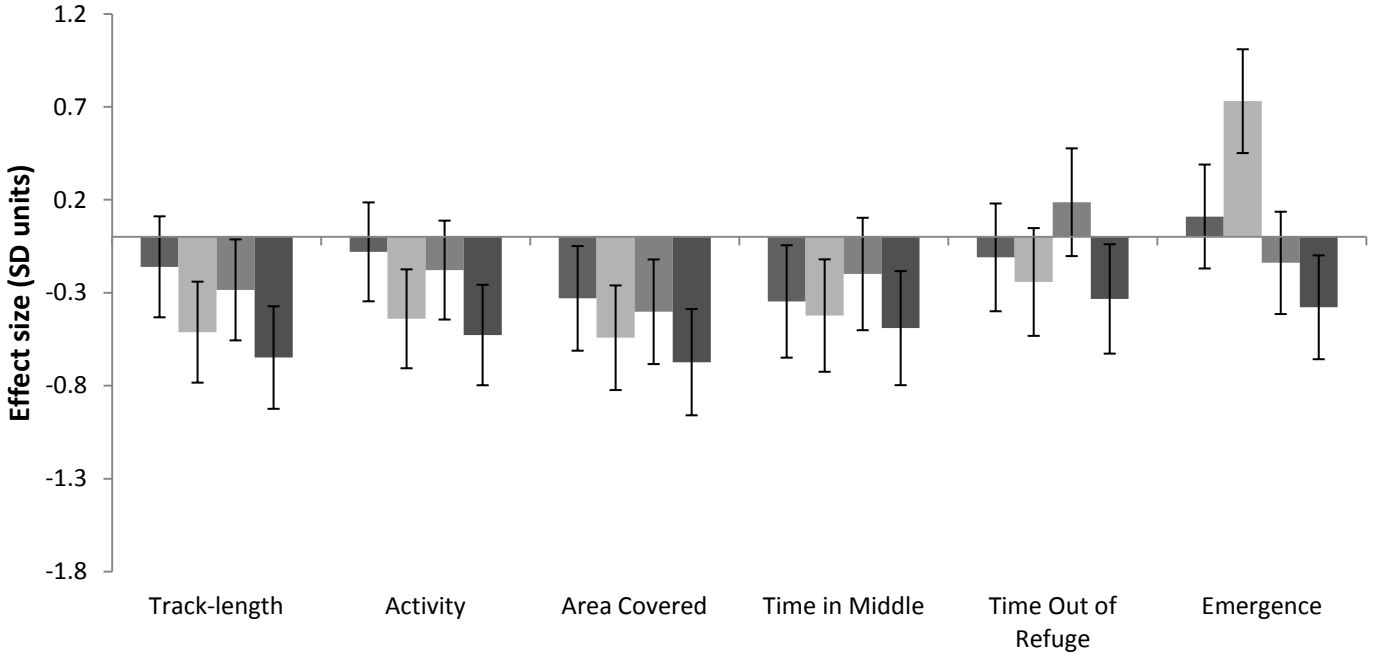
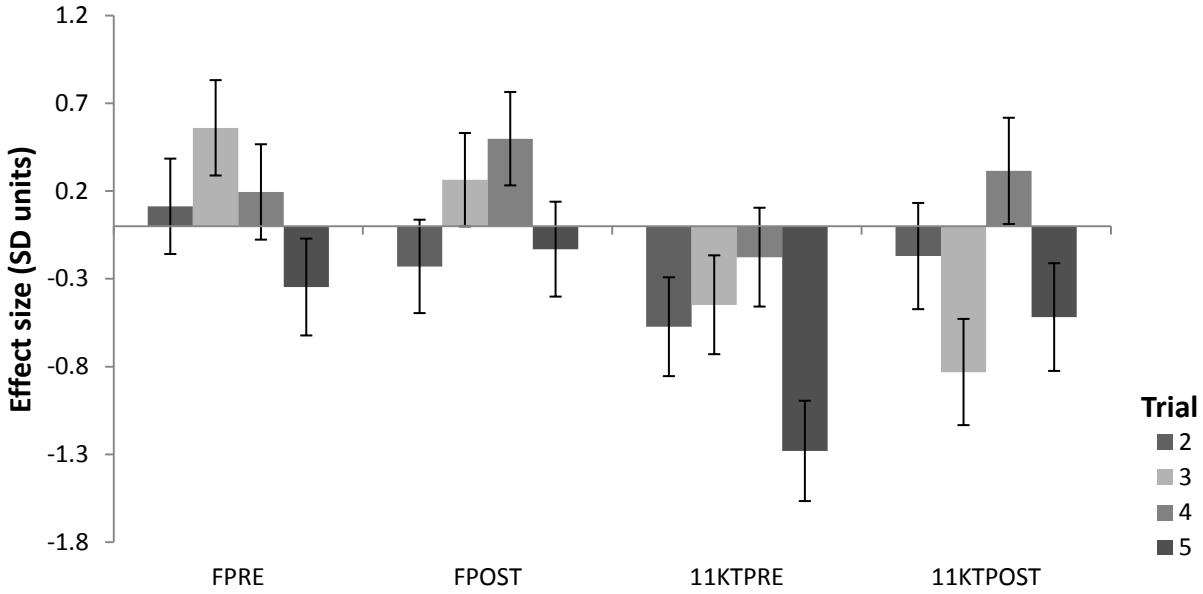


Figure A2.4 Estimated effects of trial number (*Trial*) from univariate models of a) behavioural and b) endocrine traits (see Appendix 1, Tables A1.6 and A1.7 for full results). *Trial* was fitted as a multilevel factor and effect sizes are shown (in standard deviation units) relative to the predicted mean at trial 1. Error bars denote ± SE. There is a general pattern of decrease with trial number across behavioural traits, though *Trial* was only statistically significant for emergence from refuge (Appendix 1, Table A1.7), a result driven by notably higher emergence rates in Trial 3. Significant mean differences among trials were found for all endocrine traits except F_{POST}.

a)

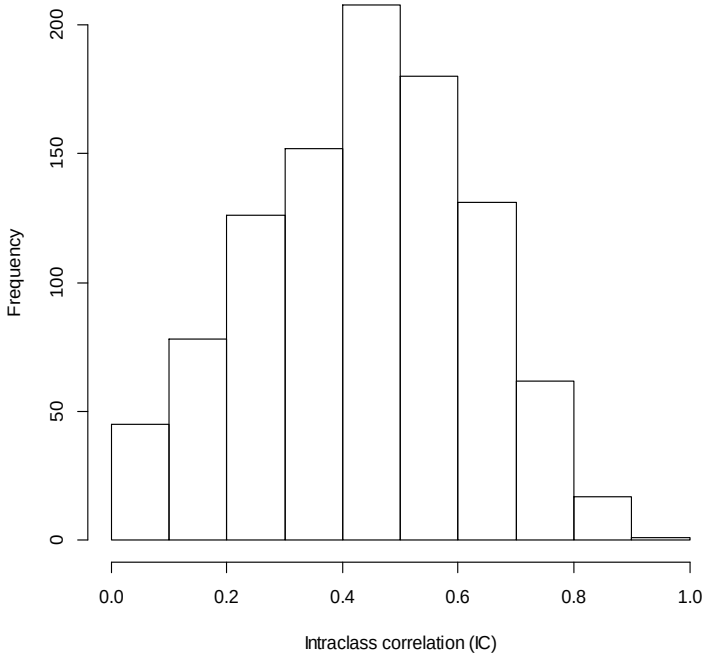


b)


1.2
0.7
0.2
-0.3
-0.8
-1.3
-1.8
Effect size (SD units)
FPRE
FPOST
11KTPRE
11KTPOST
Trial
2
3
4
5

N

Figure A2.5 Posterior distribution of the intraclass correlation (IC) of the binary trait, emergence from refuge (emREF) from an analysis modelling emREF as a categorical trait in MCMCglmm. See main text for full model details.



APPENDIX 3

Table A3.1 Timescale of data collection for each of the 6 stacks (A-G) presented in thesis chapters 2 and 5. Individuals were measured upon stack entry and four-weekly thereafter. Open field trials (OFT) and emergence and exploration trials (EET) were observed on consecutive weeks. In-tank observations (ITO) were observed (on males only) during periods of time alternate to OFT and EETs. Two sets of dyadic trials were also observed for males only, one each at the end of part 1 (prior to-) and part 2 (post-) density swap. Age represents the age in weeks of the youngest fish in each stack at the time of observation and the stage of the experiment (weeks) is given, along with the number of individuals observed and the percentage of the total represented (not presented for male only measures (-)).

Stack	Entry	Measure 1	Measure 2	Measure 3	OFT 1	EET 1	Measure 4
A	01/03/2011	29/03/2011	26/04/2011	24/05/2011	31/05/2011	07/06/2011	21/06/2011
B	25/03/2011	22/04/2011	20/05/2011	17/06/2011	24/06/2011	01/07/2011	15/07/2011
D&E	09/05/2011	06/06/2011	04/07/2011	01/08/2011	08/08/2011	15/08/2011	29/08/2011
F	08/06/2011	06/07/2011	03/08/2011	31/08/2011	07/09/2011	14/09/2011	28/09/2011
G	21/07/2011	18/08/2011	15/09/2011	13/10/2011	20/10/2011	27/10/2011	10/11/2011
Age	12	16	20	24	25	26	28
Weeks	0	4	8	12	13	14	16
No. fish	384	384	383	380	378	378	374
% total	100	100	99.7	99.0	98.4	98.4	97.4

Stack	ITO 1	Measure 5	OFT 2	EET 2	Measure 6	ITO 2	Dyadic 1
A	11/07/2011	19/07/2011	26/07/2011	02/08/2011	16/08/2011	23/08/2011	30/08/2011
B	29/07/2011	12/08/2011	22/08/2011	29/08/2011	09/09/2011	16/09/2011	23/09/2011
D&E	19/09/2011	26/09/2011	03/10/2011	10/10/2011	24/10/2011	31/10/2011	07/11/2011
F	12/10/2011	26/10/2011	02/11/2011	09/11/2011	23/11/2011	30/11/2011	07/12/2011
G	24/11/2011	08/12/2011	15/12/2011	22/12/2011	05/01/2012	12/01/2012	19/01/2012
Age	30	32	33	34	36	37	38
Weeks	18	20	21	22	24	25	26
No. fish	158	369	356	356	359	152	152
% total	-	96.1	92.7	92.7	93.5	-	-

Cont...

Density swap

Stack	Measure 7	ITO 3	Measure 8	OFT 3	EET 3	Measure 9	ITO 4
A	13/09/2011	27/09/2011	11/10/2011	18/10/2011	25/10/2011	08/11/2011	22/11/2011
B	07/10/2011	21/10/2011	04/11/2011	11/11/2011	18/11/2011	02/12/2011	16/12/2011
D&E	21/11/2011	05/12/2011	19/12/2011	26/12/2011	02/01/2012	16/01/2012	30/01/2012
F	21/12/2011	04/01/2012	18/01/2012	25/01/2012	01/02/2012	15/02/2012	29/02/2012
G	02/02/2012	16/02/2012	01/03/2012	08/03/2012	15/03/2012	29/03/2012	12/04/2012
Age	40	42	44	45	46	48	50
Weeks	28	30	32	33	34	36	38
No. fish	336	146	313	291	291	277	141
% total	87.5	-	81.5	75.8	75.8	72.1	-

Stack	Measure 10	OFT 4	EET 4	Measure 11	ITO 5	Dyadic 2	Measure 12
A	06/12/2011	13/12/2011	20/12/2011	03/01/2012	10/01/2012	17/01/2012	31/01/2012
B	30/12/2011	06/01/2012	13/01/2012	27/01/2012	03/02/2012	10/02/2012	24/02/2012
D&E	13/02/2012	20/02/2012	27/02/2012	12/03/2012	19/03/2012	26/03/2012	09/04/2012
F	14/03/2012	21/03/2012	28/03/2012	11/04/2012	18/04/2012	25/04/2012	09/05/2012
G	26/04/2012	03/05/2012	10/05/2012	24/05/2012	25/05/2012	07/06/2012	21/06/2012
Age	52	53	54	56	56	58	60
Weeks	40	41	42	44	44	46	50
No. fish	232	220	220	197	74	74	183
% total	60.4	57.3	57.3	51.5	-	-	47.7