INDIVIDUAL AND GENETIC DIFFERENCES IN FEARFULNESS: EFFECTS ON FEATHER PECKING AND MEAT QUALITY OF TURKEYS

By

Marisa Anna Erasmus

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Animal Science – Doctor of Philosophy

2014

ABSTRACT

INDIVIDUAL AND GENETIC DIFFERENCES IN FEARFULNESS: EFFECTS ON FEATHER PECKING AND MEAT QUALITY OF TURKEYS

By

Marisa Anna Erasmus

Research with several species has demonstrated that individual differences in behavior or temperament influence behavior, well-being and economically important characteristics such as meat quality. In particular, differences in fearfulness are related to feather pecking in laying hens and meat quality in pigs and cattle. There has been scant research examining the behavior and well-being of commercial turkeys and even less is known about fear responses of turkeys or the possible relationships between fearfulness, feather pecking and meat quality. Fearfulness was examined in male commercial turkeys using four tests of fear, including tonic immobility (TI), open field, (OF), voluntary approach (VA), and novel object (NO) tests. Changes in fear responses over time were described and the reliability of all four tests of fear was assessed. Although behavior changed over time, most TI (latency to vocalize and number of vocalizations), OF (latency to ambulate, numbers of steps taken, squares entered, and defecations) and VA (latencies to move within two body lengths, one body length, approach and peck) test responses were reliable as indicated by significant, moderate correlation coefficients. The only reliable test measure for the NO test was the latency to peck the NO. The validity of the four tests was assessed by comparing inter-test correspondence at two ages (4-6 weeks and 8-10 weeks) using two strategies: 1) turkeys showing extreme behavioral responses in each test were selected and classified as high responders (HR) or low responders (LR) and their behavior was compared across test situations; 2) test measures from each test were correlated with test measures from all other tests. At 4-6 weeks, OF test behavior tended to be correlated with VA

test behavior. At 8-10 weeks, OF test behavior was correlated with TI test behavior and tended to be correlated with NO test behavior. The OF test was subsequently used to 1) examine the relationship between fearfulness and feather pecking in males and females of a commercial (COMM) and randombred (RB) turkey strain, and 2) examine the relationship between fearfulness and meat quality in COMM and RB males. No associations were found between physiological (corticosterone levels) and behavioral OF responses and feather pecking or meat quality. Some strain differences were observed in feather pecking behavior. Specifically, plumage damage was worse in COMM compared to RB turkeys. Furthermore, a large percentage of turkeys (> 39%) developed feather pecking. In conclusion, this research demonstrates that behavior of turkeys in fear tests is reliable under the conditions of this research, but caution is needed when interpreting responses to the NO test. New information was presented regarding feather pecking behavior of turkeys. In contrast to laying hens, there does not appear to be an association between feather pecking behavior and fear responses in an OF test, nor is there an association with meat quality, contrary to what has been demonstrated with pigs and cattle. Results may be useful in the development of welfare assessment programs for turkeys because fear tests are often used to assess welfare. Moreover, results provide insights into feather pecking behavior and identify important areas for future research regarding feather pecking and the relationship between stress, meat quality and turkey temperament.

To my family and especially Ouma Anna for teaching me to be creative, follow my dreams and "byt vas".

ACKNOWLEDGEMENTS

The road to the PhD is filled with bumps, forks, and adventures. Without the help and support of my family, friends and colleagues, I would not have been able to complete this journey.

First I would like to thank my adviser, Dr. Janice Swanson, for providing me with this extraordinary opportunity and the freedom to explore my own research questions. Janice, thank you so much for your patience, advice and for your willingness to listen no matter the issue. Thank you for reminding and encouraging me to stay focused on the goal. Your guidance and mentorship have shaped me into a more critical, confident scientist.

I am grateful to the members of my guidance committee, Dr. Dalen Agnew, Dr. Darrin Karcher, Dr. Janice Siegford and Dr. Tina Widowski. Dalen – thank you for your willingness to provide assistance and for helping me understand and incorporate immunohistochemistry into the study of animal behavior and welfare so that I may see behavior from a different perspective. Darrin – thank you for listening, providing advice and always having an open door. Thank you for pushing me to "think outside the box". I am also grateful to you for finding turkeys for my experiments; they would not have been possible otherwise. Janice – thank you for making me a part of your lab and for helping me develop my own personal mentorship and teaching style. Through working and collaborating with you I have come to appreciate and enjoy teaching. Tina – thank you for seeing me through another degree with turkeys. It is because of your guidance and mentorship that I am where I am today. I know I can always count on you for advice and support.

I would like to express my gratitude to Dr. Steve Bursian for his kindness and for always being willing to listen and provide advice. I would like to thank Angelo Napolitano for his patience, help and many lunch time conversations as I ran turkey project after turkey project.

This work would not have been possible without the help of the Animal Behavior and Welfare Group and all of the student volunteers who helped with data collection at the farm and watched hours and hours of turkey video. Many thanks to Dana Campbell, Courtney Daigle, Melissa Elischer, Jackie Jacobs, Maja Makagon, David Orban, and Ali Witaifi for their support and for keeping me (mostly) sane.

To my wonderful parents, Koos and Merlynn; sister, Annette; and brother, Louw – you have supported me and provided encouragement no matter which path I chose and reminded me to stay focused and stay positive. And lastly, to my husband, Jeffrey, this would not have been possible without you and I am glad we are able to share our adventures and complete this journey together.

PREFACE

Chapter 2 was formatted for publication in *Applied Animal Behaviour Science* and corresponds to the peer-reviewed version of Erasmus, M. and J. Swanson. 2014. Temperamental turkeys: reliability of behavioural responses to four tests of fear. *Applied Animal Behaviour Science.* 157: 100-108.

Chapter 3 was formatted for publication in *Animal Behaviour*, Chapter 4 was formatted for publication in *Physiology & Behavior* and Chapter 5 was formatted for publication in *Poultry Science.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

Ach = acetylcholine

ACTH = adrenocorticotropic hormone

AD = avoidance distance

ATP = adenosine triphosphate

COMM = commercial

CRH = corticotropin releasing hormone

d = day

DFD = dark firm dry

EEG = electroencephalogram

GABA = gamma-aminobutyric acid

h = hour

HFP = high feather pecking

HPA = hypothalamic-pituitary-adrenal

HR = high responder

HS = high stress

LAL = long attack latency

LFP = low feather pecking

LR = low responding

LS = low stress

LTI = long tonic immobility

min = minute

MP = moving person

NO = novel object

NPECK = non-feather pecking

OF = open field

PAG = periaqueductal gray matter

PECK = feather pecking

PSE = pale, soft exudative

PVN = paraventricular nucleus

QTL = quantitative trait locus

RB = randombred

s = second

SAL = short attack latency

SAM = sympathetic-adrenal-medullary

SP = stationary person

STI = short tonic immobility

TI = tonic immobility

TT = touch

VA = voluntary approach

wk = week

CHAPTER 1: GENERAL INTRODUCTION AND LITERATURE REVIEW

GENERAL INTRODUCTION

The wild turkey is native to North America and was domesticated in Mexico over 2000 years ago (Buss, 1989; Crawford, 1992). Initially, turkeys were selected for plumage pattern and color so that they could be used in shows, and it was only in the 20th century that selection for meat production and conformation became important (Appleby et al., 2004). Along with selection for production traits, there was also selection against dark feathering, resulting in most domestic turkeys being white by the 1960s (Appleby et al., 2004). Selection for increased breast muscle yield resulted in turkeys no longer being able to mate naturally and the modern commercial turkey industry relies on artificial insemination to produce turkeys.

Although commercial turkey production comprises over 240 million turkeys annually in North America (United States Department of Agriculture National Agricultural Statistics Service, 2014), there has been very little research into the behavior and welfare of turkeys compared to other poultry species. Feather pecking (the pecking at, plucking and sometimes removal of feathers from conspecifics) is one of the main welfare problems of commercial turkey production (Marchewka et al., 2013), but this abnormal behavior is poorly understood. The majority of information regarding injurious pecking, particularly feather pecking, has come from research with laying hens, and suggests that feather pecking is a complex, multifactorial problem (Dalton et al., 2014). Research specifically examining feather pecking in turkeys has focused on environmental factors (e.g. Crowe and Forbes, 1999; Duggan et al., 2014; Lewis et al., 1998; Martenchar et al., 2001; Moinard et al., 2001; Sherwin et al., 1999), and the effects of genetics, development and sex (Busayi et al., 2006; Hughes and Grigor, 1996). However, research with

laying hens has demonstrated that individual differences in behavior such as fearfulness also play an important role in the development of feather pecking (reviewed in Rodenburg et al., 2013).

Individual differences in behavior, or animal temperament, are also associated with other economically important characteristics such as meat quality. Meat quality is an important area of research in turkey production because of the occurrence of pale, soft, exudative meat (PSE), which is meat that is characterized by its pale color and inability to hold water (Sosnicki et al., 1998). Research examining turkey meat quality has primarily focused on genetic (e.g. Updike et al., 2005; Werner et al., 2008), environmental (e.g. Sarica et al., 2011) and pre-slaughter stress (e.g. Owens and Sams, 2000) factors. As with feather pecking, there is no scientific literature examining whether turkey temperament is related to meat quality.

Scientific literature regarding poultry temperament has focused on fearfulness, which can be assessed using a number of different fear tests. Although tests used for hens and quail are widely used for turkeys, research has not confirmed whether these tests are reliable or valid for use in turkeys and very little is known about how turkeys respond to fear tests, or about their fear-related behavior in general. Fear tests are also used to draw conclusions regarding animal well-being. Using fear tests for turkeys without knowing whether these tests are reliable or valid may lead to improper conclusions being drawn.

The objectives of this research were to provide more information regarding fear responses of turkeys and how these responses differ between sexes and genetic strains. Furthermore, this research examined whether fearfulness, as part of turkey temperament, is related to feather pecking and meat quality.

LITERATURE REVIEW

Personality, temperament, coping styles and fearfulness

Many studies have documented the existence of correlated groups of individual characteristics (Koolhaas et al., 2010) and personality traits (Stamps and Groothuis, 2010) in different animal species. Interest in studying whether animals differ in behavior and whether these differences are consistent across different situations and contexts has increased greatly in recent years. The study of animal temperament has many applications ranging from human medicine to animal breeding and production. Indeed, rodents are used as models for diseases and disorders such as depression, anxiety and post-traumatic stress (reviewed in Bourin et al., 2007, and Neumann et al., 2011). In production animal species, differences in temperament are associated with differences in disease susceptibility (see Koolhaas et al., 1999) and other economically important characteristics such as production, reproduction (see Jones and Boissy, 2011), and meat quality (e.g. Voisinet et al., 1997). There is evidence that individuals that differ in feather pecking behavior (discussed below) differ in other behavioral characteristics as well. Temperament is relevant to animal welfare because animals that differ in how they respond to stressors in their environment may be at risk of reduced welfare if they are not able to successfully cope with stressors. Animal well-being concerns "the state of the individual as regards its attempts to cope with its environment" (Broom, 1986), and animals may develop abnormal behaviors if they are unable to successfully cope with challenges (Wechsler et al., 1995).

The study of animal personality is complicated, and made even more so by the various terms that are used across disciplines. The terminology used in relation to individual variation in animal behavior includes personality, temperament, coping styles, behavioral profiles and

behavioral syndromes (see Groothuis and Carere, 2005; Koolhaas et al., 2010; and Stamps and Groothuis, 2010). In addition to the terms proactive and reactive, some other terms that are used to differentiate between animals at the extremes of the behavioral spectrum include shy vs. bold, active vs. passive, hawk vs. dove (Koolhaas et al., 2010) and fast vs. slow (see Groothuis and Carere, 2005). Some of the specific behaviors or traits that are most often studied as part of, and that are used to describe animal personality include activity, aggression, sociability, friendliness, emotionality, confidence, anxiety, exploration, psychopathy, affinity, playfulness, dominance, reactivity and fearfulness (reviewed in Gosling, 2001). Although there are some differences in the literature regarding the definitions of terms relating to personality; temperament, personality and coping styles are generally considered to have the same meaning (Cockrem, 2007; Stamps and Groothuis, 2010).

Personality is defined as individual behavioral differences that are consistent across contexts and time, and it concerns the behavior of individuals relative to one another rather than the "absolute level" of an individual's behavior (Stamps and Groothuis, 2010).

Behavioral syndromes refer to sets of behaviors that are correlated across different situations (Sih et al., 2004; Jacobs, 2009), where a situation is defined as "a given set of conditions at one point in time" (Sih et al., 2004). Stamps and Groothuis (2010) distinguish between behavioral syndrome and personality, stating that behavioral syndrome refers to a suite of behaviors that are correlated across time or across contexts; whereas personality refers to behavior that is correlated across both time and context. Groothuis and Carere (2005) use the term behavioral profiles as an adaptation of the definition of behavioral syndromes to include both behavior and underlying physiology.

Temperament is often used instead of personality because the term personality is perceived as being anthropomorphic (Gosling, 2001). In the animal behavior literature, temperament is defined as the "characteristics of individuals that describe and account for consistent patterns of feeling and behaving" and is usually described in terms of fearfulness or coping styles (Jones and Boissy, 2011).

The term coping styles has been used widely in the biomedical and stress physiology literature to refer to the strategies that individuals use to cope with environmental challenges (see Groothuis and Carere, 2005), and a considerable amount of research has been dedicated to the study of coping styles. Coping styles are defined as consistent behavioral and physiological stress responses of individuals (Koolhaas et al., 1999) and "alternative response patterns in reaction to a stressor" (Koolhaas et al., 2010). Therefore, coping styles are specifically concerned with the responses to stressors. A stressor is defined as a stimulus that induces stress, which is the state that results from activation of the hypothalamo-pituitary adrenocortical (HPA) axis and glucocorticoid (stress hormone) secretion due to a stressor (Cockrem, 2007). Two coping styles have been identified, namely proactive and reactive. Animals with proactive coping styles are characterized as being bold, fast, aggressive, less fearful and having high sympathetic nervous system activity and low corticosterone levels in response to stressors; whereas reactive animals are slow, shy and cautious, more fearful and have lower sympathetic nervous system activity and higher corticosterone levels in response to stressors (Cockrem, 2007; Carere et al., 2010). Research with mice and pigs suggests that the two types of coping styles differ in behavioral flexibility, such that animals with reactive coping styles behave based on environmental cues and stimuli and fare better under "variable and unpredictable environmental conditions" (Koolhaas et

al., 2010). In contrast, proactive animals are more likely to develop routines and do better under stable conditions (Koolhaas et al., 2010).

By definition, coping styles include both the behavioral as well as the physiological response to a stressor. Although individual differences in physiological stress responses are associated with individual differences in personality (reviewed in Cockrem, 2013), the link between behavior and physiology is still debated. Carere et al. (2010) propose three mechanisms to explain the relationship between behavior and glucocorticoid levels: behavior determines particular physiological responses, such as an increase in glucocorticoid secretion, or vice versa, or thirdly, behavior and stress physiology are both determined by other factors and are therefore correlated, but do not determine one another (i.e. one does not cause the other). The authors state that the third mechanism may provide a more comprehensive explanation of the link between behavior and physiology (Carere et al., 2010). Along the same lines, Koolhaas et al. (2010) distinguish between the quality (how animals respond (Coppens et al., 2010)) and the quantity (how strongly animals respond (Coppens et al., 2010)) of the behavioral and physiological responses to stress. Koolhaas et al. (2010) propose that qualitative and quantitative stress responses should be considered as independent dimensions because individuals can vary along both dimensions such that proactive and reactive copers may have either high or low stress reactivity (rather than proactive animals having only low stress reactivity and reactive animals having only high stress reactivity as stated in Cockrem (2007)).

Various genetic selection lines have been developed to study the relationship between different behavioral and physiological traits and these genetic lines possess characteristics that are consistent with proactive and reactive coping styles. Separate genetic lines of long (LAL) and short (SAL) attack latency mice have been developed based on aggressive behavior which is

measured using the latency of the mice to attack an intruder that is introduced into their cage (reviewed in Groothuis and Carere, 2005 and Cockrem, 2007). SAL mice display behavioral characteristics and show physiological responses that are consistent with the proactive coping style, whereas LAL mice possess characteristics consistent with the reactive coping style (see Groothuis and Carere, 2005; and Carere et al., 2010). For bird species, the FAST (proactive coping style) and SLOW (reactive coping style) genetic lines of the great tit (*Parus major*, a passerine bird*)* were developed based on whether the birds took a long (SLOW) or short (FAST) time to explore a novel environment or novel object (reviewed in Groothuis and Carere, 2005; Cockrem, 2007; Carere et al., 2010). Two lines of quail were developed based on their tonic immobility reactions (LTI – long tonic immobility and STI – short tonic immobility), as well as two lines differing in corticosterone responses to restraint stress (LS – low stress and HS – high stress) (see Cockrem, 2007). The quail lines are also consistent with proactive (LS and STI lines) and reactive (HS and LTI lines) coping styles (Cockrem, 2007). In chickens, high (HFP) and low (LFP) feather pecking lines of white leghorns were developed to examine differences in feather pecking behavior and the relationship between feather pecking and other behavior such as fearfulness (see Carere et al., 2010; Rodenburg et al., 2013). Although the physiological data (corticosterone levels in response to stress) suggest that HFP chickens are proactive copers and LFP chickens are reactive copers (Carere et al., 2010), behavioral data show the opposite: HFP chickens (proactive) are less active in a novel environment and have longer durations of tonic immobility compared to LFP chickens (reactive) (reviewed in Groothuis and Carere, 2005). However, behavioral data are not consistent with what is seen in other species, and thus there may be differences among animal species due to domestication or genetic selection that should be taken into consideration when classifying animals as having a particular behavioral profile

(Groothuis and Carere, 2005). Furthermore, behavior may differ depending on which test is used (e.g. TI vs. OF).

Although Cockrem's (2007) description of proactive and reactive coping styles includes differences in fearfulness, other researchers consider fearfulness to be an altogether separate dimension of personality that is related to emotionality and not coping style (e.g. Koolhaas et al., 1999; Jones and Boissy, 2011). Fearfulness refers to how susceptible an individual is to being frightened (Boissy, 1995; Jones, 1996), whereas fear is the individual's reaction to a perceived threat or danger (Forkman et al., 2007).

In stark contrast to the behavioral ecology literature where personality of wild bird species has been studied (e.g great tits - Carere et al., 2010, and geese - Kralj-Fišer et al., 2010), research with domestic bird species, primarily laying hens and quail, has focused on fear and fearfulness, and the terms temperament and personality have not traditionally been used. The interest in fear in particular, rather than other aspects of poultry temperament, can be traced back to the release of the Brambell Report in 1965 (Brambell, 1965) and the development of the Five Freedoms which include freedom from hunger and thirst, freedom from discomfort, freedom from pain, injury or disease, freedom to express normal behavior, and freedom from fear and distress (Farm Animal Welfare Council, 1979). Consequently, fear and fearfulness have been important areas of poultry behavior and welfare research and the ability to identify individual differences in fearfulness and characteristics of fearfulness have received a great deal of attention. Moreover, the development of tests for assessing fear and identifying fearful individuals has been a major part of fear research.

Tests used to assess fear in poultry

Because an individual animal's subjective state of fear can only be inferred indirectly, a number of different tests have been developed for assessing fear in poultry (reviewed in Forkman et al., 2007 and Jones, 1996). Stimuli that elicit fear have been classified into categories based on the properties of the stimuli (Gray, 1979; reviewed in Boissy, 1995 and Forkman et al., 2007). Categories of fear-eliciting stimuli include:

1) innate fears or stimuli associated with factors that are important to the species from an evolutionary standpoint (e.g. a particular predator)
2) the specific characteristics associated with the stimulus, such as novelty, suddenness, intensity and duration
3) stimuli that are perceived as threatening because previous experience and learning have changed the animal's perception of the stimulus (e.g. conditioned fear)
4) stimuli associated with social signals such as alarm calls or pheromones; for many farm animal species, social isolation induces fear

Considering the characteristics of stimuli that induce fear, the most common types of fear tests that are used for poultry expose poultry to predatory stimuli and/or restraint (e.g. tonic immobility, and human approach or avoidance tests), or novelty (e.g. novel arena (open field) and novel object tests), or sometimes elements of predatory stimuli and novelty are combined. The most common poultry fear tests include the tonic immobility, open field, novel object, human approach or avoidance and emergence tests (reviewed in Jones, 1996 and Forkman et al., 2007).

Tonic immobility (TI) is an innate, unlearned response to capture and physical restraint, characterized by temporary inhibition of the righting response; i.e. temporary paralysis (Gallup et al., 1974). TI occurs in a wide range of animal species, including mammals, birds and reptiles and has been referred to as death feigning and animal hypnosis (Gallup, 1974; Jones, 1986; Leite-Panissi et al., 2006). According to Jones (1986), TI was "discovered" in 1636 by Daniel Schwenter but TI occurring in a hen has been documented as early as 1562 (see Hoagland, 1928). Tonic immobility is believed to be the last reaction to an attack by a predator and functions to protect the animal by improving its chances of escape (Suarez and Gallup, 1983). Because TI is preceded by fear, TI is used as a test of fear for poultry. Specifically, the duration of TI is used as an indicator of the level of fear of the bird. Longer TI duration is associated with greater levels of fear prior to TI induction (Jones, 1996). General procedures for assessing TI are described in Forkman et al. (2007). Briefly, hens and quail are placed on their side on a table or on their back in a U-shaped cradle. One of the observer's hands is placed on the bird's sternum and the other hand is placed on the head. Chicks are usually restrained on a cloth on a table top, also with one hand on the sternum and one on the chick's head. The bird is held in this manner for 15 s and then released. If the bird rights itself within 10 s the procedure is repeated again, usually up to a maximum of three times. Once TI has been achieved (if the bird remains on its back for 10 s or longer), the amount of time taken before the bird moves its head for the first time, and the duration of TI (amount of time taken before the bird rights itself) are recorded. The number of inductions is also recorded. This methodology has been adapted for turkeys. Turkeys are usually restrained on their side with one hand on the wing and the other hand holding the shanks (Noble et al., 1996).

The open field (OF) test (also known as the novel arena test) was developed by Hall in 1938 to measure "emotionality" in rodents (Boissy, 1995). During the open field test, rodents are exposed to several threatening stimuli, including novelty, bright light, and the absence of landmarks (Archer, 1973). All of these stimuli are believed to induce fear (or anxiety). Rodents that spend more time in the center of the open field are believed to be less anxious (Archer, 1973). The OF test was adapted for use in poultry in the 1980s because it was easy to use, standardize and implement (Boissy, 1995) and is based on similar concepts to the rodent model. However, species differences need to be taken into account when using the OF test. Specifically, poultry are gregarious animals that have "evolved to hide in the undergrowth" and there is little evidence that poultry find bright light aversive, therefore increasing light intensity in the open field as is normally done for rodents is not likely to increase fear in poultry (Forkman et al., 2007). Furthermore, poultry do not display thigmotaxis (staying close to the wall) as rodents do and the time spent in the center of the OF may, therefore, not be a good indicator of a lack of fear in poultry as it is for rodents. Nonetheless, the OF test for poultry is intended to measure fear of novelty, and subjects animals to social isolation (because animals are removed from their flock and tested individually) and is also used to assess social reinstatement motivation (Forkman et al., 2007). It is believed that birds that are inactive and silent experience greater fear than birds that move around the arena and vocalize (Forkman et al., 2007). Procedures for testing birds in an OF test are described in Jones (1996): birds are removed from their home pen or cage and individually tested in an unfamiliar arena that is usually larger and more brightly lit than their home pen for a prescribed period of time. The open field is usually constructed as a square arena that is divided into different areas by constructing a grid on the floor of the arena, but circular arenas have also been used. Measures that are recorded include the latency to ambulate, the

number of steps taken and areas entered, the latency to vocalize and the number of vocalizations emitted (e.g. Clarke and Jones, 2000).

The emergence test (also known as the hole-in-the-wall test) is another test that was originally used to assess emotionality of rodents (Archer, 1973) and similar to the open field test, has been applied for assessing fear in poultry. The emergence test is based on the premise that a bird that is in a sheltered area, such as its home cage, will take longer to emerge into an adjacent unfamiliar area (which could occur through a hole in the cage wall) if the bird is fearful (Jones, 1996; Forkman et al., 2007). The unfamiliar environment is usually more brightly lit than the sheltered area. Forkman et al. (2007) concluded that the emergence test is a version of the OF test, and the OF test is the better test to use.

Similar to the open field test, the novel object (NO) test also assesses the birds' level of fear of novelty (neophobia). An unfamiliar, inanimate object is presented to birds in their home cage, and it is believed that the birds' level of avoidance of the object is indicative of the level of fear. Birds located farther away and taking longer to approach the novel object are interpreted as experiencing a greater level of fear. The novel object test is fast and easy to implement (Forkman et al., 2007), and has been incorporated into welfare assessment programs to evaluate the welfare of poultry (Welfare Quality®, 2009a), pigs (Welfare Quality®, 2009b) and cattle (Welfare Quality®, 2009c).

Up until relatively recently, laying hens were housed in conventional cages where interactions with humans were limited to the interactions occurring at the front of the cage. However, as laying hen housing systems are transitioning from caged systems to alternatives such as aviaries, tests for assessing hens' responses to humans in these larger systems have been developed. Tests include the avoidance distance (approaching person) test, the moving person

test, the stationary person test and the touch test (Raubek et al., 2007). With the avoidance distance (AD) test, the observer approaches a group of birds and determines the distance at which birds withdraw (Raubek et al., 2007), or counts the number of birds within a prescribed distance of the observer (Welfare Quality®, 2009a). With the stationary person (SP) test, the person stands in the litter area with his or her back to a wall and counts the average number of birds within a prescribed distance of the person (Raubek et al., 2007). The moving person (MP) test combines elements of the AD and SP tests so that episodes of moving are interspersed with stationary periods and the number of birds within a prescribed distance of the person when the person is moving and stationary is counted (Raubek et al., 2007). The touch (TT) test is similar to the SP test, except that the person attempts to touch three birds and also counts the number of birds within reach (Raubek et al., 2007). In general, birds that take less time to approach the person or show less avoidance when approached are perceived as being less fearful. The avoidance distance test is used as part of the Welfare Quality® Assessment Protocol for broiler chickens (Welfare Quality®, 2009a).

The aforementioned fear tests and versions thereof are have been used for wild birds (e.g. Carrere et al., 2005; Kralj-Fišer et al., 2010) and other species, such as pigs (e.g. Brown et al., 2009), fish (e.g. Martins et al., 2011; Castanheira et al., 2013) and cattle (e.g. Mazurek et al., 2011) to classify animals as having a particular coping style or temperament.

Reliability and validity of fear tests

Before fear tests can be used to draw conclusions about animal welfare or to classify animals as having a particular temperament, the tests need to be evaluated for reliability and

validity. Martin and Bateson (1993) define a reliable test as one that has few random errors (precise) and is consistent over repeated applications; whereas a valid test is one that is accurate (free from systematic errors), specific (the measurement measures what it is supposed to) and scientifically valid (the measurement provides information relevant to the questions being investigated). In addition to consistency and precision, reliability includes sensitivity (the test measure changes when there are small changes in the true value of the measure) and resolution (the test measure can detect small changes in the true value) (reviewed in Waiblinger et al., 2006). Validity also includes convergent validity (independent test measures that are supposed to measure the same thing are in fact correlated) and discriminant validity (independent test measures that are not supposed to measure the same thing are not related), which form part of the specificity of the test (see Waiblinger et al., 2006). In their review, Waiblinger et al. (2006) further divide scientific validity into internal validity (whether the method used answers the research question) and external validity (whether the results can be applied to situations outside of the experimental environment).

The majority of information regarding the reliability and validity of fear tests for poultry is presented in a review by Forkman et al. (2007) who concluded that TI is a reliable and valid test of fear for hens and quail because previous research has shown TI to be repeatable when birds are tested multiple times, and because there is a quantitative trait locus (QTL) for TI, which suggests that the response is stable within individuals. Furthermore, TI is believed to be a valid test of fear because research has shown that TI reactions are correlated with reactions in other tests of fear: quail selected for long TI duration show increased fear responses to novel objects, novel environments and humans compared to quail selected for short TI durations (reviewed in Jones, 1996). Additionally, TI is influenced by other situations that affect fear responses.

Situations that are believed to cause fear, such as administering electric shocks to the animals or housing animals in poor environments, are associated with increased duration of TI, whereas situations believed to be associated with reduced levels of fear, such as housing animals in enriched environments, are associated with reduced TI duration (reviewed in Forkman et al., 2007).

Similar to the TI test, the OF test is reliable for testing fear in chickens and quail because ambulation in the OF is heritable and because there is a QTL that is related to OF behavior, indicating that the response is stable (reviewed in Forkman et al., 2007; Jensen et al., 2014). The OF test is considered to be valid because behavior in the OF is affected by or associated with other situations that affect fear. For example, electrical shocks induce fear in animals, and animals that are shocked before OF testing vocalize and walk less (see Forkman et al., 2007). Additionally, the OF test is believed to assess fear of novelty, and increasing the degree of novelty present in the arena is associated with reduced activity levels and vocalizations and increased latency to move (see Forkman et al., 2007). Conversely, situations believed to reduce fear, such as housing birds in enriched environments, has been associated with greater activity and shorter vocalization latency (see Forkman et al., 2007).

Although the emergence test has been validated and is correlated with the OF test and TI, there is little information available regarding the reliability of the test, it is similar to the OF test, and therefore the OF test is a better test to use (Forkman et al., 2007).

Forkman and colleagues (2007) concluded that the NO test is reliable for testing fear in laying hens and quail because it is repeatable between days and weeks. Furthermore, birds reared in enriched environments were more likely to approach the NO, indicating that it is a valid test of

fear, and birds housed in the top tier of battery cages (associated with high fear) that have high TI durations also display longer latencies to approach the NO (reviewed in Forkman et al., 2007).

Raubek et al. (2007) assessed the reliability of the AD, SP, MP and TT tests for laying hens. They concluded that the AD and TT tests were the "most promising" tests for assessing hens' reactions to humans because these tests had high repeatability within and between days (Raubek et al., 2007). Graml et al. (2008) assessed the validity of the AD and TT tests and concluded that the tests are valid because repeated handling of birds reduced the distance at which hens moved away from the observer and increased the number of hens that approached the person.

In conclusion, research with hens and quail has demonstrated the reliability and validity of the OF, TI, NO, AD and TT tests for assessing fear in hens and quail. However, there has been no research to date examining the reliability and validity of any of these fear tests for turkeys.

Fear and stress physiology

Emotions, especially fear, are important from an evolutionary perspective because they promote survival. Therefore, fear responses are widely accepted as being "hard-wired" during evolution (Davis, 1992). Fear is a "functional behavior system" that has been regarded as defensive behavior (Misslin, 2003). Some of the most common defensive behaviors occurring in response to predatory attacks and stimuli that induce stress include the "fight or flight" response, freezing, tonic immobility, hypoalgesia (reduced pain sensitivity) and reactions mediated by the autonomic nervous system (e.g. piloerection) (reviewed in Misslin, 2003).

Brain regions involved in fear responses in mammals include the amygdala, hippocampus, thalamus, prefrontal cortex and brain stem structures (reviewed in Shin and

Liberzon, 2010). The amygdala is the most important brain region in mammals involved in the response to fear-inducing stimuli. Through projections to the brainstem, the amygdala causes the expression of fear responses, and projections to the cortex result in the "fear experience" and cognitive processing of emotional stimuli (LeDoux, 2003). In birds, the fear response is mediated by the nuclei teaniae amygdala, posterior pallial amygdala, and the subpallial amygdala, which are homologous to the amygdala in mammals (Reiner et al., 2004).

A stimulus that elicits a fear response can be considered an emotional stressor because it has both an affective component (the subjective experience of fear) and a physiological stress component resulting in activation of the stress-related pathways (Armario et al., 2012). Cannon (1914) is credited with developing the "fight or flight" concept which is concerned with the actions of the sympathetic-adrenal medullary (SAM) axis. Hans Selye is credited with developing the concept of stress, or the General Adaptation Syndrome as it was called, which is concerned with activity of the hypothalamic-pituitary-adrenal (HPA) axis (Selye, 1932). The SAM and HPA axes are the main pathways involved in the stress response. Upon detection of a threat, the SAM and HPA axes are activated. The SAM axis is faster-acting and initiates the immediate responses to a stressor, whereas the HPA axis is slower-acting and is involved in responses that have metabolic and anti-inflammatory effects (reviewed in Mellor et al., 2000). The activation and effects of the SAM and HPA axes have been reviewed in several books and manuscripts (e.g. Matteri et al., 2000; Rodrigues et al., 2009; Amario et al., 2012).

The SAM axis is activated when centers in the brainstem receive signals from the amygdala. Activation of the SAM axis results in the fight/flight response and the release of epinephrine (adrenaline) and norepinephrine (noradrenaline) from sympathetic nerves and the adrenal medulla. Epinephrine and norepinephrine circulate throughout the body and result in

various autonomic effects, including increased blood pressure and heart rate, and divert energy to muscle and inhibit digestion.

In response to signals from the amygdala, the paraventricular nucleus (PVN) of the hypothalamus releases corticotrophin-releasing hormone (CRH) and vasopressin. Corticotrophin-releasing hormone and vasopressin bind to receptors in the anterior pituitary gland, which releases adrenocorticotropin-releasing hormone (ACTH). Adrenocorticotropin-releasing hormone binds to receptors in the cortex of the adrenal gland, causing the secretion of glucocorticoids. Corticosterone is the main glucocorticoid in birds, rodents, amphibians and reptiles; cortisol is the main glucocorticoid in fish and most mammals. Glucocorticoids are then released into the bloodstream and bind to receptors throughout the body and brain. Glucocorticoids result in the release of glucose that is used for energy, and glucocorticoids also have a negative feedback response that reduces the expression of CRH in the PVN of the hypothalamus.

Further, fear processing and activation of the amygdala have monoaminergic, endocrine and autonomic consequences (reviewed in Rodrigues et al., 2009). Monoaminergic consequences include activation of the SAM axis and the release of the neurotransmitters norepinephrine, acetylcholine, serotonin and dopamine throughout the brain, causing increased arousal and vigilance and autonomic effects (e.g. increased heart rate). Endocrine consequences arise as a result of stimulation of the HPA axis, causing the release of CRH and ACTH which lead to the secretion of glucocorticoids (e.g. corticosterone). Physiological changes occurring during fear, therefore, include increased heart rate, change in respiratory rate (breathing), decreased salivation, and increased body temperature. Furthermore, animals may display behavior associated with the fight/flight response, such as freezing or escape and piloerection (erection of

hair) or pteroerection (erection of feathers) and pupil dilation following the detection of a threat and the initiation of fear. Other behavioral indicators of fear include urination, defecation, vocalization, scanning and vigilance, and grooming (Davis, 1992).

Physiology of tonic immobility

Physiologically, the TI response is different from responses to other tests of fear because the TI response has been regarded as a reflex (e.g. Klemm, 1976), and because fear does not cause TI but the duration of TI is related to the degree of fear experienced before TI (reviewed in Jones, 1986).

The TI response is mediated by areas of the brain that are involved in fear and motor inhibition including the amygdala (Donatti and Leite-Panissi, 2009; Leite-Panissi et al., 2003), hypothalamus (Oliveira et al., 1997) and periaqueductal gray matter (PAG) (Monassi et al., 1999) as well as the reticular formation which is responsible for maintaining the waking state (Moruzzi and Magoun, 1949; reviewed in Jones, 2008). Although TI is characterized by temporary motor inhibition, the animal remains conscious (reviewed in Jones, 1986 and Jones 2008).

Gentle et al. (1989) examined electroencephalogram (EEG) activity during TI and reported that the EEG of adult hens was characterized by high amplitude, slow wave activity interspersed with low amplitude, fast wave activity during the period of immobility. Furthermore, hens had reduced neck muscle tone during TI, an elevated heart rate during TI induction which declined during the immobility period, an elevated shank temperature during the immobility period (Gentle et al., 1989. In contrast, Nash et al. (1976) reported that heart rate was lower during TI induction than before induction, and that cloacal temperature was also lower

from the onset of TI to TI termination compared to pre-TI levels. More recent results from a study with rabbits found that heart rate is decreased during TI (Giannico et al., 2014).

Neurotransmitters involved in TI include acetylcholine (Ach) which is involved in freezing and immobility (reviewed in Monassi et al., 1997 and Jones, 2008; Thompson et al., 1974; da Silva and Menescal-de-Oliveira, 2006; Leite-Panissi et al., 2003), gamma-aminobutyric acid (GABA) (reviewed in Jones, 2008; Donatti et al., 2009; Leite-Panissi and Menescal-de-Oliveira, 2002), and serotonin (Hennig, 1980; Wallnau and Gallup, 1977; Dennis et al., 2013; reviewed in Clerici and Veneroni, 2011). Increased levels of serotonin are associated with longer TI durations (e.g. Hill et al., 1994; Dennis et al., 2013), whereas GABA reduces TI duration as indicated by the injection of GABA agonists in the PAG (Monassi et al., 1999; reviewed in Clerici and Veneroni, 2011). Cholinergic neurons stimulate activation of the cerebral cortex during the waking state and during TI (reviewed in Jones, 2008). Furthermore, the cholinergic system mediates analgesia (inability to feel pain) which occurs during TI (da Silva et al., 2006; Leite-Panissi et al., 2003). Hormones that are involved in TI include CRF and ACTH. Activation of CRF in the amygdala of guinea pigs results in increased TI duration (Donatti and Leite-Panissi, 2011) and CRF results in the release of ACTH during TI (Donatti and Leite-Panissi, 2011; Farabollini et al., 1990).

Injurious pecking

The majority of research regarding injurious or damaging pecking in poultry has been conducted with laying hens. The various forms of injurious pecking in laying hens include feather pecking, defined as the pecking, pulling and sometimes removal of feathers of conspecifics; cannibalism, which is the pecking and consumption of skin and tissue; and vent

pecking, the pecking of the area at the top of the cloaca which usually occurs shortly after the onset of lay (Savory, 1995). Injurious pecking in turkeys includes head pecking, feather pecking and cannibalism. Head pecking is more prevalent in males than females (see Dalton et al., 2013), whereas feather pecking occurs in both male and female turkey flocks. Unlike the other forms of injurious pecking, head pecking in turkeys is believed to be related to aggressive pecking that is used to establish or maintain the dominance hierarchy (reviewed in Dalton et al., 2013), whereas feather pecking and cannibalism have similar manifestations as in laying hens, and are not related to aggression.

In order to control and reduce injurious pecking, laying hens and turkeys are beak-trimmed within a few days after hatch. With beak trimming, the tip of the beak is removed using a hot blade or through infra-red beak trimming, which are the two main methods of beak trimming (reviewed in Nicol et al., 2013). In addition to beak trimming, injurious pecking is also controlled by reducing the light intensity in barns. However, both methods of controlling injurious pecking adversely affect well-being. Beak trimming is associated with pain (e.g. Freire et al., 2008; Jongman et al., 2008; reviewed in Nicol et al., 2013) and reduced light intensity is associated with altered eye development and reduced activity levels (e.g. Nickla et al., 2001; reviewed in Nicol et al., 2013).

Feather pecking is the best-studied form of injurious pecking in poultry. Two forms of feather pecking are recognized, including gentle and severe feather pecking. Gentle feather pecking causes little damage and may be directed at food or dust particles, seldom eliciting a reaction from the recipient (Savory, 1995). In contrast, severe feather pecking occurs when a bird forcefully pecks at and pulls feathers from another bird, resulting in feather damage and feather loss, and causing the recipient to react or move away (Savory, 1995). Initially, birds that are

recipients of feather pecking may respond "by moving away or by confronting the pecker", but birds may stop reacting to being pecked as pecking continues (Rodenburg et al., 2013). Severe feather pecking may lead to skin and tissue damage and cannibalism, which in turn may lead to mortality (Savory, 1995; reviewed in Rodenburg et al., 2013). Feather pecking is recorded in one of two ways, either by recording plumage damage (indirectly), or by directly observing bird behavior to identify individuals performing feather pecking. Plumage scores are often used because plumage damage is easier to examine than feather pecking behavior (Nicol et al., 2013).

Feather pecking of laying hens

Feather pecking occurs widely among egg laying hen flocks, and feather pecking and cannibalism are the major causes of mortality in laying hen flocks housed in non-cage systems (Nicol et al., 2013). Gentle feather pecking is observed as early as one day of age and remains stable or decreases in prevalence after the rearing period, whereas severe feather pecking is observed during the laying period (after 18 weeks of age) and usually increases throughout lay (Nicol et al., 1999; Pötzsch et al., 2001; Lambton et al., 2010; reviewed in Nicol et al., 2013). Because gentle feather pecking is not associated with much plumage damage (Rodenburg et al., 2013), most research has been aimed at understanding the causes and development of severe feather pecking.

In a recent review, Rodenburg et al. (2013) concluded that severe feather pecking is related to foraging and feeding behavior, in line with the redirected foraging and ground pecking hypothesis proposed by Blokhuis (1989). Indeed, the importance of access to litter as a foraging substrate in reducing feather pecking has been confirmed by several studies (reviewed in Rodenburg et al., 2013). Severe feather pecking may also be related to dietary factors such as the

presence of loose feathers in the litter, which leads to feather eating and further feather pecking; or to the amount of fiber, magnesium, protein and amino acids (lysine, cystine and methionine) present in the laying hen diet (Al Bustany and Elwinger, 1987a; 1987b; Elwinger et al., 2002; reviewed in Rodenburg et al., 2013). The maternal environment may also influence feather pecking behavior in the offspring through the deposition of maternal hormones in the eggs, which affect development of the brain and behavior of the chick (Freire et al., 2006; de Haas et al., 2014; reviewed in Rodenburg et al., 2013). Indeed, high maternal levels of corticosterone and whole-blood serotonin are associated with increased levels of severe feather pecking in chicks (de Haas et al., 2014). Lastly, individual and genetic factors play important roles in the development of feather pecking behavior. Feather pecking is known to differ among genetic strains of hens; indeed, most research regarding laying hen feather pecking is conducted with two strains that have been divergently selected for high (HFP) and low (LFP) feather pecking, respectively (Kjaer and Sørensen, 1997; Kjaer et al., 2001).

Several studies have confirmed a link between feather pecking and temperament characteristics; specifically, fearfulness and the hen's ability to cope with stress (e.g. Blokhuis and Beutler, 1992; Jones et al., 1995; Rodenburg et al., 2004; reviewed in Rodenburg et al., 2013). In general, the performance of feather pecking and the severity of feather damage are associated with increased levels of fearfulness in various genetic lines of chickens, including in chickens from high and low feather pecking lines (e.g. Jones et al., 1995; Rodenburg et al., 2010; reviewed in Rodenburg et al., 2013). Furthermore, genetic lines differing in feather pecking behavior also differ in physiological responses to stress (hens from the HFP line appear to be more susceptible to stress), and in serotinergic and dopaminergic neurotransmission in the brain (van Hierden et al., 2002; 2004; 2005; reviewed in Rodenburg et al., 2013). Indeed, hens from

the high feather pecking line have lower serotonin turnover vs. hens from the lower feather pecking line (van Hierden et al., 2002), and dopamine turnover is lower in chicks from high vs. low feather pecking lines (van Hierden et al., 2002; 2005; reviewed in Rodenburg et al., 2013).

In conclusion, feather pecking is a multifactorial problem related to early experience, feeding and foraging behavior, maternal effects, genetics, and the ability to cope with fear and stress (Rodenburg et al., 2013). Some other factors related to environment and management that are associated with increased risk of feather pecking in commercial facilities include not providing access to perches, increased stocking density and group size, more than one or two dietary changes during the rearing period and more than three dietary changes during the laying period (reviewed in Nicol et al., 2013).

Feather pecking of turkeys

In contrast to chickens, very little is known about the development and causes of feather pecking of turkeys, but feather pecking has long been recognized as an important welfare problem (Hocking, 1993; Marchewka et al., 2013). Feather pecking of turkeys occurs as early as 5 days of age (Moinard et al., 2001) and has been reported to account for up to 58% of culls and mortalities in a commercial facility (Duggan et al., 2014).

Few studies have systematically examined feather pecking of turkeys, and only two studies have been conducted since 2006. Most recently, a study examining injurious pecking among turkeys in a commercial facility (Duggan et al., 2014) demonstrated that the housing environment significantly affects the prevalence of injurious pecking. Indeed, mortality and culling was higher and feather condition was worse in turkeys housed in curtain-sided barns compared to turkeys housed in mechanically ventilated barns where the light intensity was lower

(Duggan et al., 2014). Results from another study revealed that feather pecking is affected by genetic strain and sex (Busayi et al., 2006). Specifically, the frequency of gentle and "strong" feather pecking was higher among randombred turkeys than among turkeys of a commercial strain, but damaging pecking was more severe among commercial turkeys. Furthermore, there were age-dependent sex differences in feather pecking, and more injuries in males than females (Busayi et al., 2006). Older studies examining feather pecking among turkeys have found that other factors such as environmental enrichment, light intensity, light source (incandescent vs. fluorescent) and diet affect feather pecking (reviewed in Dalton et al., 2013).

As with chickens, it has been concluded that injurious pecking of turkeys is consistent with the redirected foraging hypothesis of feather pecking of chickens, and is a multifactorial problem that is associated with genetic, environmental and nutritional factors (Dalton et al., 2013). There is no published literature examining the relationship between fearfulness, stress reactivity and feather pecking of turkeys.

Stress, temperament and meat quality

Meat quality concerns the attributes of meat that influence consumer acceptance, including appearance, tenderness, juiciness, flavor (Wilson, 1960) and functionality (Fletcher, 2002). The most important attributes determining meat quality are appearance which includes color and defects, and texture which is affected by processing procedures and deboning (Fletcher, 2002). The major factors affecting poultry meat color include heme pigments, such as myoglobin and haemoglobin; pre-slaughter factors, such as genetic strain, stress before slaughter, and feed; and the slaughter, chilling and processing procedures (reviewed in Fletcher, 2002). Color is also affected by pH such that meat with a high pH is darker in color whereas a lower pH

is associated with a lighter color (Fletcher, 2002). Color is measured using one of three methods, including visual assessment, spectrophotometry and reflectance calorimetry (Castigliego et al., 2012). The CIE (International Commission on Illumination) is the most common calorimetric method and is used to determine L* which refers to lightness, a* which refers to redness or greenness, and b* which refers to yellowness or blueness (Castigliego et al., 2012). Texture, on the other hand, is determined by the "contractile state of the myofibrillar proteins" which is a factor of the rate at which rigor mortis develops (Fletcher, 2002). Rigor mortis development affects the ultimate pH and color (Castigliego et al., 2012) and is indirectly measured using the R-value which is an indicator of the level of adenosine-triphosphate (ATP) breakdown (Calkins et al., 1982).

Pre-slaughter stress and meat quality

Two conditions associated with stress and meat quality are pale, soft, exudative (PSE) meat in pigs and dark, firm, dry (DFD) meat in pigs and cattle. PSE meat in pigs results when pigs experience stress before slaughter, which results in rapid glycolysis, increased lactic acid production and a decrease in muscle pH (reviewed in Broom and Johnson, 1993) and accelerated rigor mortis (Sosnicki et al., 1998). These factors result in water loss and pale muscle color. PSE-like meat has also been documented in chickens and turkeys (Fletcher, 2002), and has increased over the years in turkeys (Sosnicki et al., 1998) as turkeys have been selectively bred for higher growth rates and breast muscle yield (Strasburg and Chaing, 2009).

In contrast to PSE meat, DFD meat results when there is little or no glycogen in the muscle before slaughter, resulting in low lactic acid production and high pH (see Broom and Johnson, 1993). High levels of physical activity before slaughter result in DFD meat because

glycogen reserves are depleted (see Terlouw, 2005). DFD meat has been extensively studied in cattle, but DFD-like meat has also been reported in chickens (Fletcher, 2002).

Temperament and meat quality

The relationship between temperament and meat quality is based on the premise that animals that differ in temperament also differ in stress reactivity, and therefore differ in meat quality because conditions at slaughter are associated with a number of stressful events, including catching and crating, transportation, separation from conspecifics and unfamiliarity. Numerous studies using different breeds of cattle have confirmed that cattle with poor temperaments (more fearful, flightier animals) have higher stress-susceptibility to procedures before and at slaughter (e.g. Petherick et al., 2002) and produce meat with a higher pH (e.g. Cafe et al., 2011; Ribeiro et al., 2012), and darker color (e.g. Cafe et al., 2011; Ribeiro et al., 2012; Voisinet et al., 1997). Similarly, research with pigs has found that temperament characteristics such as aggression are correlated with pigs' responses to a novel object, and pigs that are more aggressive and that show less fear of a novel object are more reactive to pre-slaughter procedures, resulting in increased glycolysis and reduced pH (Terlouw, 2005). In contrast, pigs that approach and touch a human more often during a "human exposure test" administered during the rearing period exhibit reduced stress reactivity to pre-slaughter procedures, resulting in lower muscle metabolic activity and reduced glycolysis (Terlouw, 2005).

Very few studies have been conducted to assess whether meat quality is related to temperament in poultry, and there are no such studies with turkeys. Remingnon et al. (1998) assessed meat quality (pH and drip loss) in the LTI and STI lines of quail that had been subjected to physical restraint before slaughter. Their results demonstrated that pH and drip loss were

higher in quail from the LTI line. In contrast, Debut et al. (2003) found no relationship between meat quality and TI of quail assessed one week before slaughter.

CHAPTER OVERVIEW

The chapters of this dissertation examine fearfulness, feather pecking and meat quality of turkeys and aim to determine whether fearfulness is associated with feather pecking and meat quality. Chapter 2 examines the behavior of male commercial turkeys in four commonly used tests of fear. The reliability of each fear test is also evaluated as well as whether behavior changes over time. In order to examine test reliability, turkeys were repeatedly tested in each fear test. Chapter 3 addresses the validity of the four fear tests used in Chapter 2. A strategy for assessing test validity is to determine whether animals behave in the same way across different test situations; in other words, whether animals that are classified as being more fearful in one test are also more fearful in other tests relative to other animals. Therefore, test validity was assessed using inter-test correspondence. Chapter 4 examines feather pecking behavior in male and female turkeys of a commercial and randombred strain, and examines whether feather pecking behavior is associated with fearfulness in an open field test. Open field test reliability for males and females of both strains is examined and discussed in relation to the results presented in Chapter 2. Lastly, Chapter 5 examines differences in meat quality between commercial and randombred turkey males, and examines whether meat quality is associated with differences in open field behavior.

REFERENCES

REFERENCES

Al Bustany, Z., and K. Elwinger. 1987a. Comparison between barley/fish meal- and maize/soybean meal-based diets with various lysine and protein levels fed to different strains of laying hens. Acta Agric. Scand. 37: 41-49.

Al Bustany, Z., and K. Elwinger. 1987b . Response of laying hens to different dietary lysine intakes. A comparison of some commercial hybrids with strains selected on a low protein diet. Acta Agric. Scand. 37: 27-40.

Appleby, M. C., J. A. Mench, and B. O. Hughes, eds. 2004. Origins. Pages 5-8 in: Poultry Behaviour and Welfare. CABI Publishing, Cambridge, MA, USA.

Archer, J. 1973. Tests for emotionality in rats and mice - review. Anim. Behav. 21: 205-235.

Armario, A., N. Daviu, C. Munoz-Abellan, C. Rabasa, S. Fuentes, X. Belda, H. Gagliano, and R. Nadal. 2012. What can We Know from Pituitary-Adrenal Hormones About the Nature and Consequences of Exposure to Emotional Stressors? Cell. Mol. Neurobiol. 32: 749-758.

Blokhuis, H.J. 1989. The development and causation of feather pecking in the domestic fowl. Vakgroep-Ethologie, Thesis Landbouw Universiteit Wageningen, Wageningen. http://edepot.wur.nl/202211. Accessed Jun. 2014.

Blokhuis, H.J. and A. Beutler. 1992. Feather pecking damage and tonic immobility response in two lines of white leghorn hens. J. of Anim. Sci. 70: 170.

Boissy, A. Fear and fearfulness in animals. 1995. Q. Rev. Biol. 70: 165-191.

Bourin, M., B. Petit-Demoulière, B. N. Dhonnchadha, and M. Hascöet. 2007. Animal models of anxiety in mice. Fundam. Clin. Pharm. 21: 567–574.

Brambell, F. W. R. 1965. Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. (Command Rep. 2836). London: Her Majesty's Stationery Office.

Broom, D. M. 1986. Indicators of poor welfare. Br. Vet. J. 142: 524-526.

Broom, D. M., and K. G. Johnson. 1993. Chapter 5 Assessing animal welfare: short term responses. Page 106 in Stress and animal welfare. Broom, D. M., and K. G. Johnson, eds. Kluwer Academic Publishers, Cambridge, MA, USA.

Brown, J. A., C. Dewey, C. F. M. Delange, I. B. Mandell, P. P. Purslow, J. A. Robinson, E. J. Squires, and T. M. Widowski. 2009. Reliability of temperament tests on finishing pigs in group-housing and comparison to social tests. Appl. Anim. Behav. Sci. 118: 28-35.

Busayi, R. M., C. E. Channing, and P. M. Hocking. 2006. Comparisons of damaging feather pecking and time budgets in male and female turkeys of a traditional breed and a genetically selected male line. Appl. Anim. Behav. Sci. 96: 281-292.

Buss, E. G. 1989. Genetics of turkeys: Origin and development. World. Poult. Sci. J. 45:27-52.

Cafe, L. M., D. L. Robinson, D. M. Ferguson, G. H. Geesink, and P. L. Greenwood. 2011. Temperament and hypothalamic-pituitary-adrenal axis function are related and combine to affect growth, efficiency, carcass, and meat quality traits in Brahman steers. Dom. Anim. Endocrinol. 40: 230-240.

Calkins, C. R., T. R. Dutson, G. C. Smith, and Z. L. Carpenter. 1982. Concentration of creatine-phosphate, adenine-nucleotides and their derivatives in electrically stimulated and non-stimulated beef muscle. J. Food Sci. 47: 1350-1353.

Cannon, W.B. 1914. The emergency function of the adrenal medulla in pain and the major emotions. Am. J. Physiol. 33: 356– 372.

Carcia, R. V. and S. Harvey. 2000. Chapter 19 Adrenals. Pages 500-501 in Sturkie's Avian Physiology, 5th rev. ed. Whittow, G. C., ed. Academic Press, New York, NY, USA.

Carere, C., D. Caramaschi, and T. W. Fawcett. 2010. Covariation between personalities and individual differences in coping with stress: Converging evidence and hypotheses. Curr. Zool. 56: 728-740.

Castanheira, M. F., M. Herrera, B. Costas, L. E. C. Conceicao, and C. I. M. Martins. 2013. Can We Predict Personality in Fish? Searching for Consistency over Time and across Contexts. PLoS One 8: 9.

Castigliego, L., A. Armani, and A. Guidi. 2012. Meat color. Pages 82-90 in Handbook of Meat and Meat Processing, 2nd rev. ed. Hui, Y. H., ed. CRC Press, New York, NY, USA.

Clarke, C. H., and R. B. Jones. 2000. Effects of prior video stimulation on open-field behaviour in domestic chicks. Appl. Anim. Behav. Sci. 66: 107-117.

Clerici, C. A, and L. Veneroni (eds.) 2011. Chapter I. Pages 13-25 in: The impossible escape: studies on the tonic immobility in animals from a comparative psychology perspective (Animal science, issues and professions). Nova Science Publishers Inc. Hauppage, NY, USA. Available online: http://www.scribd.com/doc/103780538/The-impossible-escape-studies-on-the-tonic-immobility-in-animals-from-a-comparative-psychology-perspective

Cockrem, J. F. 2007. Stress, corticosterone responses and avian personalities. J Ornithol. 148: S169–S178.

Cockrem, J. F. 2013. Corticosterone responses and personality in birds: Individual variation and the ability to cope with environmental changes due to climate change. Gen. Comp. Endocr. 190: 156-163.

Coppens, C. M., S. F. de Boer, and J. M. Koolhaas. 2010. Coping styles and behavioural flexibility: towards underlying mechanisms. Philos. T. Roy. Soc. B. 365: 4021-4028.

Crawford, R. D. 1992. Introduction to Europe and diffusion of domesticated turkeys from the Americas. Arch. Zootec. 41: 307-314.

Crowe, R., and J. M. Forbes. 1999. Effects of four different environmental enrichment treatments on pecking behaviour in turkeys. Br. Poult. Sci. 40: S1, 11-12.

Dalton, H. A., B. J. Wood, and S. Torrey. 2013. Injurious pecking in domestic turkeys: development, causes and potential solutions. World's Poult. Sci. J. 69: 865-875.

da Silva, L. F. S., and L. Menescal-de-Oliveira. 2006. Cholinergic modulation of tonic immobility and nociception in the NRM of guinea pig. Physiol. & Behav. 87: 821-827.

Davis, M. 1992. The role of the amygdala in fear and anxiety. Ann. Rev. Neurosci. 15: 353-375.

Debut, M., C. Berri, E. Baeza, N. Sellier, C. Arnould, D. Guemene, N. Jehl, B. Boutten, Y. Jego, C. Beaumont, and E. Le Bihan-Duval. 2003. Variation of chicken technological meat quality in relation to genotype and preslaughter stress conditions. Poult. Sci. 82: 1829–1838.

de Haas, E. N., J. E. Bolhuis, B. Kemp, T. G. G. Groothuis, and T. B. Rodenburg. 2014. Parents and Early Life Environment Affect Behavioral Development of Laying Hen Chickens. PLoS One 9: 12.

Dennis, R. L., A. G. Fahey, and H. W. Cheng. 2013. Alterations to embryonic serotonin change aggression and fearfulness. Aggressive Behav. 31: 91-98.

Donatti A. F., and C. R. A. Leite-Panissi. 2009. GABAergic antagonist blocks the reduction of tonic immobility behavior induced by activation of 5-HT2 receptors in the basolateral nucleus of the amygdala in guinea pigs. Brain. Res. Bull. 79: 358–64.

Donatti, A. F., and C. R. A. Leite-Panissi. 2011. Activation of corticotropin-releasing factor receptors from the basolateral or central amygdala increases the tonic immobility response in guinea pigs: An innate fear behavior. Biobehav. Brain Res. 225: 23-30.

Duggan, G., T. Widowski, M. Quinton, and S. Torrey. The development of injurious pecking in a commercial turkey facility. 2014. J. Appl. Poult. Res. 23: 1-11.

Elwinger, K., R. Tauson, M. Tufvesson, and C. Hartmann. 2002. Feeding of layers kept in an organic feed environment. 11th. European Poultry Conference, Bremen.

Farabollini F., F. Facchinetti, C. Lupo, and G. Carli. 1990. Time course of opioid and pituitary–adrenal hormone modification during the immobility reaction in rabbits. Physiol. Behav. 47: 337–41.

Farm Animal Welfare Council. 1979. Press Statement. http://www.fawc.org.uk/pdf/fivefreedoms1979.pdf. Accessed Jun. 2014.

Fletcher, D. L., M. Qiao, and D. P. Smith. 2000. The relationship of raw broiler breast meat color and pH to cooked meat color and pH. Poult. Sci. 79:784–788.

Forkman, B., A. Boissy, M. –C. Maunier-Salaün, E. Canali, and R. B. Jones, R. B. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol. Behav. 92: 340-374.

Freire, R., P. C. Glatz, and G. Hinch. 2008. Self-administration of an analgesic does not alleviate pain in beak trimmed chickens. Asian-Aust. J. Anim. Sci. 21: 443-448.

Freire, R., S. van Dort, and L. J. Rogers. 2006. Pre- and post-hatching effects of corticosterone treatment on behavior of the domestic chick. Horm. Behav. 49: 157-165.

Gallup, G. G. 1974. Animal hypnosis - factual status of a fictional concept. Psychol. Bull. 81: 836-853.

Gentle, M. J., R. B. Jones, and S. C. Woolley. 1989. Physiological changes during tonic immobility in *Gallus gallus* var *domesticus.* Physiol. & Behav. 46: 843-847.

Giannico, A. T., L. Lima, R. R. Lange, T. R. Froes, and F. Montiani-Ferreira. 2014. Proven cardiac changes during death‐feigning (tonic immobility) in rabbits (*Oryctolagus cuniculus*). J. Comp. Physiol. A. 200:305–310.

Gosling, S. D. 2001. From mice to men: what can we learn about personality from animal research? Psychol. Bull. 127: 45-86.

Graml., C., S. Waiblinger, and K. Niebuhr. 2008. Validation of tests for on-farm assessment of the hen-human relationship in non-cage systems. Appl. Anim. Behav. Sci. 111: 301-310.

Gray, J. A. 1979. Emotionality in male and female rodents. A reply to Archer. Br. J. Psychol. 70: 425-440.

Groothuis, T. G. G., and C. Carere. 2005. Avian personalities: characterization and epigenesis. Neurosci. Biobehav. Rev. 29: 137-150.

Hennig, C. W., J. F. McIntyre, D. D. Moriarty, J. M. Picerno, and J. L. Allen. 1988. Differential cholinergic influences on the immobility response in various strains of domestic fowl. Pharmacol. Biochem. Behav. 30: 625-634.

Hill, W. L., T. M. Fleming, and E. M. Shrier. 1994. Tonic immobility and high-intensity calls in a precocial chick as a function of age, diet and time of day. Dev. Psychobiol. 27: 331-342.

Hoagland, H. 1928. On the mechanism of tonic immobility in vertebrates. J. Gen. Phys. 11: 715-741.

Hocking, P. 1993. Welfare of turkeys. Pages 125-138 in Proceedings of the Fourth European Symposium on Poultry Welfare. Savory C J and B. O. Hughes, eds. Universities Federation for Animal Welfare, Potters Bar, UK.

Hughes, B. O., and P. N. Grigor. 1996. Behavioural time-budgets and beak related behaviour in floor-housed turkeys. Anim. Welf. 5:189-198.

Jacobs, D. S. 2009. Animal personality and biological markets: rise of the individual. Afr. Zool. 44: 271-282.

Jensen, P., M. Johnsson, and D. Wright. 2014. Domestication related anxiety genes identified in chickens. 48th Congress for the International Society for Applied Ethology. Vitoria-Gasteiz, Spain. Pp. 134 (Abstr.).

Jones, B. E. 2008. Modulation of cortical activation and behavioral arousal by cholinergic and orexinergic systems. Ann. N.Y. Acad. Sci. 1129: 26–34.

Jones, R. B. 1986. The tonic immobility reaction of the domestic-fowl - a review. Worlds Poult. Sci. J. 42: 82-96.

Jones, R. B. 1996. Fear and adaptability in poultry: insights, implications and imperatives. World's Poultry Science Journal. 52: 131-174.

Jones, B., and A. Boissy. 2011. Fear and other negative emotions. Page 81 in Animal Welfare. 2nd rev. ed. Appleby, M. C, J. A. Mench, I. A. S. Olsson, and B. O. Hughes, eds. CABI International, Cambridge, MA, USA.

Jongman, E.C., P. C. Glatz, and J. L. Barnett. 2008. Changes in behaviour of laying hens following beak trimming at hatch and re-trimming at 14 weeks. Asian-Aust. J. Anim. Sci. 21: 291-298.

Kjaer, J.B., and P. Sørensen. 1997. Feather pecking behaviour in White Leghorns, a genetic study. Br. Poult. Sci. 38: 333–341.

Kjaer, J.B., P. Sørensen, and G. Su. 2001. Divergent selection on feather pecking behaviour in laying hens (Gallus gallus domesticus). Appl. Anim. Behav. Sci. 71: 229–239.

Klemm, W. R. 1976. Identity of sensory and motor systems that are critical to the immobility reflex ("animal hypnosis"). J. Neurosci. Res. 2: 57-69.

Koolhaas, J. M., S. F. de Boer, C. M. Coppens, and B. Buwalda. 2010. Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. Front. Neuroendocrinol. 31: 307-321.

Koolhaas, J. M., S. M. Korte, S. F. De Boer, B. J. Van Der Vegt, C. G. Van Reenen, H. Hopster, I. C. De Jong, M. A. W. Ruis, and H. J. Blokhuis. 1999. Coping styles in animals: current status in behavior and stress-physiology. Neurosci. Biobehav. Rev. 23: 925-935.

Kralj-Fiser, S., B. M. Weiss, and K. Kotrschal. 2010. Behavioural and physiological correlates of personality in greylag geese (Anser anser). J. Ethol. 28: 363-370.

Lambton, S.L., T. G. Knowles, C. Yorke, and C. J. Nicol. 2010. The risk factors affecting the development of gentle and severe feather pecking in loose housed laying hens. Appl. Anim. Behav. Sci. 123: 32-42.

LeDoux, J. 2003. The Emotional Brain, Fear, and the Amygdala. Cellular and Molecular Neurobiology. 23: 727-738.

Leite-Panissi, C. R. A., N. C. Coimbra, and L. Menescal-de-Oliveira. 2003. The cholinergic stimulation of the central amygdala modifying the tonic immobility response and antinociception in guinea pigs depends on the ventrolateral periaqueductal gray. Brain Res. Bull. 60: 167–178

Leite-Panissi, C. R. A., A. A. Ferrarese, A. L. B. Terzian, and L. Menescal-de-Oliveira. 2006. Serotoninergic activation of the basolateral amygdala and modulation of tonic immobility in guinea pig. Brain Res. Bull. 69: 356-364.

Leite Panissi C. R. A., and L. Menescal-de-Oliveira. 2002. Central nucleus of the amygdala and the control of tonic immobility in guinea pigs. Brain Res. Bull. 58: 13–9.

Lewis, P. D., G. C. Perry, and C. M. Sherwin. 1998. Effect of photoperiod and light intensity on the performance of intact male turkeys. Anim. Sci. 66: 759-767.

Marchewka, J., T. T. N. Watanabe, V. Ferrante, and I. Estevez. 2013. Review of the social and environmental factors affecting the behavior and welfare of turkeys (Meleagris gallopavo). Poult. Sci. 92:1467-1473.

Martrenchar, A., D. Huonnic, and J. P. Cotte. 2001. Influence of environmental enrichment on injurious pecking and perching behaviour in young turkeys. Br. Poult. Sci. 42: 161-170.

Martin, P., and P. Bateson, eds. 1993. The reliability and validity of measures. Pages 114-116 in Measuring behaviour: an introductory guide. 2nd rev. ed. Cambridge University Press, New York, NY, USA.

Martins, C. I. M., P. I. M. Silva, L. E. C. Conceicao, B. Costas, E. Hoglund, O. Overli, and J. W. Schrama. 2011. Linking Fearfulness and Coping Styles in Fish. PLoS One 6: 8.

Matteri, R. L., J. A. Carroll, and C. J. Dyer. 2000. Neuroendocrine responses to stress. Pages 43-76 in Biology of Animal stress - Basic Principles and Implications for Animal Welfare. Moberg, G. P., and J. A. Mench, J. A., eds. CABI Publishing, New York, NY, USA.

Mazurek, M., M. McGee, M. A. Crowe, D. J. Prendiville, X. Boivin, and B. Earley. 2011. Consistency and stability of behavioural fear responses of heifers to different fear-eliciting situations involving humans. Appl. Anim. Behav. Sci. 131: 21-28.

Mellor, D. J., C. J. Cook, and K. J. Stafford. 2000. Quantifying some responses to pain as a stressor. Page 174 in Biology of Animal stress - Basic Principles and Implications for Animal Welfare. Moberg, G. P., and J. A. Mench, J. A., eds. CABI Publishing, New York, NY, USA.

Moinard, C., P. D. Lewis, G. C. Perry, and C. M. Sherwin. 2001. The effects of light intensity and light source on injuries due to pecking of male domestic turkeys (Meleagris gallopavo). Anim. Welfare 10: 131-139.

Monassi, C. R., A. Hoffman, and L. Menescal-de-Oliveira. 1997. Involvement of the cholinergic system and periaqueductal gray matter in the modulation of tonic immobility in the guinea pig. Physiol. & Behav. 62: 53-59.

Monassi, C. R., C. R. Leite-Panissi, and L. Menescal-de-Oliveira. 1999. Ventrolateralperiaqueductal gray matter and the control of tonic immobility. Brain Res. Bull. 50: 201-8.

Moruzzi, G. and H. W. Magoun. 1949. Brain stem reticular formation and activation of the EEG. Electroencephalogr. Clin. Neurophysiol. 1: 455–473.

Nash, R. F., G. G. Gallup, and D. A. Czech1976. Pshychophysiological correlates of tonic immobility in the domestic chicken (*Gallus gallus*). Physiol. & Behav. 17: 413-418.

Neumann, I. D., G. Wegener, J. R. Homberg, H. Cohen, D. A. Slattery, J. Zohar, J. D. A. Olivier, and A. A. Mathe. 2011. Animal models of depression and anxiety: What do they tell us about human condition? Prog. Neuro-Psychoph. 35: 1357-1375.

Nickla, D.L., C. F. Wildsoet, and D. Troilo. 2001. Endogenous rhythms in axial length and choroidal thickness in chicks: Implications for ocular growth regulation. Invest. Ophthalmol. Vis. Sci. 42: 584-588.

Nicol, C.J., N. G. Gregory, T. G. Knowles, I. D. Parkman, and L. J. Wilkins. 1999. Differential effects of increased stocking density, mediated by increased flock size, on feather pecking and aggression in laying hens. Appl. Anim. Behav. Sci. 65: 137-152.

Nicol, C. J., M. Bestman, A. M. Gilani, E. N. De Haas, I. C. De Jong, S. Lambton, J. P. Wagenaar, C. A. Weeks, and T. B. Rodenburg. 2013. The prevention and control of feather pecking: application to commercial systems. Worlds Poult. Sci. J. 69: 775-788.

Noble, D. O., J. W. Anderson, and K. E. Nestor. 1996. Range and confinement rearing of four genetic lines of turkeys. 2. Effects on behavior and tonic immobility. Poult. Sci. 75: 165-171.

Oliveira L, A. Hoffmann, and L. Menescal-de-Oliveira. 1997. Participation of the medial and anterior hypothalamus in the modulation of tonic immobility in guinea pigs. Physiol. Behav. 62: 1171–8.

Owens, C. M., and A. R. Sams. 2000. The influence of transportation on turkey meat quality. Poult. Sci. 79: 1204-1207.

Petherick, J. C., R. G. Holroyd, V. J. Doogan, and B. K. Venus. 2002. Productivity, carcass and meat quality of lot-fed Bos indicus cross steers grouped according to temperament. Aust. J. Exp. Agr. 42: 389-398.

Pötzsch, C.J., K. Lewis, C. J. Nicol, and L. E. Green. 2001. A cross-sectional study of the prevalence of vent pecking in laying hens in alternative systems and its associations with feather pecking, management and disease. Appl. Anim. Behav. Sci. 74: 259-272.

Raubek, J., K. Niebuhr, and S. Waiblinger. 2007. Development of on-farm methods to assess the animal-human relationship in laying hens kept in non-cage systems. Anim. Welfare. 16: 173-175.

Reiner, A., D. J. Perkel, L. L. Bruce, A. B. Butler, A. Csillag, W. Kuenzel, L. Medina, G. Paxinos, T. Shimizu, G. Striedter, M. Wild, G. F. Ball, S. Durand, O. Gunturkun, D. W. Lee, C. V. Mello, A. Powers, S. A. White, G. Hough, L. Kubikova, T. V. Smulders, K. Wada, J. Dugas-Ford, S. Husband, K. Yamamoto, J. Yu, C. Siang, and E. D. Jarvis. 2004. Revised nomenclature for avian telencephalon and some related brainstem nuclei. J. Comp. Neurol. 473: 377-414.

Remignon, H., A. D. Mills, D. Guemene, V. Desrosiers, M. Garreau-Mills, M. Marche, and G. Marche. 1998. Meat quality traits and muscle characteristicsin high or low fear lines of Japanese quails (*Coturnix japonica*) subjected to acute stress. Br. Poult. Sci. 39: 372-378.

Ribeiro, J. d. S., T. d. M. Goncalves, M. M. Ladeira, R. R. Tullio, F. R. Campos, J. A. Garcia Bergmann, O. R. Machado Neto, and J. R. Reis de Carvalho. 2012. Reactivity, performance, color and tenderness of meat from Zebu cattle finished in feedlot. Rev. Bras. Zootecn. 41: 1009-1015.

Rodenburg, T.B., A. J. Buitenhuis, B. Ask, K. A. Uitdehaag, P. Koene, J. J. Van Der Poel, J. A. M. Van Arendonk, and H. Bovenhuis. 2004. Genetic and phenotypic correlations

between feather pecking and open-field response in laying hens at two different ages. Behav. Genet. 34: 407-415.

Rodenburg, T.B., E. N. de Haas, B. L. Nielsen, and A. J. Buitenhuis. 2010. Fearfulness and feather damage in laying hens divergently selected for high and low feather pecking. Appl. Anim. Behav. Sci. 128: 91-96.

Rodenburg, T. B., M. M. van Krimpen, I. C. de Jong, E. N. de Haas, M. S. Kops, B. J. Riedstra, R. E. Nordquist, J. P. Wagenaar, M. Bestman, and C. J. Nicol. 2013. The prevention and control of feather pecking in laying hens: identifying the underlying principles. Worlds Poult. Sci. J. 69: 361-373.

Rodrigues, S. M., J. E. LeDoux, and R. M. Sapolsky. 2009. The influence of stress hormones on fear circuitry. Annu. Rev. Neurosci. 32: 289–313.

Sarica, M., N. Ocak, S. Turhan, C. Kop, and U. S. Yamak. 2011. Evaluation of meat quality from 3 turkey genotypes reared with or without outdoor access. Poult. Sci. 90: 1313-1323.

Savory, C. J. 1995. Feather pecking and cannibalism. Worlds Poult. Sci. J. 51: 215-219.

Selye, H., 1932. The general adaptation syndrome and the diseases of adaptation. J. Clin. Endocrin. 6: 117–152.

Sherwin, C. M., P. D. Lewis, and G. C. Perry. 1999. Effects of environmental enrichment, fluorescent and intermittent lighting on injurious pecking amongst male turkey poults. British Br. Poult. Sci. 40:592-598.

Shin, L. M., and I. Liberzon. 2010. The neurocircuitry of fear, stress and anxiety disorders. Neuropsychopharm. Rev. 35: 169-191.

Sih, A., A. Bell, and J. C. Johnson. 2004. Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol. Evol. 19: 372-378.

Sosnicki, A. A., M. L Greaser, M. Pietrzak, E. Pospiech, and V. Sante.1998. PSE-like syndrome in breast muscle of domestic turkeys: a review. J. Muscle Foods. 9: 13-23.

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

Strasburg, G. M. and W. Chiang. 2009. Pale, soft, exudative turkey-the role of ryanodine receptor variation in meat quality. Poult. Sci. 88: 1497-1505.

Suarez, S.D. and G. G. Gallup. 1983. Social reinstatement and open field testing in chickens. Anim. Learn. Behav. 11: 119-126.

Terlouw, C. 2005. Stress reactions at slaughter and meat quality in pigs: genetic background and prior experience - A brief review of recent. Livest. Prod. Sci. 94: 125-135.

Thompson, R. W.; J. Piroch, D. Fallen, and D. Hatton. 1974. A central cholinergic inhibitory system as a basis for tonic immobility (animal hypnosis) in chickens. J. Comp. Physiol. Psychol. 87: 507-512.

United States Department of Agriculture National Agricultural Statistics Service. 2014. Poultry – production and value. USDA, Washington, DC, USA. http://www.nass.usda.gov/Publications/Todays_Reports/reports/plva0414.pdf. Accessed Jun. 2014.

Updike, M. S., H. N. Zerby, M. S. Sawdy, G. Kaletunc, and M. P. Wick. 2005. Turkey breast meat functionality differences among turkeys selected for body weight and/or breast yield. Meat Sci. 71:706–712.

van Hierden, Y.M., J. M. Koolhaas, and S. M. Korte. 2004. Chronic increase of dietary - tryptophan decreases gentle feather pecking behaviour. Appl. Anim. Behav. Sci. 89: 71-84.

van Hierden, Y.M., J. M. Koolhaas, L. U. Kost'al, P. Vyboh, M. Sedlackova, M. Rajman, M. Jurani, and M. S. Korte. 2005. Chicks from a high and low feather pecking line of laying hens differ in apomorphine sensitivity. Physiol. .Behav. 84: 471-477.

van Hierden, Y.M., S. M. Korte, E. W. Ruesink, C. G. van Reenen, B. Engel, G. A. H. Kortebouws, J. M. Koolhaas, and H. J. Blokhuis. 2002. Adrenocortical reactivity and central serotonin and dopamine turnover in young chicks from a high and low feather-pecking line of laying hens. Physiology & Behavior 75: 653-659.

Voisinet, B. D., T. Grandin, S. F. O'Connor, J. D. Tatuma, and M. J. Deesing. 1997. Bos Indicus-cross feedlot cattle with excitable temperaments have tougher meat and a higher incidence of borderline dark cutters. Meat Sci. 4: 367-377.

Waiblinger, S., X. Boivin, V. Pedersen, M. V. Tosi, A. M. Janczak, E. K. Visser, and R. B. Jones. 2006. Assessing the human-animal relationship in farmed species: A critical review. Appl. Anim. Behav. Sci. 101: 185-242.

Wallnau, L. B., and G. G. Gallup. 1977. A serotonergic, midbrain-raphe model of tonic immobility. Biobehav. Rev. 1: 35-43.

Wechsler, B. 1995. Coping and coping strategies: a behavioural review. Appl. Anim. Behav. Sci. 43: 123-134.

Welfare Quality®. 2009a. Welfare Quality® Assessment Protocol for Poultry. Welfare Quality Consortium, Lelystad, Netherlands.

Welfare Quality®. 2009b. Welfare Quality® Assessment Protocol for Pigs. Welfare Quality Consortium, Lelystad, Netherlands.

Welfare Quality®. 2009c. Welfare Quality® Assessment Protocol for Cattle. Welfare Quality Consortium, Lelystad, Netherlands.

Werner, C., J. Riegel, and M. Wicke. 2008. Slaughter performance of four different turkey strains, with special focus on the muscle fiber structure and the meat quality of the breast muscle. Poult. Sci. 87:1849-1859.

Wilson, G. D. 1960. Chapter 8 quality factors in meat and meat foods. Page 259 in The Science of Meat and Meat Products. Salisbury, G. W., and E. W. Crampton, eds. W. H. Freeman and Company, San Francisco, CA, USA.

CHAPTER 2: TEMPERAMENTAL TURKEYS: RELIABILITY OF BEHAVIORAL RESPONSES TO FOUR TESTS OF FEAR

ABSTRACT

Open field (OF), tonic immobility (TI), voluntary approach (VA) and novel object (NO) tests are used to assess fear responses, activity levels and coping styles of poultry. Fear tests are also used as part of welfare assessment programs. Little is known about fear responses of turkeys. Specifically, it is unknown whether turkeys' fear responses are reliable which is important when developing tests for assessing welfare. This study examined the short- (between days) and long-term (between weeks) changes in, and test-retest reliability of, turkeys' fear responses. Male commercial turkeys were housed in groups of four to six in 16 pens. Turkeys were individually tested in OF (n = 60) and TI (n = 66) tests. VA and NO tests were administered to groups of turkeys in their pens (n = 16). Turkeys were tested three times in each test. The first (Period 1) and second (Period 2) tests were administered on consecutive days between 4 and 6 weeks, and the third test (Period 3) between 8 and 10 weeks. Other than increased sitting and reduced standing during OF testing and more birds approaching and pecking the observer during VA testing (Period 2 vs. 1), frequencies of responses did not differ between test periods. However, test responses (e.g. latencies to ambulate, vocalize, approach and peck) differed between periods depending on the fear test used. All VA test measures differed between periods, whereas most TI test measures did not. Most OF and NO test measures differed between weeks, but not between days. Except for the number of vocalizations (r_S = 0.39), most OF test responses were moderately reliable (r_S > 0.40). The latency to vocalize ($r_S \geq 0.51$) and number of vocalizations ($r_S \geq 0.59$) were the most reliable TI test measures, whereas reliability

of TI duration ($r_S \leq 0.31$) was low. All VA test measures were moderately to highly repeatable between days and weeks (e.g. latency to peck: $r_S \geq 0.67$). Reliability was lowest for the NO test, where only the latency to peck was moderately repeatable ($r_S = 0.61$). These findings suggest that although some fear responses of turkeys change over time and with repeated testing, most OF, TI and VA responses are reliable between days and weeks. However, few NO test measures were reliable under the conditions of this study. Further research is needed to assess the validity of OF, TI, VA and NO tests for assessing fear responses of turkeys.

INTRODUCTION

Fear and fearfulness have been major topics of poultry behavior and welfare research since 1965 with the release of the Brambell Report (Brambell, 1965) and the development of the Five Freedoms (Farm Animal Welfare Council, 1979). Fear is defined as the "reaction to the perception of actual danger" (Forkman et al., 2007), and fearfulness is defined as the susceptibility of an individual to being frightened (Boissy, 1995; Jones, 1996). Not only is fear a negative affective state that is indicative of suffering, but it is associated with numerous adverse effects on productivity and welfare (reviewed in Jones, 1996), such as increased injury (Reed et al., 1993), reduced production (Hemsworth et al., 1994; Cransberg et al., 2000; de Haas et al., 2013), depressed growth (Jones, 2002; Schütz et al., 2004), and feather pecking (Vestergaard et al., 1993; de Haas et al., 2010).

Underlying fearfulness cannot be measured directly because it is a subjective state. Therefore, inferences about a bird's level of fearfulness are based on the bird's responses to situations believed to induce fear. Fearful birds exhibit more intense fear responses compared to

less fearful birds (Jones, 1996). Moreover, birds that show greater fear responses have greater physiological responses to other stressors (Jones, 1989; Calandreau et al., 2011), which affects the bird's ability to cope with its environment. Fearfulness is considered by some to be a characteristic associated with coping style (Cockrem, 2007), which is defined as "a coherent set of behavioral and physiological stress responses which is consistent over time and which is characteristic to a certain group of individuals" (Koolhaas et al., 1999). However, others consider fearfulness and coping styles to be separate dimensions of personality (Jones and Boissy, 2011), where coping styles are discussed in relation to stress physiology (Groothuis & Carrere, 2005) and fear is associated with emotional reactivity which is distinct from coping style (Koolhaas et al., 1999). Irrespective of whether fearfulness is a component of coping style or an independent dimension associated with emotion, how the animal responds to environmental challenges affects the animal's welfare (Broom, 1988). To this end, fear tests have been incorporated into welfare assessment programs, such as the Welfare Quality® Assessment Protocols for poultry (Welfare Quality®, 2009a), pigs (Welfare Quality®, 2009b) and cattle (Welfare Quality®, 2009c).

The most common types of fear tests for poultry involve exposure to predatory stimuli (e.g. human approach or avoidance), restraint (e.g. tonic immobility), or novelty (e.g. novel object), or all three. However, fear tests are only usable if they are reliable. Reliable tests have few random errors and are consistent over repeated applications (Martin and Bateson, 1993). Reliability is distinct from validity, which refers to accuracy, freedom from systematic errors; specificity, whether the measurement measures what it is supposed to; and scientific validity, whether the measurement provides information relevant to the questions being investigated (Martin and Bateson, 1993). Forkman et al. (2007) assessed the reliability and validity of a number of fear tests used for various domestic farm animal species, including hens and quail.

Open field, tonic immobility and novel object tests are reliable and valid tests of fear for hens and quail (Forkman et al., 2007). Less information is available regarding birds' fear responses to humans, but Raubek et al. (2007) concluded that repeatability between consecutive days is high ($r_S > 0.80$) for the avoidance distance, stationary person and the touch test (performed in the litter area). Furthermore, Graml et al. (2008) demonstrated the validity of the tests used by Raubek et al. (2007) for assessing the reactions of hens toward humans.

In contrast to chickens and quail, there is scant research examining the welfare of turkeys (Marchewka et. al., 2013). Research examining fear or fearfulness per se is nearly nonexistent, even though turkey production has increased dramatically over the past three decades. Fear responses of turkeys have not been described or characterized, there is no information available regarding the reliability of any fear tests for turkeys, and results from studies with other poultry species may not be transferrable to turkeys. With increasing public interest in animal welfare and development of animal welfare assessment programs, there is an urgent need for information regarding fear and fearfulness of turkeys. The objectives of this study were to describe turkeys' responses to four tests of fear (tonic immobility, open field, novel object and voluntary approach tests) and to assess the short- and long-term reliability of the four tests.

MATERIALS AND METHODS

All procedures were approved by the Institutional Animal Care and Use Committee of Michigan State University.

Animals and housing

Commercial male turkeys (Hybrid Converter) were brooded in groups of 24 in one of four littered (wood shavings) pens (2.40 m x 3.05 m) from 1 day to 2 weeks of age. At 2 weeks, birds were distributed throughout the barn and randomly assigned to one of 16 littered pens (2.40 m x 3.05 m) in the same barn and housed in groups of four to six where they remained until study completion at 14 weeks. Poults were brooded under heat lamps (35 °C to 37 °C) for the first 7 days. After heat lamps were removed, the temperature was reduced to, and maintained at 30 °C for 7 days. Thereafter, the temperature was gradually reduced by 1 to 3 °C per week over a period of 6 weeks to reach a final temperature of between 13 and 18 °C. A photoperiod of 24L:0D was applied for the first 7 days. Thereafter, the photoperiod was reduced by 1 hr per day for 7 days. A photoperiod of 16L:8D was maintained from 14 days to 14 weeks. Lights came on at 06:00 h and were turned off at 22:00 h. Light intensity was 20 lx for the first 10 days and then reduced to 5 lx for the remainder of the study.

Turkeys were fed a commercial starter diet (Blue Seal® Multi Flock Game Starter, Kent Nutrition Group, Muscatine, IA, USA) from 1 to 8 weeks and a commercial grower diet (Blue Seal® Multi Flock Starter/Grower, Kent Nutrition Group, Muscatine, IA, USA) from 8 to14 weeks. At 2 weeks, turkeys were individually identified using blue non-toxic livestock marker (Prima Tech marking stick, Neogen Corporation, Lansing, MI, USA) which was applied to a different area for each bird in the pen (e.g. shoulders, left or right wing) and was visible on overhead cameras. Marker was reapplied biweekly to ensure markings remained visible. Body weight was recorded the day prior to testing at 4 weeks and 8 weeks.

Test procedures

For TI and OF tests, birds were individually removed from their home pens and tested in another room. VA and NO tests were conducted with groups of turkeys in their home pens to reduce confounds associated with environmental novelty, and because NO tests are used to assess welfare in groups of birds as part of welfare assessment programs (e.g. Welfare Quality®, 2009a). The same observer administered NO, VA and TI tests. Three different observers conducted OF tests during Periods 1 and 2. The same observer conducted all OF tests during Period 3. All observers wore blue coveralls and blue shoe covers. For OF and TI testing, turkeys were placed in plastic Rubbermaid® Roughneck™ bins (Rubbermaid®, Mogadore, OH, USA) measuring 67.3 (length) x 47.2 (width) x 43.4 (height) or 81.3 (length) x 50.8 (width) x 47.8 (height) containing a small amount of litter on the bottom and carried to and from testing areas.

Test schedule

Turkeys in each pen were tested three times (three test periods) in all four tests to examine test-retest reliability. Pens were randomly assigned to one of four test groups consisting of four pens. To control for possible carryover effects and to balance test order across test groups, each test group was randomly assigned to a particular test sequence (Table 2.1). Each test group was tested in only one test on a particular test day. In order to examine repeatability between days (Period 1 vs. 2), birds were tested in the same test (e.g. TI) over 2 consecutive days (Period 1 = day 1, Period 2 = day 2), followed by 2 days of no testing, followed by the next type of test (e.g. OF test) which was administered twice over 2 consecutive days (day 5 and day 6). This pattern of testing was used to control for possible carryover effects between test types. Therefore, birds were tested in all four tests over a period of 2 weeks (4 to 6 weeks of age).

Testing was then repeated 2 weeks later (Period 3) but birds were then tested only once in each test (3 days between test types), so that the first and third tests (Period 1 vs. 3) could be compared to assess repeatability between weeks. For example, birds in Group 1 were tested in the VA test at 28 days (Period 1) and 29 days (Period 2) to examine repeatability between days, and again at 56 days (Period 3) to examine repeatability between weeks (Period 1 vs. 3), and then in the NO test at 32 days (Period 1), 33 days (Period 2), and 60 days (Period 3, Table 2.1). Test order for each bird was randomized for each day of testing. All testing took place between 07:45 h and 17:45 h.

Test measures

Open field (OF) test. Turkeys were individually moved to an empty room adjacent to the room in which their home pens were located for testing. The OF test arena consisted of a square arena (length: 2.74 m, width: 2.74 m, height: 1.83 m) with a concrete floor. The testing arena was enclosed by solid black walls (1.22 m high) with grey mesh netting (0.61 m high) above the solid black sections. Black electrical tape was used to create a grid of 81 squares (each 0.30 m^2) on the floor of the test arena. Each bird was placed in the centre of the arena for 10 min. and behavior was recorded in real time (60i fields/s) from two overhead high definition camcorders (VIXIA HF M41, Canon USA, Inc. Melville, NY, USA). Behavioral data (latency to ambulate (s), latency to vocalize (s), number of vocalizations, number of squares entered, number of steps taken and escape attempts) were collected from video recordings. Ambulation was defined as two or more steps in rapid (within 4 s) succession. A bird was considered to have entered a square if at least 67% of both of the bird's feet were in the same square. An escape attempt was

defined as the bird attempting to jump and/or fly out of the test arena. Test arena walls were too high to permit successful escapes.

Tonic immobility (TI). Turkeys were individually tested in an empty pen in an empty room adjacent to the room in which their home pens were located. The observer restrained each turkey using methods similar to those of Noble et al. (1996a). Briefly, the turkey was placed on a table covered with a cloth sheet. The observer restrained the turkey on its left side by placing the left hand over the turkey's right wing and gently holding the turkey's legs with the right hand. The turkey was restrained in this manner for 15 s (induction period) and then released. If the turkey righted itself within 15 s after being released, the turkey was restrained twice more at most. If tonic immobility was not induced after a third attempt, the test was terminated and the turkey was recorded as not being inducible. If the turkey remained on its side after being induced, it was observed until it righted itself, or up to a maximum of 13 min. after which the test was terminated and the turkey was returned to its home pen. The number of inductions, number of vocalizations, latency to vocalize (s) and TI duration (s) were recorded.

Voluntary approach (VA) test. Turkeys' responses to an observer in their home pen were recorded for 15 min. The same observer conducted all tests. The observer entered the pen and stood against the wall across from the pen doors and remained motionless for the duration of the test. Turkeys' behavior during testing was recorded in real time (60i fields/s) with an overhead camera (2MB-70IR42L210 Jaguar Series infrared CCTV camera, 2M® Technology, Grand Prairie, TX, USA). Turkeys' responses, including latencies (s) to move within two body lengths, one body length, to approach, and to peck the person, were determined from video analyses. A body length was defined as the distance from the bird's shoulders to the base of its tail

(excluding tail feathers), and an approach was defined as the bird being less than one body length away and within pecking distance.

Novel object (NO) test. Procedures were similar to those for the VA test, with the exception that one of four novel objects, including a deflated beach ball, yellow bucket, wooden rod with colored tape similar to that used in the Welfare Quality® Assessment Protocol for Poultry (Welfare Quality®, 2009a), and six plastic balls strung together (two each painted red, white and blue), was placed in the center of the pen. Similar to studies with cattle (Gibbons et al., 2009), quail (Miller et al., 2005) and chickens (Hocking et al., 2001), different objects were used because objects could not be considered novel upon repeated exposure. Therefore, objects were balanced across test repeats to prevent habituation and maintain object novelty. Novel objects were balanced within test groups so that each pen in a particular test group was tested with a different object in each test period. In this manner, every object was represented on every test day, and each pen was only exposed to a particular novel object once. Each pen was therefore tested with three of the four objects (one object for each test period). Behavior during testing was recorded with an overhead camera from which behavioral data were collected. Cameras and behavioral measures that were recorded were the same as for VA testing.

Statistical analysis

All statistical analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, NC). For OF (n = 60) and TI (n = 66) tests, the experimental unit was the individual turkey. For NO and VA tests, the experimental unit was the pen (n = 16). Data were not normally distributed and transformations were ineffective. Levene's test for homogeneity of variance indicated that

variance was not equal. Body weight was included as a covariate in all analyses. Data consisted of binomial responses and frequencies, counts and latencies.

Binomial (yes/no) responses in the OF and TI tests were analyzed using the GENMOD procedure for a binomial distribution with the logit link function (e.g. Rodenburg et al., 2010). Turkey nested within pen was included as a repeated measure. Frequencies of birds in each pen that moved within two body lengths, one body length and that approached and pecked during NO and VA testing were compared between test periods using the GENMOD procedure for a poisson distribution. Pen was included as a repeated measure. For the NO test, object was also included as a covariate to account for differences between novel objects.

A characteristic of count data (e.g. number of steps or vocalizations) is the occurrence of an excess of zeros. In order to account for excess zeros, a zero-inflated Poisson model (Littell et al., 2006) was used to compare count variables between test periods (NLMIXED procedure). The parameter corresponding to pen was included in the model to account for similarities among turkeys from the same pen. The random intercept corresponding to the individual turkey was included in the model to account for repeated measures on the same individuals. For the TI test, the number of inductions was compared between test periods using the GENMOD procedure for a poisson distribution. Turkey nested within pen was included as a repeated measure.

Because fear test responses were only observed and recorded within a finite period of time (OF test: 10 min., TI: 13 min., and NO and VA tests: 15 min.), data were censored (truncated) such that many observations had values equal to the maximum test time allotted. For example, birds that did not ambulate or vocalize during OF testing were assigned a value of 600 s for latency to ambulate/vocalize. In order to account for the censored, non-normal data, the nonparametric LIFETEST procedure was used to compare latency measures (e.g. latency to

ambulate or vocalize) between test periods. Similar analyses have recently been used to examine censored observations in grey parrots (Van Zeeland et al., 2013) and grasshoppers (Karpestam et al., 2012). For the TI and OF tests, latencies from each individual bird were included in the analysis. Pen was specified in the STRATA statement, period was specified as the GROUP variable to compare latencies between periods while adjusting for pen differences, and weight was included as a covariate in the TEST statement. For the VA and NO tests, pen averages were used because pen was the experimental unit. Therefore, the average latencies to move within two body lengths, one body length, to approach and peck the object were calculated for each pen. Period was specified in the STRATA statement and weight was included as a covariate in the TEST statement. In order to determine whether responses during NO testing depended on which object was used, the LIFETEST procedure was used to compare latency measures between objects. Period was specified in the STRATA statement and object was specified as the GROUP variable to compare latencies between objects while adjusting for test period differences. The LIFETEST procedure calculates two test statistics (Log-rank, which places greater emphasis on differences in latencies occurring later in time, and Wilcoxon, which places greater emphasis on differences in latencies occurring earlier in time; SAS/STAT User's Guide, 2012) to test the effect of test period on the response variable (e.g. latency to ambulate). Both test statistics are reported. Results are presented as medians, 25-75% quartiles, χ^2 values for the log-rank and Wilcoxon test statistics, and the *P* value.

To assess test-retest reliability within each fear test, the Spearman rank correlation coefficient was used to assess correlations between Period 1 vs. Period 2 and Period 1 vs. Period 3. Correlation coefficients were defined as low: 0.2 – 0.4, moderate: 0.4 – 0.7, high: 0.7 – 0.9,

and very high: 0.9 – 1.0 (Martin and Bateson, 1993). P values < 0.05 were regarded as significant.

RESULTS

Mean (± SE) body weights were 1.68 ± 0.03 kg for Period 1 and 6.15 ± 0.08 kg for Period 3. For each type of fear test, only birds from which data for all three test periods had been collected were included in analyses.

Changes in fear responses over time and with repeated testing

Open field test. Turkeys' OF responses did not differ between observers. Analyses of binomial responses indicated that the percentages of turkeys sitting and standing differed between days (Period 1 vs. 2), but there were no other differences between test periods in any other OF responses (Table 2.2). Some turkeys attempted to escape the testing arena during Period 1 (13%) and Period 2 (8%), but no attempts were made in Period 3 (Table 2.2). The numbers of steps taken, squares entered and vocalizations produced increased over time and with repeated testing (Table 2.3). The latencies to ambulate and vocalize did not differ between days (Table 3), but were shorter for Period 3 than Period 1 (Table 2.3).

Tonic immobility. There was a trend for a higher percentage of turkeys to defecate in Period 3 (18%) vs. Period 1 (14%, Table 2.2). The number of vocalizations produced differed between test periods (Table 2.3), but no differences were found in the duration of TI or number of inductions required to induce TI.

Voluntary approach test. The percentage of turkeys that approached and pecked the person increased from Period 1 to Period 2 (Table 2.2) and latencies to move within two body

lengths, one body length, approach and peck decreased with each consecutive test period (Table 2.3).

Novel object test. NO test responses did not differ between objects for Period 1 and Period 2. However, when responses were compared between Period 1 and Period 3, latency to move within two body lengths (LIFETEST; Log-rank = 8.8, $P = 0.03$, Wilcoxon = 5.5, $P = 0.1$) differed depending on which object was used, and there were tendencies for the latency to move within one body length (LIFETEST; Log-rank = 7.1, $P = 0.07$, Wilcoxon = 5.3, $P = 0.2$) and latency to approach (LIFETEST; Log-rank = 7.5, $P = 0.06$, Wilcoxon = 7.6, $P = 0.06$) to differ between Periods 1 and 3 depending on the object used.

There was a tendency for the percentages of turkeys that moved within one body length and that approached the novel object to be higher for Period 3 than Period 1 (Table 2.2). No differences in latencies were found between Period 1 and Period 2 (Table 2.3). Latencies to move within two body lengths and one body length were shorter for Period 3 than Period 1 (Table 2.3).

Test-retest reliability

Spearman rank correlation coefficients for correlations between days and weeks are presented in Table 2.4. Significant, moderate correlations were found between Period 1 and Period 2 for all OF test measures. With the exception of a low correlation for the number of vocalizations, and no significant correlation for latency to vocalize, moderate correlations were also found between Period 1 and Period 3 for all other OF test measures. Moderate correlations were found between Period 1 vs. 2 and Period 1 vs. 3 for TI vocalization latency and number of vocalizations, whereas low correlations were found for TI duration. The number of inductions

was not significantly correlated between Period 1 and 3. Moderate to high correlations were found for all VA test measures. For the NO test, the highest correlation was found for latency to peck (Period 1 vs. 3). No other NO test measures were significantly correlated.

DISCUSSION

Previous experiments using fear tests for turkeys have used fear tests in a variety of applications, ranging from comparing fear responses of turkeys housed in different rearing systems (TI: Noble et al., 1996a), to examining the effects of dietary ascorbic acid on TI reactions (Konca et al., 2009), assessing motor activity (OF test: Kowalski et al., 2002) and activity level (OF test: Huff et al., 2007), and examining the effects of beak trimming on tameness (NO test: Noble et al., 1996b). Notably, fear responses of turkeys had not been described, characterized or assessed for test-retest reliability.

Our results indicated that responses were highly variable among turkeys with some turkeys displaying very little, if any reaction during testing, whereas other turkeys were extremely active and vocal. For example, between 27% (Period 1) and 53% (Period 3) of turkeys ambulated, with the number of steps taken reaching as high as 382 steps when tested in the OF. Few other studies have reported OF behavior of turkeys, and those that have are not directly comparable to our study due to differences in methodology and how test measures were defined and recorded. Kowalski et al. (2002) examined OF activity in two strains of commercial turkeys (BIG-6 and BUT-9) as a measure of motor activity and emotional reactivity. Results were reported as means (± SD) of activity levels in different areas of the OF (center, periphery), and episodes of grooming, vocalization, sitting, and defecation. It is not stated how episodes were defined and quantified. Another study examined “freezing and active behaviors” of male and female turkeys

of three different genetic lines in an OF test at 8 days (Huff et al., 2007). Birds were also tested for the latency to reach an area of the OF adjacent to another pen of turkeys that was visible to the test bird. OF test results were used to compare activity levels between sexes and strains (Huff et al., 2007), and not to examine fear per se.

Similar to the OF test, few studies have examined TI responses of turkeys. Tonic immobility was induced in most turkeys following only one induction, and a large percentage of turkeys (39% in Period 1 and 58% in Period 3) remained immobile for the entire 780 s testing period. Previous studies stopped TI testing at 60 s (Noble et al., 1996a; 1996b) and 600 s (Konca et al., 2009). Mean TI durations in four different turkey lines tested at 20 weeks ranged from 6 to 20 s (Noble et al., 1996a) and 278.0 ± 38.3 to 332.9 ± 39.3 s in commercial (BUT-6) turkeys tested at 16 weeks (Konca et al., 2009). Because data were not normally distributed, median TI durations were reported here, and results are not directly comparable to those of Noble et al. (1996a) and Konca et al. (2009). In an early study with chickens, it is stated that TI is characterized by a period of reduced arousal, although some birds may vocalize (Gentle et al., 1989). In contrast, as many as 42% (Period 1) and 26% (Period 3) of turkeys vocalized during TI. Interestingly, TI measures associated with vocalization were more highly correlated between test periods than TI duration. However, vocalizations are not often measured or reported as part of TI, because unlike TI duration, vocalization during TI has not been shown to be an indicator of the level of fear (see Forkman et al., 2007). Similar to OF and TI test responses, NO and VA responses varied greatly among individual turkeys, with some birds pecking within 8 s (VA test) and 18 s (NO test) and other birds not pecking at all. On average, the majority of birds (≥ 69%) pecked the person during the VA test, and between 37% (Period 1) and 62% (Period 3) of birds pecked the novel object during NO testing.

Measures that are closely spaced in time are more likely to have a greater degree of relatedness than measures that are more distantly spaced in time. Therefore, little difference would be expected between test measures from Period 1 vs. Period 2, whereas a larger difference would be expected between Period 1 and Period 3. In general, results were reflective of a greater difference between Period 1 and Period 3 test measures, with the exception of the VA test, where all measures were significantly different between days and between weeks. For most test measures, the likelihood of a behavioral response occurring (binomial responses) and frequencies of responses did not change between test periods, but the rate at which behavioral responses were performed changed. Specifically, the number of steps (OF test) and vocalizations (TI and OF tests) changed between test periods, indicating that these behavioral responses change over time and with repeated testing. Unfortunately, it was not possible to record vocalizations during VA and NO testing due to background noise from fans, heaters and other birds in the barn. Further research is needed to examine vocalization during NO and VA tests. Latencies to ambulate and vocalize during OF testing, and latencies to move within two and one body lengths during NO testing were all significantly shorter in Period 3 than Period 1, whereas there were no differences in these test measures between Period 1 vs. 2. Contrarily, TI test responses (except for the number of vocalizations) did not differ significantly between days and between weeks. Some level of habituation is to be expected with repeated testing. Therefore, changes in test responses may be indicative of habituation or age-related changes in behavior, or both, but it is not possible to distinguish between changes due to age or due to habituation.

For most test measures, test repeatability was higher between days than between weeks, as indicated by higher Spearman rank correlation coefficients for Period 1 vs. 2 than for Period 1 vs. 3. Moderately high ($r_S > 0.4$) correlations were found for most TI, OF, and VA test measures.

The highest correlations were found for VA test measures. The levels of correlation found among VA test measures are comparable to those from a study examining human-animal interactions on commercial turkey farms (Botheras et al., 2008). They reported that the number of turkeys within a prescribed distance from a person who was at times moving and at times stationary was significantly, positively correlated ($r = 0.67$, $P < 0.01$) between 8 and 12 weeks (during the stationary phase of the test). However, the numbers of turkeys within proximity of the person were not correlated between 4 and 8 weeks or 4 and 12 weeks (Botheras et al., 2008), whereas our results indicated moderate to high correlations between Period 1 (4 to 6 weeks) and Period 3 (8 to 10 weeks). Turkeys in our study received frequent human contact because they were tested in multiple tests and because marker was reapplied bi-weekly to enable individual identification of birds. The amount of human contact likely affected their response to the observer during the VA test such that few turkeys were actually fearful of the observer. Indeed, Botheras et al. (2008) stated that turkeys receiving more human contact were less fearful of humans. In our study, latencies to approach and peck the observer were shorter with each consecutive VA test, indicating reduced fearfulness of the observer over time. However, responses were highly correlated, suggesting that although fearfulness of the observer may have been reduced, turkeys' responses were reliable. It is possible that increased human contact may influence reliability of test responses, but further research is necessary to evaluate whether responses of animals experiencing more, or consistent, human contact are more reliable than responses of animals receiving little, or inconsistent, human contact.

In contrast to the VA test, the only NO test measure that was significantly correlated was the latency to peck the NO (Period 1 vs. 3). These results are in agreement with those from a study with quail, which indicated that the latency to peck the NO had the highest correlation coefficient

of all NO test measures recorded (Miller et al., 2005). In both our study and that of Miller et al. (2005), different novel objects were used and objects were balanced across test repeats. Low repeatability of the NO test probably resulted from using different objects for each test repeat. Indeed, the latency to move within two body lengths of the NO differed depending on the object used when Period 1 responses were compared to those of Period 3, indicating that some objects may have been perceived as more frightening than others. However, exposing birds to the same object repeatedly would result in the object no longer being novel, creating a dilemma between test repeatability and object novelty. As discussed by Miller et al. (2005) using quail, and Gibbons et al. (2009) using cattle, using the same object repeatedly would have increased habituation and the likelihood of turkeys responding to specific features of the object, which in turn reduces generalizability of the results.

According to Cohen (1988), a Pearson's product moment correlation coefficient of 0.1 is considered a small effect size, 0.3 is a medium effect size, and a large effect size is indicated by a correlation coefficient of 0.5 or larger. If Cohen's large effect size criterion is applied to correlations between Period 1 and 3, some measures from all four tests meet the criterion (OF test: latency to ambulate, number of steps, number of squares, and number of defecations; TI test: latency to vocalize and number of vocalizations; VA test: all measures; NO test: latency to peck). Cohen's large effect size criterion has been used by others to select fear tests measures that were reliable enough for testing the validity of fear tests in quail (Miller et al., 2006).

Most studies of test repeatability expose animals to the same conditions repeatedly and then assess the degree of relatedness (correlation) between test repeats. However, it is possible that fear responses may be consistent, not because they are characteristic of a stable personality trait, but because the conditions under which responses are elicited and conditions under which

animals are reared are kept constant (Chesler et al., 2003; Crabbe et al, 1999; Miller et al., 2005). Because no research has previously examined fear test reliability using turkeys, a first step is to determine whether test responses are consistent under stable conditions, and then further research can be conducted to examine how varying environmental conditions affect test reliability. Further research is needed to examine to what degree other variables may influence results when they are not held constant over test situations. Research is also needed to assess the validity of the four tests for assessing fear.

In their review of fear tests for hens and quail, Forkman et al. (2007) concluded that the OF, TI and NO tests are reliable for testing fear in hens and quail. Ambulation in the open field is heritable and there is a QTL related to OF behavior, suggesting that the response is stable within individuals (Forkman et al., 2007). Similarly, there is a QTL for TI and previous research has shown TI to be repeatable (reviewed in Forkman et al., 2007). The NO test is repeatable between days and weeks (Forkman et al., 2007). Repeatability of hens' responses to humans was demonstrated in a study by Raubek et al. (2007), who concluded that hens' responses to the stationary person test, avoidance distance test and touch test are repeatable between days. Results from our study are generally in agreement with results from studies with hens and quail that responses to TI, OF and VA tests are repeatable between days and weeks. However, the only NO measure that was reliable was the latency to peck.

In conclusion, fear responses of turkeys are not necessarily stable over time, but responses to TI, OF and VA tests are repeatable because most test measures are moderately correlated between days and weeks. Caution is needed when using the NO test as a welfare assessment tool, because under the conditions of this study, only the latency to peck the object

was a reliable measure. These results may aid in the development of measures for assessing welfare of turkeys, but further research examining fear test validity is needed.

ACKNOWLEDGEMENTS

We thank Hybrid Turkeys for supplying the poults, Dr. Darrin Karcher, and Erika First, Caitlin Hummel, Sarah Jager, Cheyenne Ryzenga and Shannon Williams for assistance with data collection. This work was supported by a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship and Michigan State University AgBioResearch.

APPENDIX

Table 2.1. Testing schedule and test order for each group of four pens. Each test group was tested three times in four tests of fear (VA = voluntary approach test, NO = novel object test, OF = open field test, TI = tonic immobility).

		Test used and order of testing			
Test Period	Age (d)	Group 1	Group 2	Group 3	Group 4
1	28	VA	TI	NO	OF
2	29	VA	TI	NO	OF
1	32	NO	VA	OF	TI
2	33	NO	VA	OF	TI
1	36	OF	NO	TI	VA
2	37	OF	NO	TI	VA
1	40	TI	OF	VA	NO
2	41	TI	OF	VA	NO
3	56	VA	TI	NO	OF
3	60	NO	VA	OF	TI
3	64	OF	NO	TI	VA
3	68	TI	OF	VA	NO

Table 2.2. Differences in percentages of turkeys performing a particular behavior in the open field, tonic immobility, voluntary approach and novel object tests between days (test Period 1 vs. 2) and between weeks (Period 1 vs. 3).

	Period 1	Period 2	χ^2	Period 3	χ^2
		Open Field (n = 60)			
Standing	88.0	78.0	4.5*	93.0	0.02
Sitting	20.0	28.0	4.5*	47.0	0.1
Ambulation	27.0	40.0	1.0	53.0	2.8^T
Vocalization	57.0	63.0	2.7	83.0	1.4
Escape	13.0	8.0	0.1	0	.
Defecation	87.0	78.0	1.6	95.0	0.1
		Tonic Immobility (n = 66)			
One induction	77.0	70.0	0.1	71.0	0.3
Vocalization	42.0	38.0	0.1	26.0	1.6
Immobile for full duration	39.0	52.0	1.6	58.0	0.2
Defecation	14.0	17.0	1.0	18.0	3.0^T
		Voluntary Approach (n = 16)			
Within two BL	84.3	95.0	1.7	97.7	0.5
Within one BL	79.0	92.2	2.5	95.2	1.1
Approach	72.9	89.6	5.2*	93.6	0.9
Peck	69.2	88.0	5.1*	91.1	1.2
		Novel Object (n = 16)			
Within two BL	78.1	67.5	2.1	95.0	0.2
Within one BL	61.0	55.0	0.5	81.3	3.1^T
Approach	50.4	50.0	0	73.8	3.5^T
Peck	37.1	47.7	1.2	62.1	2.2

Note: BL = body length.
$^T P < 0.10$, $^* P < 0.05$.

Table 2.3. Differences in median (25-75% range) values of open field, tonic immobility, voluntary approach and novel object test measures between Period 1 and Period 2 and Period 1 and Period 3. Testing occurred at 28 to 40 days (Period 1), 29 to 41 days (Period 2), and 56 to 68 days (Period 3).

Test measure	Period 1	Period 2	Test Statistic Period 1 vs. 2	Period 3	Test Statistic Period 1 vs. 3
			Open field test (n = 60)		
Ambulate (s)	600 (487-600)	600 (22-600)	L = 1.7, W = 1.9	550 (183-600)	L = 8.3*, W = 11.5**
Step no.	0 (0-7)	0 (0-18)	t = 6.3***	7 (0-21)	t = 8.0***
Square no.	0 (0-1)	0 (0-1)	t = 3.4**	0 (0-3)	t = 3.2*
Vocalize (s)	381 (127-600)	285 (73-600)	L = 1.4, W = 1.9	52 (10-163)	L = 22.5***, W = 29.7***
Vocalize no.	6 (0-168)	19 (0-197)	t = 5.6***	99 (10-240)	t = 11.6***
Defecate no.	2 (1-3)	2 (1-3)	t = 0.5	2 (1-3)	t = 1.5
			Tonic immobility test (n = 66)		
Induction no.	1 (1-1)	1 (1-2)	χ^2 = 1.5	1 (1-2)	χ^2 = 1.4
Dur. (s)	589 (321-780)	780 (365-780)	L = 0.6, W = 0.3	780 (493-780)	L = 2.4, W = 1.1
Vocalize (s)	780 (189-780)	780 (65-780)	L = 0.1, W = 0	780 (630-780)	L = 3.3^T, W = 2.5
Vocalize no.	0 (0-21)	0 (0-4)	t = 10.3***	0 (0-2)	t = -12.5***
			Voluntary approach test (n = 16)		
Two BL (s)	103 (59-369)	24 (18-32)	L = 7.6*, W = 10.9**	29 (24-35)	L = 11.7**, W = 10.6*
One BL (s)	244 (133-518)	32 (22-46)	L = 7.4*, W = 11.6**	48 (36-61)	L = 10.2*, W = 9.5*
Approach (s)	315 (181-559)	52 (32-75)	L = 7.0*, W = 9.7*	66 (44-194)	L = 9.8*, W = 8.9*
Peck (s)	449 (293-817)	75 (41-133)	L = 8.6*, W = 12.4**	115 (78-259)	L = 12.2**, W = 10.7*
			Novel object test (n = 16)		
Two BL (s)	69 (18-474)	143 (9-634)	L = 0.001, W = 0.3	9 (3-57)	L = 11.6**, W = 10.3**
One BL (s)	426 (33-702)	323 (2-900)	L = 0.4, W = 0.1	24 (12-388)	L = 6.3*, W = 7.3*
Approach (s)	563 (113-900)	601 (79-900)	L = 0.2, W = 0.1	153 (24-730)	L = 1.4, W = 2.9^T
Peck (s)	764 (450-900)	680 (254-900)	L = 0.02, W = 0.2	440 (97-736)	L = 2.3, W = 2.4

Note: Lat. = latency, BL = body length, L = log-rank, W = Wilcoxon.
^{T}P < 0.10, * P < 0.01, **P < 0.001, *** P < 0.0001.

Table 2.4. Spearman rank correlation (r_S) of test measures between days (Period 1 vs. 2) and weeks (Period 1 vs. 3) for four tests of fear (open field, tonic immobility, voluntary approach, and novel object tests).

Test	Behavioural response	Period 1 vs.2	Period 1 vs. 3
Open Field	Lat. ambulate	0.62****	0.66****
(*n* = 60)	Lat. vocalize	0.47***	0.17
	Step no.	0.65****	0.58****
	Square no.	0.71****	0.59****
	Vocalization no.	0.51****	0.39**
	Defecation no.	0.47***	0.50****
Tonic	Duration	0.35**	0.31*
Immobility	Lat. vocalize	0.57****	0.51****
(*n* = 66)	Induction no.	0.39**	0.08
	Vocalization no.	0.67****	0.59****
Voluntary	Lat. two BL	0.63**	0.62*
Approach	Lat. one BL	0.62*	0.62**
(*n* = 16)	Lat. approach	0.81***	0.87****
	Lat. peck	0.91****	0.67**
Novel Object	Lat. two BL	0.39	0.34
(*n* = 16)	Lat. one BL	0.38	0.23
	Lat. approach	0.29	0.15
	Lat. peck	0.35	0.61*

Note: Lat. = latency, BL = body length.
[T]P*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

REFERENCES

REFERENCES

Boissy, A. Fear and fearfulness in animals. 1995. *Q. Rev. Biol.*. 70: 165-191.

Botheras, N. A., J. A. Pempek, and D. E. Enigk. 2008. Human-animal interactions in the turkey industry. Midwest Poultry Research Program Project Database. Available online: http://www.mwpoultry.org/ProjectPDFs/08-09.pdf

Brambell, F. W. R. (1965). Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. (Command Rep. 2836). London: Her Majesty's Stationery Office.

Broom, D. M. 1988. The scientific assessment of animal welfare. *Appl. Anim. Behav. Sci.* 20: 5-19.

Calandreau L., A. Favreau-Peigné, A. Bertin, P. Constantin, C. Arnould, A. Laurence, S. Lumineau, C. Houdelier, M. A. Richard-Yris, A. Boissy, and C. Leterrier. 2011. Higher inherent fearfulness potentiates the effects of chronic stress in the Japanese quail. *Behav. Brain Res.* 225: 505-510.

Chesler, E. J.,S. G. Wilson, W. R. Lariviere, S. L. Rodriguez-Zas, and J. S. Mogil. 2003. Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neurosci. Biobehav. Rev.* 26: 907–923.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, New Jersey: L. Erlbaum. Pp 79-81.

Cockrem, J. F. 2007. Stress, corticosterone responses and avian personalities. *J Ornithol*. 148 (Suppl 2): S169–S178.

Crabbe, J.C., D. Wahlsten, and B. C. Dudek. 1999. Genetics of mouse behavior: interactions with laboratory environment. *Science*. 284: 1670–1672.

Cransberg, P. H., P. H. Hemsworth, and G. J. Coleman. 2000. Human factors affecting the behavior and productivity of commercial broiler chickens. *Br. Poult. Sci.* 41: 272-9.

de Haas, E. N., B. L. Nielsen, A. J. (Bart) Buitenhuis, T. B. Rodenburg. 2010. Selection on feather pecking affects response to novelty and foraging behavior in laying hens. *Appl. Anim. Behav. Sci.* 124: 90-96.

de Haas, E. N., B. Kemp , J. E. Bolhuis , T. Groothuis , and T. B. Rodenburg. 2013. Fear, stress, and feather pecking in commercial white and brown laying hen parent-stock flocks and their relationships with production parameters. *Poult. Sci.* 92: 2259-2269.

Farm Animal Welfare Council. 1979. Press Statement. Available Online: http://www.fawc.org.uk/pdf/fivefreedoms1979.pdf

Forkman, B., A. Boissy, M. –C. Maunier-Salaün, E.Canali, and R. B. Jones. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiol. Behav.* 92: 340-374.

Gentle, M. J., R. B. Jones, and S. C. Woolley. 1989. Physiological changes during tonic immobility in *Gallus gallus* var *domesticus. Physiol. Behav.* 46: 843-847.

Gibbons, J., A. Lawrence, and M. Haskell. 2009. Responsiveness of dairy cows to human approach and novel stimuli. *Appl. Anim. Behav. Sci.*116: 163-173.

Graml., C., S. Waiblinger, and K. Niebuhr. 2008. Validation of tests for on-farm assessment of the hen-human relationship in non-cage systems. *Appl. Anim. Behav. Sci.* 111: 301-310.

Groothuis, T. G. G., and C. Carere. 2005. Avian personalities: characterization and epigenesis. *Neurosci. and Biobehav. Rev.* 29: 137-150.

Hemsworth, P. H., G. J. Coleman, J. L. Barnett, and R. B. Jones. 1994. Behavioral responses to humans and the productivity of commercial broiler chickens. *Appl. Anim. Behav. Sci.* 41: 101-114.

Hocking, P. M., C. E. Channing, D. Waddington, and R. B. Jones. 2001. Age-related changes in fear, sociality and pecking behaviors in two strains of laying hen. *Br. Poult. Sci.* 42: 414-423.

Huff, G., W. Huff, N. Rath, A. Donoghue, N. Anthony, and K. Nestor. 2007. Differential effects of sex and genetics on behavior and stress response of turkeys. *Poult. Sci.* 86: 1294-1303.

Jones, R. B. 1989. Chronic stressors, tonic immobility and leucocytic responses in the domestic fowl. *Physiol. Behav.* 46: 439-442.

Jones, R. B. 1996. Fear and adaptability in poultry: insights, implications and imperatives. *World's Poult. Sci. J.* 52: 131-174.

Jones, R. B. 2002. Role of comparative psychology in the development of effective environmental enrichment strategies to improve poultry welfare. *Int. J. Comp. Psychol.* 15:77-106.

Jones, B., and A. Boissy. 2011. Fear and other negative emotions. In: Animal welfare. Second ed. Appleby, M. C, J. A. Mench, I. A. S. Olsson, and B. O. Hughes (eds.). pp 81.

Karpestam, E., S. Merilaita, and A. Forsman. 2012. Reduced predation risk for melanistic pygmy grasshoppers in post-fire environments. *Ecol. Evol.* 2: 2204-2212.

Konca, Y., F. Kirkpinar, and M. Çabuk. 2009. Effects of dietary ascorbic acid on blood haemoatological profile, serumbiochemical components and tonic immobility reaction of male turkeys under summer condition. *J. Poult. Sci.* 46: 105-111.

Koolhaas, J. M., S. M. Korte, S. F. De Boer, B. J. Van Der Vegt, C. G. Van Reenen, H. Hopster, I. C. De Jong, M. A. W. Ruis, and H. J. Blokhuis. 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neurosci. Biobehav. Rev.* 23: 925–935.

Kowalski, A., P. Mormède, K. Jakubowski, and M. Jedlińska-Krakowska. Comparison of susceptibility to stress in two genetic lines of turkey broilers, BUT-9 and Big-6. *Polish J. Vet. Sci.* 5: 145-150.

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. SAS® for mixed models (2nd ed.)SAS Institute Inc., North Carolina, USA, pp. 589-594.

Marchewka, J., T. T. N. Watanabe, V. Ferrante, and I. Estevez. 2013. Review of the social and environmental factors affecting the behavior and welfare of turkeys (*Meleagris gallopavo*). *Poult. Sci.* 92:1467–1473.

Martin, P., Bateson, P., 1993. Measuring Behavior, An Introductory Guide. Second ed. Cambridge University Press, New York, USA, pp. 1–222.

Miller, K. A., J. P. Garner, and J. A. Mench. 2005. The test-retest reliability of four behavioral tests of fearfulness for quail: a critical evaluation. *Appl. Anim. Behav. Sci.* 92:113-127.

Miller, K. A., Garner, J. P., and J. A. Mench. 2006. Is fearfulness a trait that can be measured with behavioral tests? A validation of four fear tests for Japanese quail. *Anim. Behav.* 71: 1323-1334.

Noble, D. O., J. W. Anderson, and K. E. Nestor. 1996a. Range and confinement rearing of four genetic lines of turkeys. 2. Effects on behavior and tonic immobility. *Poult. Sci.* 75: 165-171).

Noble, D. O., K. E. Nestor, and K. K. Krueger. 1996b. The effect of beak trimming on two strains of commercial turkeys. 2. Behavior traits. *Poult. Sci.* 75: 1468-1471.

Raubek, J., K. Niebuhr, and S. Waiblinger. 2007. Development of on-farm methods to assess the animal-human relationship in laying hens kept in non-cage systems. *Anim. Welfare.* 16: 173-175.

Reed, H. J., L. J. Wilkins, S.D. Austin, and N. G. Gregory. 1993. The effect of environmental enrichment during rearing on fear reactions and depopulation trauma in adult caged hens. *Appl. Anim. Behav. Sci.* 36: 39-46.

Rodenburg, T. B., de Haas, E. N., Nielsen, B. L., and A. J. Buitenhuis. 2010. Fearfulness and feather damage in laying hens divergently selected for high and low feather pecking. *Appl. Anim. Behav. Sci.*128: 91-96.

SAS/STAT User's Guide. 2012. SAS/STAT 12.1 user's guide the LIFETEST procedure (chapter). Available online: http://support.sas.com/documentation/onlinedoc/stat/121/lifetest.pdf

Schütz, K. E., S.Kerje, L.Jacobsson, B. Forkman, O. Carlborg, L. Andersson, and P. Jensen. 2004. Major growth QTLs in fowl are related to fearful behavior: possible genetic links between fear responses and production traits in a red junglefowl x white leghorn intercross. *Behav Genet*. 34: 121-30.

van Zeeland, Y. R. A., M. M .J. A. van der Aa, C. M. Vinke, J. T. Lumeij, and N. J. Schoemaker. 2013. Behavioral testing to determine differences between coping styles in Grey parrots (*Psittacus erithacus erithacus*) with and without feather damaging behavior. *Appl. Anim. Behav. Sci.* 148: 218-231.

Vestergaard, K. S., J. P. Kruijt, and J. A. Hogan. 1993. Feather pecking and chronic fear in groups of red junglefowl: their relations to dustbathing, rearing environment and social status. *Anim. Behav.* 45: 1127-1140.

Welfare Quality®. 2009a. Welfare Quality® Assessment Protocol for Poultry. Welfare Quality Consortium, Lelystad, Netherlands.

Welfare Quality®. 2009b. Welfare Quality® Assessment Protocol for Pigs. Welfare Quality Consortium, Lelystad, Netherlands.

Welfare Quality®. 2009c. Welfare Quality® Assessment Protocol for Cattle. Welfare Quality Consortium, Lelystad, Netherlands.

CHAPTER 3: USING INTER-TEST CORRESPONDENCE TO ASSESS THE VALIDITY OF FOUR TESTS OF FEAR FOR TURKEYS

ABSTRACT

Fear and fearfulness are the most widely studied aspects of poultry temperament, but little is known regarding turkeys' fear responses. If fearfulness is part of temperament, fear tests should be reliable and valid. This study evaluated inter-test correspondence as a measure of validity among four tests (tonic immobility (TI), open field (OF), voluntary approach (VA), and novel object (NO) tests). Male turkeys (*Meleagris gallopavo*) were housed in groups of four to six in 16 pens ($N = 16$). Turkeys were tested in all four tests between 4 and 6 weeks (Age 1) and 8 and 10 weeks (Age 2). Test correspondence was assessed using pair-wise correlations and by examining whether turkeys showing extreme responses (high (HR) vs. low (LR) responders) in each test situation differed across other test situations. At Age 1, the VA and OF tests tended to be correlated ($r_S = 0.43$, $P < 0.10$), and turkeys classified as HR in the VA test (VA-HR) tended to be more active in the OF test vs. VA-LR ($S = 141$, $P < 0.10$). At Age 2, OF responses were moderately correlated with TI vocalizations ($r_S = -0.50$, $P < 0.05$) and tended to be correlated with NO test behavior ($r_S = -0.44$, $P < 0.10$). Compared to NO-LR, NO-HR were significantly more active in the OF and took less time to peck when tested in the VA test. Overall, results reveal a weak association between turkeys' willingness to move toward an observer during the VA test and their ambulatory behavior in an OF test at Age 1. There may be stronger associations among tests at Age 2, but age was confounded with test habituation. Results did not validate that OF, TI, VA and NO tests measure the same types of fear responses. Fear responses may therefore be situation-dependent.

INTRODUCTION

Individual differences in behavior are widely recognized in different vertebrate species, ranging from fish to birds and mammals. Research into individual differences continues to increase and has widespread applications, including applications to human health and disease, food quality, and animal well-being. For instance, rodents are used as models for human disorders such as post traumatic stress disorder and anxiety (reviewed in Bourin et al., 2007, and Neumann et al., 2011) because individual differences in behavior facilitate the development and testing of drugs for treating these disorders. In domestic animal species, individual differences in behavior are used in the development of selection and breeding programs to increase favorable traits or reduce harmful or unwanted traits. Indeed, individual behavioral differences are associated with differences in disease susceptibility in rats, calves and pigs (reviewed in Koolhaas et al., 1999), and production and meat quality in steers (e.g. Petherick et al., 2002; Voisinet et al., 1997) and quail (e.g. Remignon et al., 1998).

Different terminology is used in the context of individual differences in behavior, including the terms personality, temperament, coping styles, coping strategies, behavioral syndromes and behavioral profiles, among others. Several review papers discuss the definitions of, and differences among, these terms (Cockrem, 2007; Gosling, 2001; Groothuis & Carere, 2005; B. Jones & Boissy, 2011; A. C. Jones & Gosling, 2005; Koolhaas et al., 1999; Stamps & Groothuis, 2010). In general, temperament, personality, and coping styles are considered to have similar meanings (Cockrem, 2007; Stamps & Groothuis, 2010). However, "dimensions of personality are generally described in terms of coping styles or fearfulness" (B. Jones and Boissy, 2011), where coping styles concern stress physiology (Groothuis & Carrere, 2005), and fear is a characteristic associated with emotional reactivity and not coping style (Koolhaas et al.,

1999). However, some descriptions of coping styles also include fearfulness as a coping style characteristic (e.g. Cockrem, 2007).

If fearfulness is part of temperament or coping styles, then fearfulness should be stable in individuals over time and tests used to assess fear responses should be reliable and valid even if tests differ in the specific fear stimuli used. Correspondence among fear tests can be used as a method of validating that tests are measuring fear responses, and not some other response. Indeed, Forkman et al. (2007) state that "validation through evaluation of inter-test consistency should be given more attention in future studies". In their review of fear tests for chickens and quail, Forkman et al. (2007) conclude that the open field and tonic immobility tests are reliable and valid, and the novel object test is a reliable and practical test of fear for chickens and quail. Graml et al. (2008) concluded that the voluntary approach test (stationary person test) is a valid test for assessing the hen-human relationship. However, there is conflicting evidence regarding fear test correspondence. Older studies appear to indicate some correspondence among fear tests (e.g. R. B. Jones, 1996), whereas more recent studies have failed to show correspondence, especially among tests assessing novelty and tests assessing predatory-type stimuli (e.g. Miller et al., 2006).

Several strategies have been used to validate and assess correspondence among tests intended to assess fear or other components of animal temperament. One strategy is to examine whether differences between behaviorally distinct groups of animals remain consistent across various test situations. Animals are tested in one type of fear test, such as a tonic immobility (TI) test, so that they can be assigned to behaviorally distinct groups, such as short TI and long TI durations. The groups are then tested in another test, such as a novel object test, and their behavior is compared to assess whether TI and novel object (NO) tests correspond. If tests

correspond, short TI and long TI groups will differ in their responses to the NO. This strategy has been used to assess tests of coping styles in pigs (Spake et al., 2012). A related method is to compare animals from genetic lines that have been bidirectionally selected for a particular behavioral or physiological response to ascertain whether their responses differ in other situations as well. For example, genetically selecting for animals with short and long TI durations and then comparing NO test responses between genetic lines (e.g. Hazard et al., 2008; Richard et al., 2008). A third method is to use correlations to assess the strength of the relationship between pairs of test measures from different tests (e.g. Albentosa et al., 2003). Similarly, factor analyses can also be used to evaluate whether measures from different tests are related (e.g. Mazurek et al., 2011; Miller et al., 2006). Another strategy is to use drugs, such as anti-anxiolytics, to determine whether fear and anxiety-related behavior of treated and untreated animals differ when tested in fear tests (e.g. Marín et al., 1997). Lastly, animals can be assigned to different treatments designed to increase or reduce fear (e.g. increased positive human contact to reduce fear), and responses to other fear tests can be compared among the treatment groups to determine whether groups that display reduced fear as a result of the treatment also exhibit reduced fear under other fear-eliciting conditions (e.g. Graml et al., 2008).

Most research concerning fearfulness of domestic poultry has used chickens and quail, but there is scant research examining fear or fearfulness of turkeys. Chickens are one of the most widely studied production animal species, and quail are an important species used in laboratory research. However, domestic turkey production has increased greatly in recent decades, yet there is very little research examining behavior in this species compared to chickens and quail, despite the importance of behavior to turkey welfare, management and temperament. We recently demonstrated that some measures of tonic immobility, open field, voluntary approach and novel

object tests are reliable for assessing fear responses of turkeys (Erasmus & Swanson, 2014). This study builds upon our previous results. The objectives were to assess inter-test correspondence among tonic immobility, open field, voluntary approach, and novel object tests for turkeys using two strategies, including 1) calculating correlations between test measures, and 2) evaluating whether differences among behaviorally distinct groups of turkeys remain consistent across test situations.

MATERIALS AND METHODS

All procedures were approved by the Institutional Animal Care and Use Committee of Michigan State University.

General procedures

A full description of materials and methods is presented elsewhere (Erasmus & Swanson, 2014). Briefly, commercial male turkeys (Hybrid Converter) were brooded in four littered (wood shavings) pens (2.4 m x 3.05 m) in groups of 24 from 1 day to 2 weeks of age. At 2 weeks, turkeys were randomly assigned to one of 16 littered pens (2.4 m x 3.05 m) in the same barn and housed in groups of four to six until 14 weeks.

Turkeys were tested three times (Period 1: 28 to 41 d, Period 2: 29 to 42 d, and Period 3: 56 to 68 d) in each of four tests of fear (tonic immobility (TI), open field (OF), voluntary approach (VA) and novel object (NO)). Only data from Period 1 and Period 3 were used here to assess inter-test correspondence at two ages (Period 1 and Period 3). Test periods will consequently be referred to as Age 1 (28 to 41 d; formerly Period 1) and Age 2 (56 to 68 d; formerly Period 3) throughout.

For TI and OF tests, birds were individually removed from their home pens and tested in another room. Voluntary approach and NO tests were conducted with groups of turkeys in their home pens to reduce confounds associated with environmental novelty, and because NO tests are used to assess welfare in groups of birds as part of welfare assessment programs (e.g. Welfare Quality®, 2009). The TI test lasted until the turkey righted itself, or for a maximum of 13 min. The number of inductions, number of vocalizations, latency to vocalize (s) and TI duration (s) were recorded. Turkeys were individually tested in the OF, consisting of a square arena (length: 2.74 m, width: 2.74 m, height: 1.83 m) with a concrete floor that was divided into 81 squares (each 0.30 m^2) for 10 min. OF test measures included the latency to ambulate (s), latency to vocalize (s), number of vocalizations, number of squares entered, number of steps taken and escape attempts. Turkeys were tested in VA and NO tests for 15 min. Four different objects were used to assess NO test responses to avoid habituation. Test measures that were recorded included latencies to move within two body lengths, one body length, approach and peck.

Based on results from Erasmus & Swanson (2014), test measures that had between-week (Period 1 vs. Period 3) test-retest correlation coefficients of $r_S \geq 0.5$ were retained for analysis of inter-test correspondence. The following test measures were thus included in the analyses: OF test latency to ambulate ($r_S = 0.66$), number of steps ($r_S = 0.58$), number of squares ($r_S = 0.59$), and number of defecations ($r_S = 0.50$); TI test latency to vocalize ($r_S = 0.51$) and number of vocalizations ($r_S = 0.59$); all VA test measures (latency to move within two body lengths: $r_S = 0.62$, latency to move within one body length: $r_S = 0.62$, latency to approach: $r_S = 0.87$, and

latency to peck: r_S = 0.67); and the latency to peck the object during the NO test (r_S = 0.61) (Erasmus & Swanson, 2014).

Two strategies were used to assess correspondence among fear tests. Test measures were correlated to examine pair-wise associations between measures from different tests. Secondly, the two turkeys from each pen showing the highest and lowest extremes in responses to fear test measures, respectively, were selected for further analysis to determine whether birds showing extremes in their behavioral responses in one test also differed in responses to other tests. These birds were classified as either a low responder (LR; fewest vocalizations, longest latency to ambulate, approach or peck) or a high responder (HR; most vocalizations, shortest latency to ambulate, approach or peck) in each test. In a study with pigs, the top and bottom 10% of high and low responding pigs were selected for further study. Here, we selected one HR and one LR from each pen. Each HR and LR therefore represented 17-25% of turkeys in the pen.

For each fear test, the test measure that was most reliable when test-retest reliability was assessed between weeks (Erasmus & Swanson, 2014) was used as the basis for choosing HR and LR for each test. For the OF test, the two turkeys from each pen with the shortest (open field high responder; OF-HR) and longest (open field low responder; OF-LR) latencies to ambulate were selected. For the TI test, turkeys were chosen based on the number of vocalizations (TI-HR had the most vocalizations and TI-LR had the fewest vocalizations). For the VA test, turkeys were selected based on the latency to approach the observer (VA-HR: fastest approach vs. VA-LR: slowest approach), and for the NO test turkeys were selected based on the latency to peck to NO (NO-HR: shortest peck latency vs. NO-LR: longest peck latency). Similar to methods of Spake et al. (2012), we verified that turkeys that were classified as HR and LR in each fear test differed significantly. Specifically, OF-HR and OF-LR differed significantly in the latency to

ambulate (Age 1: $U = 53.0$, $P < 0.0001$; Age 2: $U = 106.0$, $P < 0.0001$). Turkeys classified as TI-HR and TI-LR differed significantly in the number of vocalizations produced (Age 1: $U = 5.09$, $P < 0.0001$; Age 2: $U = 5.03$, $P < 0.0001$), VA-HR and VA-LR differed in the latency to approach (Age 1: $U = 109.0$, $P = 0.0002$; Age 2: $U = 332.0$, $P = 0.01$) and NO-HR and NO-LR differed in the latency to peck (Age 1: $U = 73.0$, $P < 0.0001$; Age 2: $U = 91.0$, $P < 0.0001$).

There were several pens in which none of the birds responded (all low responders). In these cases, data from only one randomly selected bird was used in further analyses and the bird was classified as LR. Turkeys classified as HR and LR in each test were not necessarily the same birds that were classified as HR or LR in other tests. Therefore, sample sizes varied for each test (OF-HR: $N = 9$, OF-LR: $N = 15$; TI-HR: $N = 14$, TI-LR: $N = 15$; VA-HR: $N = 12$, VA-LR: $N = 15$; NO-HR: $N = 10$, NO-LR: $N = 15$).

Statistical analysis

All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC). Pen ($N = 16$) was the experimental unit. For pair-wise correlations, data from each test measure were averaged across birds in each pen to obtain a single value for each test measure for each pen. Data were not normally distributed and transformations were ineffective. Therefore, test measures were correlated across all four tests of fear using the Spearman rank correlation coefficient. Correlation coefficients were defined as low: 0.2 – 0.4, moderate: 0.4 – 0.7, high: 0.7 – 0.9, and very high: 0.9 – 1.0 (Martin & Bateson, 1993). The Mann-Whitney U test (Wilcoxon rank sum test; SAS PROC NPAR1WAY) was used to assess differences in test responses between high and low responders. The exact Wilcoxon test was specified. The Wilcoxon rank

sum test was used to verify that behavior of turkeys that were classified as LR and HR in each test differed significantly.

RESULTS

Pair-wise correlations of test measures

Spearman rank correlation coefficients for pair-wise comparisons among test measures are presented in Table 3.1.

Age 1. There was a tendency for the number of steps taken and number of squares entered in the OF test to be moderately, positively correlated with the latency to peck the observer during the VA test. No other significant correlations were found.

Age 2. A greater number of significant correlations were observed between test measures at Age 2 compared to Age 1. The number of vocalizations during TI testing was moderately correlated with all OF test measures. The number of steps taken and number of squares entered in the OF tended to be correlated with the latency to peck the NO. There were no significant correlations between TI and NO, or TI and VA test measures.

Differences between high (HR) and low (LR) responders

Age 1. Turkeys classified as OF-LR and OF-HR did not differ in their responses to any other fear tests (Table 3.2). Similarly, TI-HR and TI-LR, and NO-HR and NO-LR did not differ in their responses to other tests. However, there was a tendency for VA-HR to take less time to ambulate and to enter more squares when tested in the OF test compared to VA-LR (Table 3.2).

Age 2. A greater number of differences were observed between HR and LR turkeys at Age 2 (Table 3.3). Turkeys classified as OF-HR tended to take less time to move within two

body lengths and approach the observer, and took significantly less time to move within one body length of the observer in the VA test compared to OF-LR, but OF-HR and OF-LR did not differ in TI or NO behavior. There was a tendency for TI-HR to defecate more when tested in the OF test compared to TI-LR. Turkeys classified as VA-HR tended to take less time to peck the NO during NO testing compared to VA-LR. Turkeys classified as NO-HR were more active in the OF test and took significantly less time to peck the observer during the VA test compared to NO-LR.

DISCUSSION

Two strategies were used to assess inter-test correspondence between four tests of fear for turkeys. Results from comparisons between high and low responders were in agreement with results from pair-wise correlations at both ages, indicating that these two strategies yielded similar results. Specifically, the tendency for a positive correlation between the latency to peck the observer during the VA test and the numbers of squares entered and steps taken during OF testing at Age 1 were in agreement with the tendency for VA-HR and VA-LR to differ in the latency to ambulate and number of squares entered when tested in the OF test. At Age 2, significant correlations were found between the TI and OF tests, and TI-HR and TI-LR tended to differ in the number of defecations produced in the OF test. Furthermore, the numbers of steps taken and squares entered in the OF test and latency to peck the NO tended to be correlated, and turkeys that took less time to peck the NO (NO-HR) took less time to ambulate, took more steps and defecated more in the OF than NO-LR at Age 2.

Interestingly, there were cases where HR and LR in one type of test differed in their responses in another test type, but the reverse was not observed. For example, at Age 1 VA-HR

and VA-LR tended to differ in the latency to ambulate in the OF test, but turkeys classified OF-HR and OF-LR did not differ in their responses to the VA test. Similarly, NO-HR and NO-LR differed in OF behavior at Age 2, but OF-HR and OF-LR did not differ in their NO responses. Because high and low responding individuals were selected separately for each test, HR and LR birds were not the same birds in one test as HR and LR birds in another test. In other words, there was some overlap between tests in the individual turkeys that were represented, but VA-HR turkeys for example were not all the same turkeys as OF-HR turkeys. The reason for differences in which turkeys were used to represent HR and LR is that individual turkeys' behavior varied within pens from test situation to test situation, such that the same two birds were not always the highest and lowest responders in the pen across all test situations. A possible explanation for the inconsistency in individual turkeys' behavior is that there may have been different floor (high numbers of birds showing little or no reaction) and ceiling (high numbers of birds showing the maximum response) effects depending on the test used. Indeed, a greater number of turkeys displayed the same behavior in the NO test (e.g. did not peck the NO) than in the OF test, making it impossible in some cases to separate high and low responders in the NO test. Increasing NO test duration may have lead to greater observed differences in behavioral responses, but further research is needed to understand floor and ceiling effects of various fear tests.

We previously reported that turkeys' behavior in fear tests changed over time and with repeated testing (Erasmus & Swanson, 2014). For example, activity levels in the OF increased over time (shorter latency to ambulate and greater numbers of steps taken), whereas latencies to approach and peck during VA and NO testing decreased over time (Erasmus & Swanson, 2014). Similarly, inter-test correspondence appears to change over time and with repeated testing. In

contrast to results from Age 1, there were more instances of inter-test correspondence and stronger relationships between different test situations when turkeys were tested at Age 2. Results from pair-wise comparisons and differences between HR and LR suggest that there may be correspondence between turkeys' responses in a VA test and an OF test between 4 and 6 weeks, but the relationship between these two tests is not very strong. Conversely, there appear to be strong relationships between turkeys' behavior in OF and NO tests, and OF and TI tests, and a weak relationship between VA and NO tests at 8 to 10 weeks of age.

A study assessing the validity of four tests of fear for quail also reported age-related differences in test correspondence: correlations between test measures were different when quail were tested at 42 to 49 days of age compared to tests at 61 to 65 days of age (Miller et al., 2006). Other researchers have only assessed inter-test correspondence at one age (e.g. Albentosa et al., 2003) or have taken the average of fear responses across different ages (e.g. Hocking et al., 2001). Our results suggest that not only does behavior in each fear test change over time (Erasmus & Swanson, 2014), but relationships between tests also change over time and with repeated testing. The differences in test correspondence between the two ages may have been affected by repeated testing of turkeys, and by the amount of human contact that turkeys received in this study. Unfortunately, the effects of repeated testing and habituation could not be separated from the effects of age. Additional research examining test correspondence in different groups of turkeys at different ages is necessary to fully understand how habituation and repeated testing may influence inter-test correspondence.

Correspondence between tests, or lack thereof, may be explained by similarities and differences in what the tests are actually measuring. Because fear is a subjective, complex emotional state that is difficult to quantify, test measures are by necessity directed at specific

components of animals' behavioral reactions that are quantifiable, such as the numbers of vocalizations produced and steps taken. Consequently, behavioral reactions, and thus test measures, may differ depending on the test used, making it challenging to interpret and examine associations among different types of fear tests. Alternatively, test measures may assess similar underlying behavioral patterns that are consistent within individuals and may therefore be more likely to correspond. For example, the general activity level of an animal, manifested as its locomotor behavior, may be a consistent characteristic of that animal, and therefore, it may be expected that a test measure in one test would correspond to a measure in another test if both assess locomotor behavior. It is possible that the correspondence that was observed among VA, NO and OF tests may be due to all three of these tests assessing components of turkeys' activity levels and motor patterns. Contrarily, test measures that assess vastly different behavioral patterns, such as vocalization vs. number of steps taken, may be less likely to correspond among tests.

Another explanation for a lack of test correspondence is that animals' responses to fear tests may be reflective of other motivations or behavioral patterns and not necessarily fear per se (Boissy, 1995). Some tests may measure exploratory motivation, social reinstatement motivation or locomotion/general activity level (e.g. open field test) whereas other tests may only measure fear of novelty (e.g. novel object test conducted in home pen). Indeed, the OF test not only measures fear of novelty, but it also subjects animals to social isolation (Forkman et al., 2007) and has been used to assess activity levels of turkeys (Huff et al., 2007).

Reviews specifically examining fear and fearfulness in poultry have concluded that there are associations among certain fear tests, which indicates that these fear tests are assessing the same underlying responses (R. B. Jones, 1996) and are valid (Forkman et al., 2006). A number of

studies with birds and other animals have shown that animals maintain their position in a fear hierarchy, indicating stability of the fear response across situations (R. B. Jones, 1996). Furthermore, birds in top tiers of battery cages had greater avoidance of a NO, longer TI durations and increased latencies to approach humans (reviewed in Forkman, 2007). However, more recent research has cast doubt upon the association among fear tests and the notion that different tests measure the same, single, underlying fear state. For example, chickens that received more human contact displayed shorter avoidance distances in an approaching human test, but did not differ in their responses to a novel object compared to chickens receiving less human contact (Graml et al., 2008). These results suggest that fear of humans and fear of novel objects are independent. Similarly, Miller et al. (2006) reported poor correspondence among four different tests of fear in quail (emergence test, predator test, novel object test, and novel food test), indicating that measures in one fear test do not correspond with measures recorded in other fear tests. When genetic selection for TI duration was used as a different strategy to examine the correspondence between fear tests, quail selected for long (LTI) and short (STI) TI durations did not differ in responses to novel objects, suggesting that TI duration is not associated with fear of novelty (Richard et al., 2008).

Our results reveal a complex relationship among TI, OF, NO and VA tests. When results from Age 1 are considered, our results are in agreement with others (Graml et al., 2008; Miller et al., 2006; and Richard et al., 2008) that there is poor correspondence among fear tests. However, there may be stronger associations among fear test measures as animals age. The lack of agreement in the literature and the contradictory findings regarding fear test correspondence may be due to differences in methodologies and strategies used. Alternatively, the contradictory findings and lack of agreement may indicate that tests are measuring specific responses to

specific components or stimuli present during testing, and are not capturing a larger, general fear response. Indeed, several researchers have stated that fear responses may depend largely on the context (Miller et al., 2006) and that fear is multidimensional (Richard et al., 2008; 2010).

In conclusion, there appears to be a weak association between turkeys' willingness to move toward an observer in their pen and their ambulatory behavior in an open field test between 4 and 6 weeks of age. There may be stronger associations among test situations between 8 and 10 weeks of age, but further research is needed to examine age-related differences in fear test correspondence independent of test habituation and repeated testing. Results did not validate that the OF, TI, VA and NO tests are measuring underlying fearfulness as a unitary construct, or that these tests are measuring the same types of fear responses. This study only assessed convergent validity (whether there is agreement among tests intended to measure similar behavioral responses) and not discriminant validity (whether tests disagree with other tests that are intended to measure different responses; Cozby, 1993). In addition to evaluating discriminant validity, additional research is needed to examine how external environmental conditions, genetics, and sex influence turkeys' OF, TI, VA and NO responses.

ACKNOWLEDGEMENTS

We thank Hybrid Turkeys for supplying the poults, and Dr. Darrin Karcher and Angelo Napolitano for their assistance with this project. We also thank Erika First, Caitlin Hummel, Sarah Jager, Cheyenne Ryzenga and Shannon Williams for assistance with data collection. M.E. was supported by a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship and this work was supported by Michigan State University AgBioResearch.

APPENDIX

Table 3.1. Spearman rank correlation coefficients (r_S) of open field (OF), tonic immobility (TI), voluntary approach (VA), and novel object (NO) test measures.

	NO test	VA test				TI test	
	Lat. peck	Lat. Two BL	Lat. one BL	Lat. approach	Lat. peck	Lat. voc.	Voc. no.
			Age 1 (N = 16)				
OF test							
Lat. amb.	-0.31	-0.31	0.14	0.06	-0.01	-0.03	0.06
Step no	-0.14	0.35	0.29	0.42	0.46 ^{T}	-0.36	0.32
Square no.	-0.1	0.35	0.28	0.37	0.43 ^{T}	-0.35	0.3
Def. no.	0.13	-0.2	-0.33	-0.29	-0.22	-0.09	0.04
TI test							
Lat. voc.	-0.02	0.07	0.1	0.14	0.006		
Voc. no.	0.04	-0.12	-0.02	-0.1	-0.01		
VA test							
Lat. two BL	-0.21						
Lat. one BL	-0.2						
Lat. approach	-0.23						
Lat. peck	-0.15						
			Age 2 (N = 16)				
OF test							
Lat. amb.	0.28	0.15	0.11	0.1	0.17	-0.42	0.50*
Step no	-0.50^{T}	-0.29	-0.27	-0.007	0.1	0.25	-0.51*
Square no.	-0.44 ^{T}	-0.37	-0.22	0.02	0.15	0.38	-0.56*
Def. no.	-0.36	-0.33	-0.17	0.01	0.15	0.34	-0.50*
TI test							
Lat. voc.	-0.04	-0.24	-0.17	-0.13	-0.19		
Voc. no.	0.05	0.18	0.19	0.02	-0.04		
VA test							
Lat. two BL	0.39						
Lat. one BL	0.26						
Lat. approach	0.04						
Lat. peck	0.22						

Note: Lat. = latency, BL = body length, amb = ambulate, voc. = vocalize.
$^{T}P < 0.10$, $^{*}P < 0.05$.

Table 3.2. Wilcoxon rank sum test differences in median (25-75% quartiles) values of open field (OF), tonic immobility (TI), voluntary approach (VA) and novel object (NO) test measures for turkeys classified as high (HR) or low (LR) responders in each test at Age 1 (4 to 6 weeks).

	OF test classification		TI test classification		VA test classification		NO test classification	
	OF-HR (N = 9)	OF-LR (N = 15)	TI-HR (N = 14)	TI-LR (N = 15)	VA-HR (N = 12)	VA-LR (N = 15)	NO-HR (N = 10)	NO-LR (N = 15)
OF test responses								
Lat. amb. (s)			600 (376-600)	600 (513-600)	600 (358-600)[a]	600 (600-600)[b]	600 (513-600)	600 (358-600)
Step no.			0 (0-8)	0 (0-2)	0 (0-63)	0 (0-2)	0 (0-14)	0 (0-18)
Square no.			0 (0-1)	0 (0-0)	0 (0-16)[a]	0 (0-0)[b]	0 (0-3)	0 (0-1)
Def. no.			3 (1-4)	2 (1-3)	2 (1-4)	1 (1-3)	2 (1-3)	2 (1-3)
TI test responses								
Voc. Lat. (s)	780 (482-780)	780 (320-780)			780 (303-780)	680 (137-680)	430 (132-780)	780 (140-780)
Voc. no.	0 (0-21)	0 (0-13)			0 (0-36)	2 (0-27)	2 (0-27)	0 (0-39)
VA test responses								
Lat. two BL (s)	37 (25-87)	88 (44-327)	74 (25-390)	74 (35-387)			203 (71-900)	74 (15-387)
Lat. one BL (s)	90 (27-688)	243 (89-773)	118 (29-461)	246 (77-688)			292 (114-900)	220 (21-688)
Lat. app. (s)	91 (44-698)	244 (116-776)	369 (91-900)	363 (103-858)			400 (124-900)	225 (62-858)
Lat. peck (s)	340 (53-900)	370 (221-777)	451 (172-900)	477 (157-900)			463 (221-900)	354 (227-900)

Note: amb. = ambulate, app. = approach, lat. = latency, def. = defecation, voc. = vocalization, BL = body length.

[a, b]Rows within test classifications lacking common superscripts differ ($P < 0.10$).

Table 3.2. (cont'd).

	OF test classification		TI test classification		VA test classification		NO test classification	
	OF-HR (N = 9)	OF-LR (N = 15)	TI-HR (N = 14)	TI-LR (N = 15)	VA-HR (N = 12)	VA-LR (N = 15)	NO-HR (N = 10)	NO-LR (N = 15)
NO test responses								
Lat. peck (s)	900 (339-900)	900 (283-900)	899 (115-900)	900 (353-900)	900 (416-900)	900 (44-900)		

Note: amb. = ambulate, app. = approach, lat. = latency, def. = defecation, voc. = vocalization, BL = body length.
[a, b]Rows within test classifications lacking common superscripts differ ($P < 0.10$).

Table 3.3. Wilcoxon rank sum test differences in median (25-75% quartiles) values of open field (OF), tonic immobility (TI), voluntary approach (VA) and novel object (NO) test measures for turkeys classified as high (HR) or low (LR) responders in each test at Age 2 (8 to 10 weeks).

	OF test classification		TI test classification		VA test classification		NO test classification	
	OF-HR (N = 13)	OF-LR (N = 17)	TI-HR (N = 13)	TI-LR (N = 16)	VA-HR (N = 15)	VA-LR (N = 15)	NO-HR (N = 13)	NO-LR (N = 16)
OF test responses								
Lat. amb. (s)			600 (117-600)	528 (178-600)	410 (103-600)	600 (256-600)	508 (74-600)[e]	600 (600-600)[f]
Step no.			7 (1-8)	6 (0-37)	5(2-37)	1 (0-10)	8(7-10)[e]	0 (0-3)[f]
Square no.			0 (0-1)	1 (0-5)	0 (0-5)	0 (0-1)	1 (0-2)[a]	0 (0-0)[b]
Def. no.			2 (1-3)[a]	2 (1-2)[b]	1 (1-2)	1 (1-2)	2 (1-3)[c]	1 (1-2)[d]
TI test responses								
Voc. Lat. (s)	780 (630-780	780 (262-780)			780 (184-780)	780 (780-780)	780 (310-780)	780 (780-780)
Voc. no.	0 (0-2)	0 (0-1)			0 (0-4)	0 (0-0)	0 (0-102)	0 (0-0)
VA test responses								
Lat. two BL (s)	17 (12-26)[a]	34 (23-43)[b]	23 (13-37	26 (15-59)			17 (11-26)[a]	29 (24-45)[b]
Lat. one BL (s)	26 (20-40)[c]	49 (34-79)[d]	36 (29-46)	40 (21-70)			29 (21-40)[a]	47 (37-61)[b]
Lat. app. (s)	37 (26-55)[a]	89 (38-116)[b]	45 (37-112)	53 (25-85)			38 (26-46)	56 (37-101)
Lat. peck (s)	55 (36-78)	91 (55-178)	83 (55-133)	65 (35-137)			51 (29-83)[a]	80 (58-167)[b]

Note: amb. = ambulate, app. = approach, lat. = latency, def. = defecation, voc. = vocalization, BL = body length.
[a-f] Rows within test classifications lacking common superscripts differ ([a, b] $P < 0.10$, [c, d] $P < 0.05$, [e, f] $P < 0.01$).

Table 3.3. (cont'd).

	OF test classification		TI test classification		VA test classification		NO test classification	
	OF-HR (N = 13)	OF-LR (N = 17)	TI-HR (N = 13)	TI-LR (N = 16)	VA-HR (N = 15)	VA-LR (N = 15)	NO-HR (N = 13)	NO-LR (N = 16)
NO test response								
Lat. peck (s)	46 (25-432)	407 (32-900)	80 (25-900)	863 (61-900)	38 (18-900)[a]	407 (83-900)[b]		

Note: amb. = ambulate, app. = approach, lat. = latency, def. = defecation, voc. = vocalization, BL = body length.
[a-f] Rows within test classifications lacking common superscripts differ ([a, b] $P < 0.10$, [c, d] $P < 0.05$, [e, f] $P < 0.01$).

REFERENCES

REFERENCES

Albentosa, M. J., Kjaer, J. B., & Nicol, C. J. (2003). Strain and age differences in behaviour, fear response and pecking tendency in laying hens. *British Poultry Science*, 44, 333-344.

Boissy, A. Fear and fearfulness in animals. (1995). *The Quarterly Review of Biology,* 70, 165-191.

Bourin, M., Petit-Demoulière, B., Dhonnchadha, B. N., & Hascöet, M. (2007). Animal models of anxiety in mice. *Fundamental & Clinical Pharmacology*, 21, 567–574.

Cockrem, J. F. (2007). Stress, corticosterone responses and avian personalities. *Journal of Ornithology*, 148 (Suppl 2), S169–S178.

Cozby, P. C. (1993). Methods in Behavioral Research. Mountain View, California: Mayfield.

Erasmus, M. A., & Swanson, J. C. (2014) Temperamental turkeys: reliability of behavioural responses to four tests of fear. *Applied Animal Behaviour Science*, 157: 100-108.

Forkman, B., Boissy, A., Maunier-Salaün, M. -C., Canali, E, & Jones, R. B. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior*, 92, 340-374.

Gosling, S. D. (2001). From mice to men: what can we learn about personality from animal research? *Psychological Bulletin*, 127, 45-86.

Graml., C., Waiblinger, S., & Niebuhr, K. (2008). Validation of tests for on-farm assessment of the hen-human relationship in non-cage systems. *Applied Animal Behaviour Science*, 111, 301-310.

Groothuis, T. G. G., & Carere, C. (2005). Avian personalities: characterization and epigenesis. *Neuroscience and Biobehavioral Reviews*, 29, 137-150.

Hazard, D., Leclaire, S. Couty, M. & Guémené, D. (2008). Genetic differences in coping strategies in response to prolonged and repeated restraint in Japanese quail divergently selected for long or short tonic immobility. *Hormones and Behavior*, 54, 645-653.

Hocking , P. M., Channing , C. E., Waddington, D. & Jones, R. B. (2001). Age-related changes in fear, sociality and pecking behaviours in two strains of laying hen. *British Poultry Science,* 42, 414–423.

Huff, G., Huff, W., Rath, N., Donoghue, A., Anthony, N., & Nestor, K. (2007). Differential effects of sex and genetics on behavior and stress response of turkeys. *Poultry Science*, 86, 1294-1303.

Jones, A. C., & Gosling, S. D. (2005). Temperament and personality in dogs (*Canis familiaris*): a review and evaluation of past research. *Applied Animal Behaviour Science*, 95, 1-53.

Jones, R. B. (1996). Fear and adaptability in poultry: insights, implications and imperatives. *World's Poultry Science Journal*, 52, 131-174.

Jones, B., & Boissy, A. (2011). Fear and other negative emotions. In M. C. Appleby, J. A. Mench, I. A. S. Olsson, & B. O. Hughes (Eds.), *Animal welfare* (2nd ed.). (pp. 81). Cambridge, MA, USA: CAB International North America.

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H. et al. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, 23, 925–935.

Marín, R. H., Martijena, I. D., & Arce, A. (1997). Effect of Diazepam and a β-Carboline on open-field and T-maze behaviors in 2-day-old chicks. *Pharmacology Biochemistry and Behavior*, 58, 915-921.

Martin, P., & Bateson, P. (1993). *Measuring behaviour: an introductory guide* (2nd ed.). New York, USA: Cambridge University Press.

Miller, K. A., Garner, J. P., & Mench, J. A. (2006). Is fearfulness a trait that can be measured with behavioral tests? A validation of four fear tests for Japanese quail. *Animal Behaviour*, 71, 1323-1334.

Mazurek, M., McGeeb, M., Crowe, M. A., Prendiville, D. J., Boivin, X., & Earley, B. (2011). Consistency and stability of behavioural fear responses of heifers to different fear-eliciting situations involving humans. *Applied Animal Behaviour Science*, 131, 21-28.

Neumann, I. D. Wegener, G., Homberg, J. R, Cohen, H., Slattery, D. A., Zohar, J., et al. (2011). Animal models of depression and anxiety: What do they tell us about human condition? *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 35, 1357-1375.

Petherick, J. C., Holroyd, R. G., Doogan, V. J., & Venus, B. K. (2002). Productivity, carcass and meat quality of lot-fed *Bos indicus* cross steers grouped according to temperament. *Australian Journal of Experimental Agriculture*, 42, 389-398.

Remignon, H., Mills, A. D., Guémené, D., Desrosiers, V., Garreau-Mills, M., Marche, M. et al. (1998). Meat quality traits and muscle characteristics in high or low fear lines of Japanese quails (*Coturnix japonica*) subjected to acute stress. *British Poultry Science*, 39, 372–378.

Richard, S., Wacrenier-Ceré, N., Hazard, D., Saint-Dizier, H., Arnould, C., & Faure, J. M. (2008). Behavioral and endocrine fear responses in Japanese quail upon presentation of a novel object in the home cage. *Behavioural Processes*, 77, 313-319.

Richard, S., Land, N., Saint-Dizier, H., Leterrier, C., & Faure, J. M. (2010). Human handling and presentation of a novel object evoke independent dimensions of fear in Japanese quail. *Behavioural Processes*, 85, 18-23.

Spake, J. R., Gray, K. A., Cassady, J. P. (2012). Relationship between backtest and coping styles in pigs. *Applied Animal Behaviour Science*, 140, 146-153.

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

Suarez, S.D. & Gallup, G. G. (1983). Social reinstatement and open field testing in chickens. *Animal Learning & Behavior*, 11, 119-126.

Voisinet, B. D., Grandin, T., O'Connor, S. F., Tatuma, J. D., & Deesing, M. J. (1997). *Bos Indicus*-cross feedlot cattle with excitable temperaments have tougher meat and a higher incidence of borderline dark cutters. *Meat Science*, 4, 367-377.

Welfare Quality®. 2009. Welfare Quality® Assessment Protocol for Poultry. Welfare Quality Consortium, Lelystad, Netherlands.

CHAPTER 4: FEATHER PECKING AND OPEN FIELD BEHAVIOR OF TURKEYS

ABSTRACT

Feather pecking among commercial turkeys is an economic and welfare concern, but there has been scant research on this topic. Research with chickens has provided evidence that feather pecking may be associated with differences in temperament characteristics, such as fearfulness. The objectives of this research were to 1) examine fear responses using the open field (OF) test in males and females of a commercial (COMM) and randombred (RB) turkey strain, 2) compare feather pecking between sexes and strains, and 3) ascertain whether behavioral and physiological OF responses are associated with feather pecking. Turkeys were housed in same-sex, same-strain groups of 14-20 in 16 pens. Turkeys were individually tested three times in an OF test (n = 16 per pen). Males were tested at 1 (Test 1), 4 (Test 2), and 11 (Test 3) wk; females were tested at 1 (Test 1), 5 (Test 2) and 12 (Test 3) wk. Pre-and post-test plasma corticosterone levels (n = 6 per pen) were assessed during Test 3. Prior to Test 3, behavioral observations were conducted to identify birds that developed feather pecking so that OF responses could be compared between birds that pecked and birds that did not. Results revealed some sex differences in OF behavior between COMM males and females in Tests 2 and 3 and between RB males and females in Test 3. Strain differences in OF behavior were found in all tests. Corticosterone levels following testing were higher in RB females compared to COMM females. There were no differences in the frequency of feather pecking, but feather damage at 13 wk was worse in COMM compared to RB birds. No significant relationships were found between feather pecking and behavioral or physiological OF responses. Overall, these results indicate some sex and strain differences in OF behavior, but the nature of the differences depended on when birds were tested. Our results do not support the hypothesis that feather

pecking is associated with differences in temperament measured using the OF test, and these findings provide further evidence of the complex, multifactorial nature of feather pecking of turkeys.

INTRODUCTION

Global demand for food continues to increase, and the projected increase in the world's population will double the current need for food (Federation of Animal Science Societies, 2012). Therefore, there is pressure to develop faster growing, higher yielding and more efficient production animals. Concurrently, public interest in animal welfare continues to grow, and the welfare state of an animal can have significant and direct impacts on food quality (Blokhuis et al., 2008). Injurious pecking, including head pecking, feather pecking and cannabalism, is an important factor affecting food quality and animal welfare among intensively housed turkeys (Hocking et al., 1993). Indeed, injurious pecking results in culling, mortality, and carcass damage. However, research examining injurious pecking among turkeys is scarce, unlike research regarding injurious pecking of chickens. In a survey of commercial turkey producers in Canada, Erasmus (2009) reported that 13-14% of culls on farms were due to injuries. More recently, injuries associated with severe pecking have been reported to account for 58% of culls and mortalities at a commercial turkey facility (Duggan et al., 2014). The increasing demand for animal products and more efficient animal production increases concerns regarding animal welfare, making research examining injurious pecking among turkeys paramount.

In their review of injurious pecking of turkeys, Dalton et al. (2013) state that severe feather pecking and canibalism are "multi-factorial products of genetics, environment, and nutrition". However, not all birds develop feather pecking. Indeed, individual differences are

known to exist between laying hens that develop feather pecking and hens that do not (Rodenburg et al., 2008). For example, several studies have confirmed the relationship between temperament (coping style and fearfulness) and feather pecking in two lines of laying hens selected for high (HFP) and low (LFP) feather pecking, respectively (e.g. Korte et al., 1997; Van Hierden et al, 2002; Rodenburg et al., 2004; Jensen et al., 2005).

Coping style is defined as a consistent set of physiological and behavioral responses to stressors (Koolhaas et al., 1999). Proactive coping styles are characterized by an active behavioral response (fight or flight) and low hypothalamo-pituitary-adrenal (HPA) axis reactivity (low corticosterone levels) whereas a reactive coping style is characterized by reduced activity (immobility) and high HPA axis reactivity (Koolhaas et al., 1999). As stated by de Haas et al. (2010), chickens from the HFP line exhibit characteristics consistent with proactive coping styles, whereas chickens from the LFP line exhibit characteristics consistent with reactive coping styles (Korte et al., 1997; Van Hierden et al, 2002; Rodenburg et al., 2004; Jensen et al., 2005). Specifically, HFP chickens had lower basal and post-test corticosterone levels following a manual restraint test compared to LFP chickens (Korte et al., 1997; Van Hierden et al., 2002; reviewed in Rodenburg et al., 2008). There is also evidence that chickens that differ in feather pecking performance differ in fearfulness, which some scientists regard as being part of coping styles (Cockrem et al., 2007), whereas others regard fearfulness as an altogether separate dimension of temperament (Jones and Boissy, 2011; Koolhaas et al., 1999).

Fearfulness is often tested using the open field test (Forkman et al., 2007). Chickens from the LFP line were less fearful when tested in an open field test (Jones et al., 1995). Moreover, open field behavior of chicks was associated with their feather pecking behavior as adults in the F2 generation of a cross between the HFP and LFP lines (Rodenburg et al., 2004). In a recent

study, anxiety (tested using a social isolation test) at 1 wk was found to be associated with feather damage at 5 wk (de Haas et al., 2014). Therefore, individual differences in temperament may be important indicators of, or contributors to, the likelihood of developing feather pecking, but this has not been examined in turkeys.

Our study examined the relationship between fearfulness and feather pecking in turkeys of two genetic strains. Specifically, we examined whether fearfulness (tested using the open field test) and feather pecking differ between randombred (RB) and commercial (COMM) turkey strains and between sexes within strains. We tested the hypothesis that behavioral and physiological responses to open field testing differ between turkeys that develop feather pecking and turkeys that do not.

MATERIALS AND METHODS

All procedures were approved by the Institutional Animal Care and Use Committee of Michigan State University.

Animals and housing

A total of 80 male and 80 female commercial (COMM) turkeys (Hybrid Converter strain, Hybrid Turkeys, Kitchener, On., Canada) and 62 male and 62 female randombred (RB) turkeys (RBC2 line established in 1966 (Nestor et al., 1969) and maintained at the The Ohio Agricultural Research and Development Center (OARDC) of The Ohio State University, Wooster, OH) were used in this study. In total, 62 male and female RB turkeys were obtained. The RB turkeys were incubated and hatched at the OARDC of the Ohio State University. COMM poults were obtained from a commercial hatchery and placed 1 d prior to placement of RB poults. Poults were

brooded in same-sex, same-strain groups in eight littered (wood shavings) pens (2.40 m x 3.05 m) in the same barn from 1 d to 2 wk of age. At 2 wk, birds were distributed throughout the barn to one of 16 littered pens (2.4 m x 3.05 m) and housed in same-sex, same-strain groups (14 to 20 birds per pen) where they remained until study completion at 14 wk. sexes and strains were randomly assigned to pens such that the location of males and females of each sex was random within the barn. At 7 wk, some COMM turkeys were removed from their pens for use in another study and the number of COMM turkeys in each pen was reduced to, and maintained at 16. Therefore, group sizes were similar across sexes and strains from 7 wk onward (14-16 birds/pen).

Temperature and diet were the same as described in Erasmus and Swanson (2014). A photoperiod of 24L: 0D was applied for the first 7 d, and then the light period was reduced by 1 hr/d for 7 d. A photoperiod of 16L: 8D was maintained from 14 d - 14 wk. Lights came on at 06:00 h and were turned off at 22:00 h. Light intensity was maintained at 20 lx.

At 1 wk, turkeys were individually identified using colored (green, black, purple, and orange) non-toxic livestock marker (Prima Tech Marking Stick, Neogen Corporation, Lansing, MI, USA) which was applied to different areas and in different color combinations for each bird in the pen (e.g. black and/or green on the shoulders, left or right wing) and was visible on an overhead camera. Marker was reapplied biweekly to ensure markings remained visible. Turkeys were weighed prior to OF testing at 23 and 73 d of age (one day older for RB birds).

Open field testing procedures

We previously reported that the open field (OF) test is reliable for assessing fear responses of commercial male turkeys (Erasmus and Swanson, 2014). In order to examine

whether OF test responses at different ages were associated with feather pecking behavior, 14-16 turkeys from each pen of male and female COMM and RB turkeys were tested in the OF test at three ages (three tests). Behavior in the OF was recorded during Test 1 (1 wk of age for males and females), Test 2 (males: 4 wk, females: 5 wk), and Test 3 (males: 11 wk, females: 12 wk). During the first test week, testing occurred over two consecutive days for males, followed by two consecutive days for females. During all other test weeks, testing occurred over four consecutive days so that a total of 24 birds could be tested each day. Each bird was only tested once during each test week. Test order was randomized for each day of testing so that COMM and RB birds from each sex were tested in random order. All testing took place between 07:45 h and 17:45 h.

Testing procedures were the same as described in Erasmus and Swanson (2014). Briefly, turkeys were individually tested in an OF test arena (2.74 m (length) x 2.74 m (width) x 1.83 m (height)). Each turkey was placed in the center of the arena for 10 min and behaviour was recorded using two overhead cameras (VIXIA HF M41, Canon USA, Inc. Melville, NY, USA) from which behavioral data (latency to ambulate, number of steps taken, number of squares entered, latency to vocalize, number of vocalizations, and number of defecations) were collected. A vocalization was defined as a sound emitted by the bird being observed. Vocalizations were distinguished from one another by pauses (periods of no sound) between vocalizations and each vocalization was counted separately. Some males gobbled during OF test 3. A sequence of gobbling sounds was counted as one vocalization if there was no pause between sounds.

Feather pecking behavior

Behavior of turkeys was video recorded on three days to examine feather pecking interactions. Video was recorded from 07:30 h to 20:30 h at 74, 75 and 81 d of age for COMM

males (one day older for RB males) and 74, 81 and 82 d for COMM females (one day older for RB females). This pattern of video recording was used so that behavior could be examined within three days prior to the last OF test, and so that there were two days of overlapping video between males and females. Four pens were recorded using one type of camera (2MB-70IR42L210 Jaguar Series infrared CCTV camera, 2M® Technology, Grand Prairie, TX, USA) and the other 12 pens were recorded using a similar camera but from a different manufacturer (Clinton Electronics CE-VF540 CCTV camera, Clinton Electronics Corp., IL, USA).

A pecking interaction was defined as "repeated pecks to the same individual" (Huber-Eicher and Sebö, 2001). The interaction was considered to have ended if there were no further pecks within a 5 s period. The frequency of feather pecking interactions was calculated for each pen as the number of feather pecking interactions per individual per day.

Individuals performing repeated (three or more pecks in succession) gentle or severe feather pecking (PECK) were identified from video recordings by recording feather pecking every 10 min. Briefly, the video was stopped every 10 min and watched for 10 s before and 10 s after the 10 min time point to verify that feather pecking behavior was occurring and to score feather pecking as gentle or severe. Because feather pecking occurs in bouts, feather pecking behavior was recorded as having occurred if 1) it was observed at exactly the 10 min time point, and 2) if the bird performing the pecking was observed to peck immediately before and immediately after the 10 min time point, pausing for 5 s or less around the 10 min time point without performing any other behavior. Sometimes birds remove and eat feathers from other birds and some birds may peck at a slower rate than others. Therefore, scoring feather pecking behavior within 5 s surrounding the 10 min time point enables feather pecking bouts to be more

fully represented because the bird has not changed behavior and can be considered to still be engaged in feather pecking behavior.

Gentle and severe feather pecking were defined according to its definition for chickens (Savory, 1995). Briefly, gentle feather pecking was defined as pecking that did not elicit a response from the recipient. Severe feather pecking included forceful pecking, pulling and/or removal of feathers resulting in feather damage and causing the recipient to react or move away. Turkeys were classified as PECK only if they were observed to feather peck on two or more occasions to reduce risks of misclassifying birds (Jensen et al., 2005). Turkeys that were not observed to perform feather pecking were classified as NPECK. For each feather pecking interaction, the behavior of the recipient bird at the time of feather pecking was also noted. Behavioral categories of the recipient bird included resource (the recipient was eating or drinking), preen, dustbathe, or other (sitting, standing, walking and environmental pecking).

The frequency of severe feather pecking interactions and the frequency of interactions occurring for each behavioral classification of the recipient bird were calculated as the percentage of interactions that were observed in each pen divided by the total number of feather pecking observations, and then divided by the number of birds in the pen. The number of birds in each pen was included in calculations to account for differences in feather pecking due to group size. Percentages are reported as the mean of the 4 pens for males and females of each strain.

Reliability of feather pecking video analysis was $r \geq 0.8$ between observers. The same observer further examined all feather pecking data that were collected and verified identities of PECK birds, gentle vs. severe pecking and behavior of the recipient during pecking.

Turkeys were examined daily for pecking wounds. Turkeys with wounds that had penetrated the skin were treated with Rooster Booster® Pick-no-more lotion (TDL Industries, Inc. Fallon,

NV, USA) to prevent further pecking. If the pecking injury worsened, the turkey was moved to a hospital pen to prevent further injury. In severe cases of head pecking, turkeys were culled if necessary.

Turkeys were scored for feather pecking damage by scoring feather condition at 13 wk using the method of Bilčík and Keeling (1999) to score the back, neck, tail, and primary left and right wing feathers. A score of 0 represented intact feathers and little to no damage, whereas a score of 5 indicated complete or almost complete feather loss.

Plasma corticosterone levels

Blood samples were collected from six birds in each pen (n = 24 birds/sex-strain combination) during the second and third test periods, but only samples from Test 3 were assayed. Blood samples were collected one day prior to OF testing in order to obtain pre-test corticosterone levels. Samples were collected 15 min (± 10 min) after the time at which testing was set to start the following day. Blood samples were then collected the following day 15 min after OF testing commenced in order to examine the magnitude of corticosterone increase in response to testing. Birds were gently restrained on their side with one person holding the bird's legs and wing and the other person collecting blood from the brachial vein using a 25 gauge needle. No more than 2.2 ml of blood was collected each time. Blood was placed into EDTA-coated Vacutainer tubes (Fisher Scientific, NH, USA) and kept on ice until blood could be centrifuged (1000 rpm). After centrifugation, plasma was collected and stored at -20 °C until assayed. The same competitive enzyme immunoassay kit (Corticosterone HS, IDS Inc., Fountain Hills, AZ) used by Huff et al. (2007) was used to determine plasma corticosterone levels.

Statistical analysis

All data were analyzed using SAS 9.4 (SAS Institute Inc., Cary, NC.). Behavioral OF test responses were not normally distributed. Test-retest reliability of OF responses was analyzed using Spearman rank correlation. Correlation coefficients were defined according to the definition of Martin and Bateson (1993) (low: 0.2 – 0.4, moderate: 0.4 – 0.7, high: 0.7 – 0.9, and very high: 0.9 – 1.0). Sex and strain differences in behavioral responses in the OF test that consisted of counts (e.g. number of vocalizations) were analyzed using the GLIMMIX procedure for a negative binomial distribution with the log link function. Data were analyzed separately for each OF test period. The main effects of sex, strain and their interaction were examined. Differences in latency measures (e.g. latency to ambulate) between sexes within strains, and between males or females of different strains were tested using the LIFETEST procedure for each test period. The variable sex or strain was specified in the STRATA statement when sex or strain differences, respectively, were examined.

The MIXED procedure was used to compare the frequency of feather pecking interactions, the percentage of birds that were classified as PECK, and the percentage of feather pecking interactions occurring for each behavioral category of the recipient bird. The main effects of sex, strain and their interaction were included in the model. The percentage of birds in each pen that were classified as PECK was log(+ 1)-transformed, and the mean percentage of birds that performed feather pecking was cosine-transformed to meet normality assumptions. Raw means are reported.

Feather scores were compared between sexes and strains using the GLIMMIX procedure for a negative binomial distribution with the log link function. The main effects of sex, strain and their interaction were included in the model.

Differences in OF responses consisting of counts were compared between PECK and NPECK turkeys within males and females of each strain using the GLIMMIX procedure for a negative binomial distribution with the log link function. Separate analyses were conducted for each OF test period to determine whether OF behavior at any of the ages tested was associated with feather pecking behavior. The main effects of sex, strain, classification (PECK or NPECK) and their three-way interaction were tested. Body weight was included as a covariate for OF Test 2 and OF Test 3 models (body weights had not been recorded for OF Test 1). Latency measures were compared between PECK and NPECK birds using the LIFETEST procedure. Pen was specified in the STRATA statement and classification was specified as the GROUP variable. Body weight was included as a covariate in the TEST statement for OF Test 2 and Test 3.

Differences in pre- and post-test corticosterone levels, and in the change in corticosterone level due to testing (post-test level – pre-test level) were compared between sexes and strains using the MIXED procedure. The main effects of sex, strain and their interaction were tested. Pen was included as a random effect. Results are reported as raw means. Similarly, differences in pre- and post-test corticosterone levels were compared between PECK and NPECK birds using the MIXED procedure. The main effects of sex, strain, pecking classification and the interaction between sex, strain and pecking classification were included in the model. Pen was included as a random effect. Pre- and post-test corticosterone levels were log-transformed to achieve normality. The change in corticosterone level was transformed using log + 1 to achieve normality.

All post-hoc analyses were conducted using Tukey's test for multiple comparisons. P values < 0.05 were considered significant.

RESULTS

At 23 d of age, body weights (mean ± SE) were 0.9 ± 0.02 kg (COMM males), 0.6 ± 0.007 kg (RB males), 0.8 ± 0.02 kg (COMM females), and 0.5 ± 0.006 kg (RB females). Body weights were 8.4 ± 0.2 kg (COMM males), 4.4 ± 0.07 kg (RB males), 6.4 ± 0.2 kg (COMM females) and 3.2 ± 0.06 kg (RB females) at 73 d of age. At 7 wk, one RB male was moved to a hospital pen to prevent further injury from pecking. Between 11 and 13 wk, 2 RB males from the same pen and 3 COMM females (2 from the same pen) were moved to the hospital pen. Two RB males from another pen were culled at 10 and 13 wk, respectively, due to head pecking injuries.

Due to problems with video color, identities of PECK birds could not reliably be determined on 1 d of video recording for one RB female pen, one COMM male pen, and one COMM female pen, but the number of pecking interactions and behavior of recipients could still be recorded.

Test-retest reliability of OF behavior

Spearman rank correlation coefficients for correlations between OF tests are presented in Table 4.1. In general, a greater number of significant correlations were found, and correlation coefficients were larger between Test 2 and Test 3 than between Tests 1 and 2, or between Tests 1 and 3.

For COMM males, the only significant correlation that occurred between Test 1 and Test 2 was a tendency for the number of squares entered to be positively correlated. Significant, moderate correlations were found between Test 1 and Test 3 for the latency to ambulate and number of steps taken, and there was a tendency for a low correlation for the latency to vocalize and the number of squares entered. Significant, moderate correlations were found between Test 2

and Test 3 for all OF test responses except for the latency to ambulate and the number of squares entered, which had low, positive correlation coefficients.

There was a tendency for a low, positive correlation between Tests 1 and 2 for the number of defecations of RB males, but no other significant correlations were found between Tests 1 and 2, or Tests 1 and 3. There were significant, moderate correlations between Tests 2 and 3 for the latency to ambulate, number of steps and number of squares, and low correlations for the latency to vocalize and numbers of vocalizations and defecations.

For COMM females, the number of squares entered was significantly, moderately correlated between Tests 1 and 2, and there was a low correlation for the latency to ambulate. Additionally, the numbers of steps and vocalizations tended to be positively correlated between Tests 1 and 2. There were moderate correlations between Tests 1 and 3 for the latency to ambulate and number of defecations, and low correlations were found for the number of squares entered. The latency to vocalize tended to be correlated between Tests 1 and 3. Significant, moderate correlations were found between Test 2 and Test 3 for all OF test responses.

Significant correlations were found between all test periods for almost all behavioral OF responses of RB females. For Test 1 vs. 2, there were moderate correlations for the latency to ambulate and number of defecations, and low correlations for all other OF responses. Significant moderate correlations were found between Test 1 and Test 3 for all OF responses. There were significant, high correlations between Tests 2 and 3 for the numbers of steps and squares, and moderate correlations for all other OF responses.

Sex and strain differences in OF test responses

Behavioral OF responses for all 3 OF test periods are presented in Table 4.2.

Test 1. There was a significant sex-strain interaction for the number of steps taken and number of squares entered. Post-hoc comparisons revealed that the number of steps taken and squares entered were significantly higher for female RB turkeys compared to male RB turkeys, and there was a tendency for the number of vocalizations to differ between RB males and females ($t_{156} = 2.5$, $P = 0.07$). Both the latency to ambulate and the latency to vocalize were significantly shorter for RB females compared to COMM males and females.

Test 2. The numbers of steps taken and squares entered did not differ between sexes or strains. However, the main effect of sex affected the numbers of vocalizations and defecations. Post-hoc comparisons revealed that the numbers of vocalizations were higher for females compared to males, and COMM and RB females defecated more compared to COMM males. COMM males took longer to ambulate compared to COMM and RB females, and COMM males took significantly longer to vocalize compared to females of both strains and RB males.

Test 3. No sex or strain differences were found in the number of steps taken, but RB females entered more squares compared to COMM males. The main effect of sex significantly affected the number of vocalizations, but post-hoc comparisons only revealed a tendency for vocalizations to differ between COMM males and females ($t_{193} = 2.4$, $P = 0.08$). Both sex and strain significantly affected the number of defecations. COMM males defecated less compared to females of both strains and RB males, and RB males defecated less compared to RB females. COMM males took longer than all other turkeys to ambulate. COMM and RB females took less time to vocalize compared to COMM and RB males, and COMM females took less time to vocalize compared to RB females.

Corticosterone level. Post-test (Test 3) corticosterone levels were significantly higher compared to pre-test levels for turkeys of either sex within both strains (Fig. 4.1). Pre-test

corticosterone levels tended to differ between female COMM and female RB turkeys (t_{72} = -2.4, P = 0.08), and post-test corticosterone levels differed significantly between COMM and RB females (Fig. 4.1). The change in corticosterone levels from pre- to post-test differed significantly between COMM and RB females (Fig. 4.1), and tended to differ between RB males and females (t_{72} = 0.4, P = 0.06).

Sex and strain differences in feather pecking

Descriptive statistics of feather pecking behavior are presented in Table 4.3.

Feather pecking behavior. The main effect of sex tended to affect the lsmean (± SE) frequency of feather pecking (males: 0.56 ± 0.06 interactions per bird per day, females: 0.75 ± 0.06, F_{12} = 4.21, P = 0.06). The main effects of sex and strain significantly affected, and the interaction between sex and strain tended to affect, the percentage of birds that were observed to perform feather pecking. Post-hoc analyses revealed that a significantly higher percentage of RB females were observed to perform feather pecking compared to RB males and COMM males and females. Sex significantly affected the percentage of feather pecking interactions that were classified as severe. However, post-hoc comparisons did not reveal differences in the percentage of feather pecking interactions or the percentage of feather pecking interactions that were classified as severe, between sexes within strains, or between males or females of different strains.

Behavior of the recipient when feather pecking occurred. The percentage of feather pecking interactions/bird occurring at a resource differed between sexes (males: 0.23 ± 0.19, females: 1.56 ± 0.19, F_{12} = 26.1, P = 0.0003) and between males of different strains (Table 4.3).

The percentage of feather pecking interactions/bird occurring when the recipient was dustbathing differed significantly between strains (COMM: 0.69 ± 0.22, RB: 0.73 ± 0.22, $F_{12} = 17.8$, $P = 0.001$). Post-hoc comparisons revealed that the percentage of interactions/bird occurring during dustbathing tended to be higher for RB females compared to COMM females ($t_{12} = -2.61$, $P = 0.09$), and COMM males ($t_{12} = 2.90$, $P = 0.06$), and was significantly higher for RB males compared to COMM males, and COMM females. The percentage of feather pecking interactions/bird that occurred when the recipient was preening did not differ between sexes, strains or sexes within strains. The main effects of sex significantly affected, and the main effect of strain tended to affect, the percentage of feather pecking interactions that occurred when the recipient was engaged in "other" behavior. Post-hoc analyses indicated that COMM females performed more feather pecking/bird when the recipient was engaged in "other" behavior compared to RB females and RB males.

Feather scores

Feather scores are presented in Table 4.4. Only 5 turkeys had feather scores > 0 (0 indicated little or no feather damage) in the back region. Therefore, data from back feather scores could not be compared between sexes and strains.

Tail feather scores tended to differ between sexes ($F_{240} = 3.4$, $P = 0.06$), and differed significantly between strains. Tail feather scores were significantly higher (indicating worse feather condition) for female COMM compared to female RB turkeys, and for male COMM compared to female RB turkeys. Neck feather scores differed significantly between strains and between sexes of different strains, but not between sexes of the same strain (Table 4.4). Neck

feather scores of male and female COMM turkeys were higher compared to male and female RB turkeys.

Feather pecking and open field responses

Open field test responses of PECK and NPECK birds of males and females of both strains are presented in Table 4.5. In Test 1, the latency to vocalize tended to differ between COMM female PECK and NPECK birds (log-rank = 3.6, $P = 0.06$; Wilcoxon = 2.1, $P = 0.1$), and between male RB PECK and NPECK birds (log-rank = 3.5, $P = 0.06$; Wilcoxon = 3.2, $P = 0.07$). No differences in OF Test 1 responses were found between male COMM or female RB PECK and NPECK birds. There were no differences in any OF responses in Test 2 or Test 3 between male or female PECK and NPECK birds of either strain. Similarly, pre- and post-test corticosterone levels did not differ between PECK and NPECK birds (Fig. 4.2).

DISCUSSION

Before any test can be used to examine differences in temperament or coping style, the reliability (test-retest repeatability; assessed in Erasmus and Swanson, 2014) and validity (inter-test correspondence, assessed in Chapter 3) of the test need to be ascertained. Previously, we reported that responses of male COMM turkeys were repeatable between 4 and 10 wk of age (Erasmus and Swanson, 2014), and that although fear responses appear to be context dependent, there may be an association between OF and tonic immobility behavior at 8 to 10 wk of age (Chapter 3).

Consistent with our previous results (Erasmus and Swanson, 2014), the majority of behavioral OF responses of COMM male turkeys were repeatable between 4 and 11 wk. The

majority of responses of COMM females, and RB males and females also had moderate correlation coefficients between 4 and 11 wk (males) and 5 and 12 wk (females). However, repeatability of OF behavior was low between 1 and 4 (males) or 5 (females) wk, indicating that OF behavior changes between 1 and 5 wk, and with repeated testing. There does not appear to be any published research examining changes in turkeys' fear responses with age, but early research with chickens suggests that behavior associated with predator avoidance increases as birds age (Suarez and Gallup, 1983). Therefore, OF and other fear-related behavior may be more consistent at later ages due to the development of this behavior over time.

Interestingly, in our previous study with COMM males the latency to ambulate had the highest correlation coefficient ($r_S = 0.66$) and the latency to vocalize had the lowest correlation coefficient ($r_S = 0.17$; Erasmus and Swanson, 2014). However, results reported here revealed that the latency to vocalize ($r_S = 0.47$) had a higher correlation coefficient than the latency to ambulate ($r_S = 0.35$). The difference in test-retest repeatability may be due to the ages at which birds were tested (and body weight differences), or differences in environmental conditions. Turkeys in the first study (Erasmus and Swanson, 2014) were reared in smaller groups (4-6 birds per pen) and under a much lower light intensity (5 lx) compared to turkeys in this study, where both the group size was larger (14-16 birds per pen) and the light intensity was higher (20 lx). Furthermore, turkeys in the present study were tested later and males had higher body weights (6.2 ± 0.08 kg (Erasmus and Swanson, 2014) vs. 8.5 ± 0.13 kg), which may have reduced their physical ability to move around the OF testing arena. In contrast, COMM females and RB turkeys had much lower body weights compared to COMM males, and therefore their ambulatory ability may not have been affected.

Open field behavior may not only vary due to environmental and management factors, but sex and genetic strain may also be important factors affecting behavior. Due to genetic selection for increased feed efficiency and increased production, the domestic turkey has changed dramatically in the past 30 years. Therefore, we wanted to examine whether behavioral and physiological responses in the OF differed between COMM and RB turkeys, and between males and females of the two strains. Differences were found for some responses at each age of testing, but differences were not the same across test periods. There were fewer strain differences than there were sex differences. In general, COMM males defecated less and ambulated and vocalized later than COMM females when tested at 4 and 5 wk, and 11 and 12 wk, respectively. Similarly, RB males defecated less and vocalized later than RB females when tested at 11 and 12 wk, respectively. Strain differences were limited to differences in vocalization latency (males: Test 2, females: Test 1 and Test 3), ambulation latency (males: Test 3, females: Test 1), and number of defecations (males: Test 3). Results reveal that sex and strain differences in OF behavior appear to depend on the age at which birds are tested, and may be affected by repeated testing. Sex differences may have been confounded by the age at testing because females were tested one week later in Tests 2 and 3 compared to males. Males and females were tested one week apart because it was not possible to test all birds in the same week. Furthermore, blood samples were collected before and after testing and in order to prevent reducing the sample size for blood samples, we elected to test all birds from the same sex in the same week rather than confound corticosterone levels and age. Further research is needed to examine age-related effects on fear responses and how male and female turkeys differ in the rate of development of fear responses.

All turkeys experienced significantly elevated post-test corticosterone levels compared to pre-test levels, indicating activation of the hypothalamic-pituitary-adrenal (HPA) axis. The HPA axis is activated when animals experience fear, and corticosterone levels are higher in birds when they are exposed to a threatening situation (Cockrem, 2007). Research with chickens and quail have provided evidence that corticosterone levels are higher in birds that show behavioral responses consistent with increased fear in tonic immobility and OF tests (reviewed in Cockrem, 2007). Unlike in the studies with chickens and quail, we did not observe a clear relationship between behavioral and physiological responses of turkeys to OF testing. Corticosterone levels following testing were higher in RB females compared to COMM females, but behavioral responses were not significantly different. If corticosterone levels were associated with fear, and if lower activity levels are indicative of greater levels fear, birds with lower activity levels may be expected to have higher corticosterone levels following OF testing. In fact, the opposite occurred: the number of steps taken and squares entered were numerically, although not statistically, higher for RB vs. COMM females, which may suggest lower fear in RB females. Alternatively, corticosterone levels may reflect differences in activity levels and not fear per se. Furthermore, inactivity in the OF may not necessarily be indicative of a high level of fear because some birds may be inactive without being fearful. Another possible explanation for the discrepancy between OF behavior and corticosterone level is that the physiological stress responses that were observed were due to capture and handling by observers, rather than conditions associated with the OF test environment. Capture and handling by human observers may be more threatening than the OF test itself, thereby masking the physiological response to OF testing. In agreement with this explanation, Hazard et al. (2008) reported that quail selected for high and low durations of tonic immobility experienced greater increases in corticosterone

levels at the beginning of the tonic immobility test when they were being restrained by an observer, but no relationship was found between tonic immobility duration (measure of fear) and corticosterone level. Another explanation is that behavioral differences in fear behavior are not consistently reflected by differences in HPA axis activity (Armario et al., 2012). In their review of emotional stressors and hormones, Armario et al. (2012) state that "the relationship between HPA activation and anxiety is far from being adequately characterized". Further research examining the relationship between OF behavior and corticosterone is needed to understand whether corticosterone levels are related to differences in activity levels in the OF.

Our second objective was to examine whether feather pecking behavior differed among sexes and strains. To date, there has been very little research examining feather pecking among turkeys, and ours is the first study to directly examine feather pecking among turkeys at the individual level. As many as 42% of COMM males (39% RB males), and 57% of COMM females (63% RB females) were observed to perform feather pecking across a 3-day period. The percentages of birds that performed feather pecking was higher overall for females, and was highest for RB females on a per bird basis. In contrast, plumage scores in the neck region were higher for COMM turkeys, and tail scores were higher for COMM vs. RB females, indicating that pecking damage was greater in COMM birds.

Plumage damage was scored at 13 wk (91 d), whereas observations of feather pecking were conducted earlier, at 75 to 83 d, which may explain why behavioral observations were not consistent with plumage scores. Another possible explanation is that some feather pecking interactions from the video may have been scored as gentle because the recipient bird did not react, and classification of the feather pecking interaction depends, in part, on the reaction of the recipient. Furthermore, Busayi et al. (2006) postulated that turkeys of different strains may differ

behaviorally in their responses to feather pecking such that turkeys of a particular strain move away faster, thereby reducing pecking damage, or birds of different strains may differ in feather density, and therefore in susceptibility to pecking damage. Further research is necessary to examine behavior of the recipient, and whether turkeys of different strains react differently to feather pecking.

Feather pecking behavior may have been influenced by the colors used to individually identify birds in the pen. Using colored livestock marker to distinguish between individuals was necessary so that markings were visible from overhead cameras. Marker color, color combinations and location were randomly assigned to birds in each pen. Color was therefore confounded with the location on the turkey's body where the marker was applied. Because each turkey in the pen was marked differently, and because feather pecking behavior may differ among pens, it was not possible to assess the effects of color on pecking behavior. Further research is needed to determine whether color markings of turkeys affect feather pecking behavior by including appropriate control groups that are not marked.

In addition to examining the frequency of feather pecking, we examined when feather pecking occurred in relation to the behavior of the recipient. Results revealed that feather pecking occurred when the recipient birds were performing a wide range of behaviors. However, the majority (> 40%) of feather pecking interactions occurred when the recipient was engaged in "other" behavior, which included sitting, standing, walking and environmental pecking. Interestingly, there were strain differences in when feather pecking occurred, with RB males pecking more when the recipient had been dustbathing, compared to COMM males, and a similar tendency between RB and COMM females. None of the COMM male recipients of feather pecking were observed to dustbathe, indicating that some behavior of turkeys may have changed

as a result of genetic selection for increased feed efficiency and faster growth. However, further research is needed to verify whether dustbathing behavior is affected by selection for faster growth rates and increased feed efficiency because this was not examined in our study.

The considerable amount of scientific literature concerning feather pecking of laying hens provides evidence of a relationship between fear responses and the propensity to develop feather pecking (reviewed in Rodenburg et al., 2013). Most recent studies have been conducted with laying hens that have been divergently selected for high or low feather pecking behavior, respectively, but an earlier study with red jungle fowl also found evidence that feather pecking and fear responses differ between birds that develop feather pecking and birds that do not. Vestergaard et al. (1993) found that the rate of pecks delivered was correlated with the duration of tonic immobility in red jungle fowl. In another study with the F2 generation of a cross between the high and low feather pecking lines, OF behavior at a young age (5 wk) was genetically correlated with pecking behavior in adulthood (Rodenburg et al., 2004). The only association between OF responses and feather pecking that were found here was a tendency for the latency to vocalize to be shorter for COMM female and RB male turkeys that were not observed to perform feather pecking compared to turkeys that developed feather pecking, when tested at 1 wk. Furthermore, there were no differences in corticosterone levels between PECK and NPECK turkeys. As discussed above, the corticosterone levels may not have been reflective of fear in the OF test, but it is possible that the OF test assesses other behavioral responses or motivations, such as social reinstatement motivation, and not just fear, as discussed in Chapter 3. Research examining the relationships between feather pecking and responses in other tests of fear, such as tonic immobility, may provide further insight into whether turkeys that differ in fearfulness differ in feather pecking.

In conclusion, results from this study support our previous work that turkeys' responses in the OF test are repeatable from 4 to 12 wk of age. Furthermore, there were some differences in OF behavior between COMM males and females at 4-5 and 11-12 wk of age, and between RB males and females at 1 and 11-12 wk of age, indicating that age, sex and strain differences should be taken into account when OF test results are interpreted. Our results did not support the hypothesis that feather pecking is associated with differences in temperament measured using the OF test. Additionally, our results provide further evidence of the complex, multifactorial nature of feather pecking among domestic turkeys, which remains poorly understood. Several areas of future research will provide insight into the physiology and behavior turkeys, including research examining the relationship between corticosterone levels and behavioral responses to various other fear tests, and research examining the behavior of the recipient bird during a feather pecking interaction to ascertain whether there are strain differences in how recipients react.

ACKNOWLEDGEMENTS

We thank Dr. Darrin Karcher, and Hybrid Turkeys and Cuddy farms for supplying the poults, and Angelo Napolitano and the staff at the Michigan State University Poultry Research and Teaching Center for their assistance. We would also like to express our gratitude to Janet Ireland for conducting corticosterone assays, and to the student members of the Michigan State University Animal Behavior and Welfare Group for the invaluable assistance with data collection. M.E. was supported by a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship and this work was supported in part by Michigan State University AgBioResearch.

APPENDIX

Table 4.1. Spearman rank correlation (r_S) of open field (OF) behavior among test periods for male and female commercial (COMM) and randombred (RB) turkeys.

	Test 1 vs. 2	Test 1 vs. 3	Test 2 vs. 3
		COMM males	
	N = 31	N = 31	N = 47
Lat. ambulate	0.12	0.40*	0.35*
Lat. vocalize	0.22	0.35^{T}	0.47***
Step no.	0.30	0.40*	0.41**
Square no.	0.31^{T}	0.35^{T}	0.31*
Vocalization no.	0.26	0.24	0.69****
Defecation no.	0.08	0.02	0.58****
		RB males	
	N = 34	N = 34	N = 47
Lat. ambulate	0.19	0.01	0.42**
Lat. vocalize	0.02	0.02	0.37*
Step no.	0.24	0.10	0.46**
Square no.	0.23	-0.05	0.51***
Vocalization no.	0.16	0.05	0.34*
Defecation no.	0.33^{T}	0.17	0.33*
		COMM females	
	N = 32	N = 32	N = 47
Lat. ambulate	0.38*	0.46*	0.41**
Lat. vocalize	0.04	0.33^{T}	0.47***
Step no.	0.32^{T}	0.27	0.54****
Square no.	0.43*	0.37*	0.55****
Vocalization no.	0.32^{T}	0.23	0.64****
Defecation no.	0.23	0.40*	0.56****
		RB females	
	N = 29	N = 29	N = 47
Lat. ambulate	0.59***	0.41*	0.60****
Lat. vocalize	0.34^{T}	0.52**	0.43**
Step no.	0.39*	0.55**	0.73****
Square no.	0.37*	0.57**	0.71****
Vocalization no.	0.37*	0.49**	0.54****
Defecation no.	0.52**	0.41*	0.42**

Note: Lat. = latency. $^{T}P < 0.10$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$

Table 4.2. Behavioral responses of commercial (COMM) and randombred (RB) male and female turkeys to open field testing at 3 ages - Test 1: wk 1; Test 2: wk 4 (males) and wk 5 (females); Test 3: wk 11 (males) and wk 12 (females).

	Step no.	Square no.	Voc. no.	Def. no.	Lat. amb.	Lat. voc.
Test 1						
COMM males	4.5 ± 0.4^{ab}	3.0 ± 0.4^{ab}	4.8 ± 0.4	0.05 ± 0.2	$502\ (236\text{-}600)^{a}$	$246\ (45\text{-}600)^{a}$
RB males	3.5 ± 0.4^{a}	1.5 ± 0.5^{a}	4.7 ± 0.4	-0.2 ± 0.2	$600\ (214\text{-}600)^{ab}$	$522\ (72\text{-}600)^{ab}$
COMM females	4.0 ± 0.4^{ab}	2.6 ± 0.5^{ab}	5.2 ± 0.4	0.07 ± 0.2	$600\ (223\text{-}600)^{a}$	$414\ (5\text{-}600)^{a}$
RB females	5.2 ± 0.4^{b}	3.6 ± 0.5^{b}	6.0 ± 0.4	0.2 ± 0.2	$216\ (47\text{-}600)^{b}$	$7\ (2\text{-}520)^{b}$
Sex	$F_{156} = 1.8$	$F_{156} = 3.3^{T}$	$F_{156} = 5.2^{*}$	$F_{156} = 1.3$		
Strain	$F_{156} = 0.02$	$F_{156} = 0.3$	$F_{156} = 0.7$	$F_{156} = 0.08$		
Sex x strain	$F_{156} = 6.9^{**}$	$F_{156} = 7.4^{**}$	$F_{156} = 1.5$	$F_{156} = 1.0$		
Test 2						
COMM males	3.9 ± 0.2	2.6 ± 0.2	5.5 ± 0.2^{a}	0.8 ± 0.1^{a}	$321\ (80\text{-}600)^{a}$	$69\ (4\text{-}216)^{a}$
RB males	4.3 ± 0.2	2.9 ± 0.2	5.9 ± 0.2^{ab}	1.0 ± 0.1^{ab}	$143\ (55\text{-}600)^{ab}$	$7\ (2\text{-}178)^{b}$
COMM females	4.3 ± 0.2	2.8 ± 0.2	6.2 ± 0.2^{b}	1.3 ± 0.1^{b}	$112\ (56\text{-}464)^{b}$	$4\ (1\text{-}28)^{b}$
RB females	4.7 ± 0.2	3.1 ± 0.2	6.3 ± 0.2^{b}	1.3 ± 0.1^{b}	$76\ (34\text{-}195)^{b}$	$3\ (1\text{-}49)^{b}$
Sex	$F_{221} = 2.8^{T}$	$F_{221} = 0.9$	$F_{221} = 9.7^{**}$	$F_{221} = 15.9^{****}$		
Strain	$F_{221} = 3.8^{T}$	$F_{221} = 1.7$	$F_{221} = 2.3$	$F_{221} = 0.7$		
Sex x strain	$F_{221} = 0.0$	$F_{221} = 0.01$	$F_{221} = 1.1$	$F_{221} = 1.1$		

Note: Lat. = latency, amb. = ambulate, voc. = vocalize, L = Log-rank, W = Wilcoxon. Numbers of steps and squares entered, vocalizations and defecations are presented as lsmean ± SE. Latencies to ambulate and vocalize are presented as median (25-75% quartiles). $^{a\text{-}c}$For each test, values within columns lacking common superscripts differ significantly ($P < 0.05$).
$^{T}P < 0.10$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

Table 4.2. (cont'd).

	Step no.	Square no.	Voc. no.	Def. no.	Lat. amb.	Lat. voc.
Test 3						
COMM males	3.0 ± 0.3	1.0 ± 0.3^{a}	4.6 ± 0.3	0.3 ± 0.1^{a}	$600\ (163\text{-}600)^{a}$	$102\ (15\text{-}600)^{a}$
RB males	3.5 ± 0.3	1.8 ± 0.3^{ab}	4.7 ± 0.3	0.6 ± 0.1^{c}	$337\ (83\text{-}600)^{b}$	$141\ (16\text{-}600)^{a}$
COMM females	3.4 ± 0.3	1.4 ± 0.3^{ab}	5.5 ± 0.3	0.8 ± 0.1^{bc}	$268\ (132\text{-}600)^{b}$	$26\ (7\text{-}117)^{b}$
RB females	3.9 ± 0.3	2.1 ± 0.3^{b}	5.4 ± 0.3	1.0 ± 0.1^{b}	$226\ (78\text{-}550)^{b}$	$84\ (14\text{-}158)^{c}$
Sex	$F_{193} = 2.0$	$F_{193} = 1.8$	$F_{193} = 8.8$**	$F_{193} = 17.4$****		
Strain	$F_{193} = 3.3^{T}$	$F_{193} = 6.1$*	$F_{193} = 0.0$	$F_{193} = 4.3$*		
Sex x strain	$F_{193} = 0.0$	$F_{193} = 0.02$	$F_{193} = 0.2$	$F_{193} = 0.04$		

Note: Lat. = latency, amb. = ambulate, voc. = vocalize, L = Log-rank, W = Wilcoxon. Numbers of steps and squares entered, vocalizations and defecations are presented as lsmean ± SE. Latencies to ambulate and vocalize are presented as median (25-75% quartiles). [a-c]For each test, values within columns lacking common superscripts differ significantly ($P < 0.05$).
[T]$P < 0.10$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

Table 4.3. Sex and strain differences in feather pecking (FP) behavior (n = 4 pens of males and females per strain). Behavior was recorded over 3 days. Raw data (mean ± SE) are presented.

	Frequency of FP interactions[A]	Birds performing FP[B] (%)	Severe FP[B] (%)	FP at resource[B] (%)	FP while dustbathing[B] (%)	FP while preening[B] (%)	FP while other[B] (%)
COMM males	0.9 ± 0.1	42.2 ± 3.9^{a}	0.8 ± 0.1	0.4 ± 0.3^{ab}	0.0 ± 0.0^{a}	1.1 ± 0.06	4.7 ± 0.3^{a}
RB males	0.7 ± 0.06	39.1 ± 6.6^{a}	1.1 ± 0.2	0.04 ± 0.04^{a}	1.5 ± 0.4^{b}	0.8 ± 0.1	4.2 ± 0.2^{a}
COMM females	1.0 ± 0.2	57.8 ± 10.0^{a}	1.5 ± 0.3	1.3 ± 0.4^{bc}	0.1 ± 0.08^{a}	1.1 ± 0.3	3.7 ± 0.1^{ab}
RB females	1.3 ± 0.4	63.3 ± 10.0^{b}	1.6 ± 0.2	1.8 ± 0.2^{c}	1.3 ± 0.4^{ab}	0.8 ± 0.1	2.8 ± 0.5^{b}
Sex	$F_{12} = 4.2^{T}$	$F_{12} = 4.8$*	$F_{12} = 7.4$*	$F_{12} = 26.1$***	$F_{12} = 0.0$	$F_{12} = 0.04$	$F_{12} = 16.0$**
Strain	$F_{12} = 0.0$	$F_{12} = 12.1$**	$F_{12} = 1.1$	$F_{12} = 0.1$	$F_{12} = 17.8$**	$F_{12} = 3.1$	$F_{12} = 4.7^{T}$
Sex x strain	$F_{12} = 1.9$	$F_{12} = 4.2^{T}$	$F_{12} = 0.1$	$F_{12} = 2.6$	$F_{12} = 0.3$	$F_{12} = 0.01$	$F_{12} = 0.6$

[A]The frequency of feather pecking interactions was calculated as the mean across 4 pens of the total number of feather pecking interactions per individual per day in each pen.

[B]The percentage was calculated as the mean across 4 pens of the total percentage of feather pecking interactions occurring over 3 d for each behavior of the recipient bird / the number of birds in the pen.

[a, b]Values within rows lacking common superscripts differ significantly ($P < 0.05$).

[T]$P < 0.10$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Table 4.4. Feather scores (lsmean ± SE) from the neck, tail, and right and left flight feathers of male and female commercial (COMM) and randombred (RB) turkeys.

	Neck	Right flight	Left flight	Tail
COMM males	0.23 ± 0.23^{a}	0.53 ± 0.10	0.58 ± 0.09	0.74 ± 0.09^{a}
RB males	-1.46 ± 0.34^{b}	0.71 ± 0.09	0.65 ± 0.09	0.53 ± 0.10^{ab}
COMM females	0.30 ± 0.23^{a}	0.66 ± 0.09	0.63 ± 0.09	0.66 ± 0.09^{a}
RB females	-0.75 ± 0.28^{b}	0.64 ± 0.09	0.63 ± 0.09	0.24 ± 0.12^{b}
Sex	$F_{240} = 2.1$	$F_{240} = 0.1$	$F_{240} = 0.03$	$F_{240} = 3.4^{T}$
Strain	$F_{240} = 25.0$****	$F_{240} = 0.7$	$F_{240} = 0.13$	$F_{240} = 10.2$**
Sex x strain	$F_{240} = 1.3$	$F_{240} = 1.2$	$F_{240} = 0.09$	$F_{240} = 1.1$

[a, b]Values within rows lacking common superscripts differ significantly ($P < 0.05$).
[T]$P < 0.10$, *$P < 0.05$, **$P < 0.01$.

Table 4.5. Behavioral open field test responses of male and female turkeys that were observed to perform feather pecking (PECK) and that did not perform feather pecking (NPECK) of a commercial (COMM) and randombred (RB) strain. Open field testing was conducted at 3 ages - (Test 1: wk 1; Test 2: wk 4 (males) and wk 5 (females); Test 3: wk 11 (males) and wk 12 (females).

		Step no.	Square no.	Voc. no.	Def. no.	Lat. amb.	Lat. voc.
Test 1							
COMM males	PECK (n = 14)	4.4 ± 0.7	2.9 ± 0.7	4.8 ± 0.6	0.1 ± 0.3	504 (336-600)	365 (12-600)
	NPECK (n = 26)	4.6 ± 0.5	3.1 ± 0.5	4.9 ± 0.5	0.1 ± 0.2	494 (219-600)	233 (58-600)
RB males	PECK (n = 15)	2.8 ± 0.7	0.8 ± 0.7	3.7 ± 0.6	-0.4 ± 0.4	600 (381-600)	600 (324-600)[a]
	NPECK (n = 24)	3.7 ± 0.5	1.8 ± 0.6	5.0 ± 0.5	-0.1 ± 0.3	600 (144-600)	183 (58-600)[b]
COMM females	PECK (n = 24)	3.5 ± 0.5	2.0 ± 0.6	4.8 ± 0.5	0.0 ± 0.2	600 (131-600)	600 (4-600)[a]
	NPECK (n = 14)	4.6 ± 0.7	3.2 ± 0.7	5.8 ± 0.6	0.3 ± 0.3	443 (129-600)	44 (4-450)[b]
RB females	PECK (n = 24)	5.1 ± 0.6	3.5 ± 0.6	6.1 ± 0.5	0.3 ± 0.3	173 (39-600)	6 (2-600)
	NPECK (n = 14)	5.4 ± 0.7	3.9 ± 0.8	6.1 ± 0.7	0.1 ± 0.3	199 (58-600)	7 (2-159)
Test 2							
COMM males	PECK (n = 23)	4.0 ± 0.4	2.5 ± 0.4	5.4 ± 0.3	0.5 ± 0.2	238 (46-600)	69 (4-256)
	NPECK (n = 33)	3.8 ± 0.5	2.5 ± 0.5	5.5 ± 0.3	0.8 ± 0.2	321 (151-600)	63 (4-213)
RB males	PECK (n = 16)	4.4 ± 0.4	3.0 ± 0.4	6.0 ± 0.3	1.1 ± 0.2	136 (51-540)	9 (2-259)
	NPECK (n = 38)	4.2 ± 0.3	2.7 ± 0.3	5.9 ± 0.2	1.0 ± 0.1	158 (62-600)	9 (2-174)
COMM females	PECK (n = 35)	4.4 ± 0.3	2.9 ± 0.3	6.1 ± 0.2	1.2 ± 0.1	112 (41-343)	4 (1-28)
	NPECK (n = 22)	4.4 ± 0.4	2.8 ± 0.4	6.2 ± 0.3	1.2 ± 0.2	127 (57-464)	3 (1-40)
RB females	PECK (n = 34)	4.6 ± 0.4	3.1 ± 0.4	6.2 ± 0.3	1.4 ± 0.2	75 (28-498)	3 (1-95)
	NPECK (n = 20)	4.4 ± 0.5	2.8 ± 0.5	6.4 ± 0.3	1.3 ± 0.2	94 (47-168)	5 (2-7)
	NPECK (n = 31)	4.0 ± 0.5	2.4 ± 0.5	4.7 ± 0.5	0.5 ± 0.2	248 (43-600)	137 (7-600)

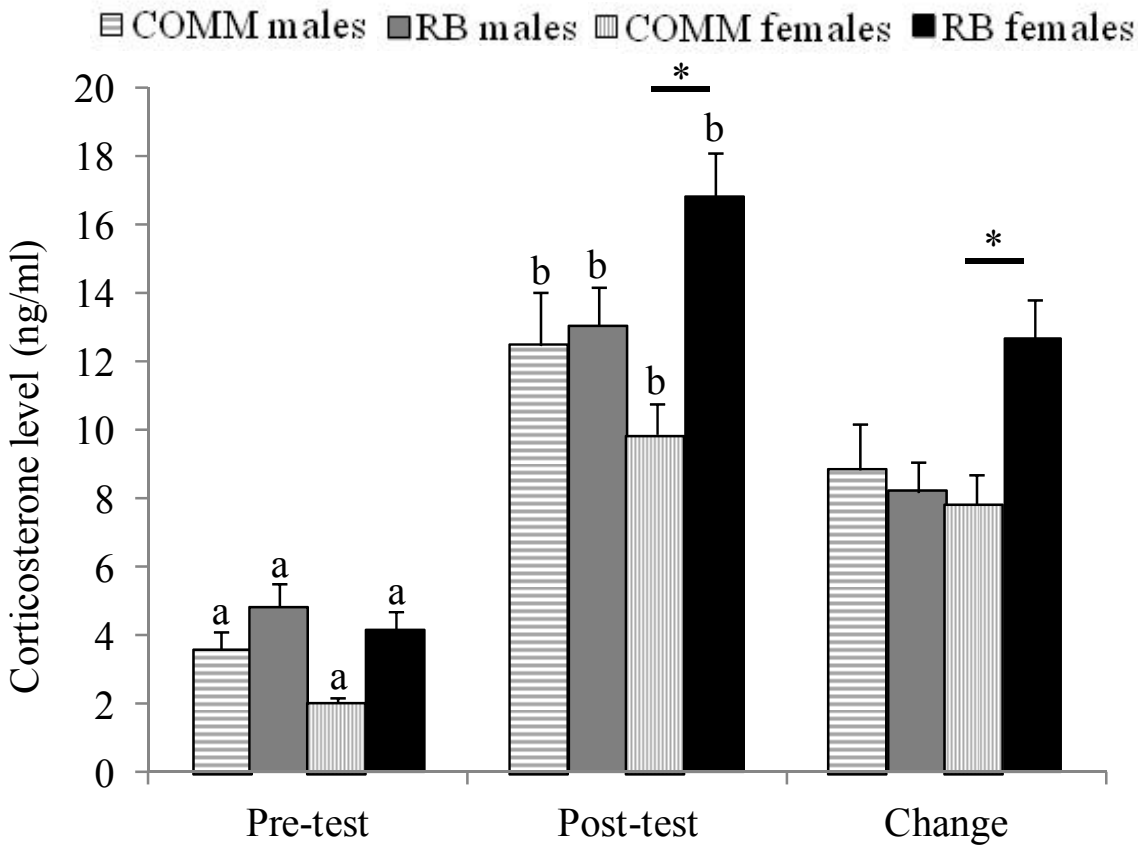
Note: Lat. = latency, amb. = ambulate, voc. = vocalize, L = Log-rank, W = Wilcoxon. Numbers of steps and squares entered, vocalizations and defecations are presented as lsmean ± SE. Latencies to ambulate and vocalize are presented as median (25-75% quartiles). [a-c]For each test, values within columns lacking common superscripts tended to differ significantly ($P < 0.10$).
[T]$P < 0.10$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

Table 4.5. (cont'd).

		Step no.	Square no.	Voc. no.	Def. no.	Lat. amb.	Lat. voc.
Test 3							
COMM males	PECK (n = 20)	1.8 ± 0.9	-0.6 ± 0.9	4.1 ± 0.9	0.4 ± 0.3	600 (113-600)	82 (12-600)
	NPECK (n = 30)	1.6 ± 1.0	-0.5 ± 0.9	4.4 ± 0.9	0.2 ± 0.3	590 (89-600)	91 (5-600)
RB males	PECK (n = 19)	4.0 ± 0.6	2.3 ± 0.6	4.9 ± 0.6	0.6 ± 0.2	166 (57-600)	63 (15-600)
COMM females	PECK (n = 28)	3.2 ± 0.4	1.3 ± 0.4	5.5 ± 0.4	0.8 ± 0.1	193 (131-600)	16 (5-87)
	NPECK (n = 19)	2.6 ± 0.6	0.3 ± 0.6	5.3 ± 0.5	0.6 ± 0.2	268 (131-600)	26 (8-117)
RB females	PECK (n = 29)	4.9 ± 0.8	3.3 ± 0.7	5.6 ± 0.8	1.0 ± 0.3	214 (72-600)	112 (14-162)
	NPECK (n = 21)	4.7 ± 0.8	3.0 ± 0.8	5.6 ± 0.8	1.1 ± 0.3	253 (131-544)	82 (49-131)
Sex		$F_{211} = 1.6$	$F_{211} = 0.7$	$F_{211} = 8.1$**	$F_{211} = 15.1$***		
Strain		$F_{211} = 0.3$	$F_{211} = 0.2$	$F_{211} = 0.7$	$F_{211} = 1.5$		
Classification		$F_{211} = 0.4$	$F_{211} = 0.3$	$F_{211} = 0.1$	$F_{211} = 0.3$		
Sex x strain x classification		$F_{211} = 0.06$	$F_{211} = 0.1$	$F_{211} = 0.3$	$F_{211} = 1.0$		

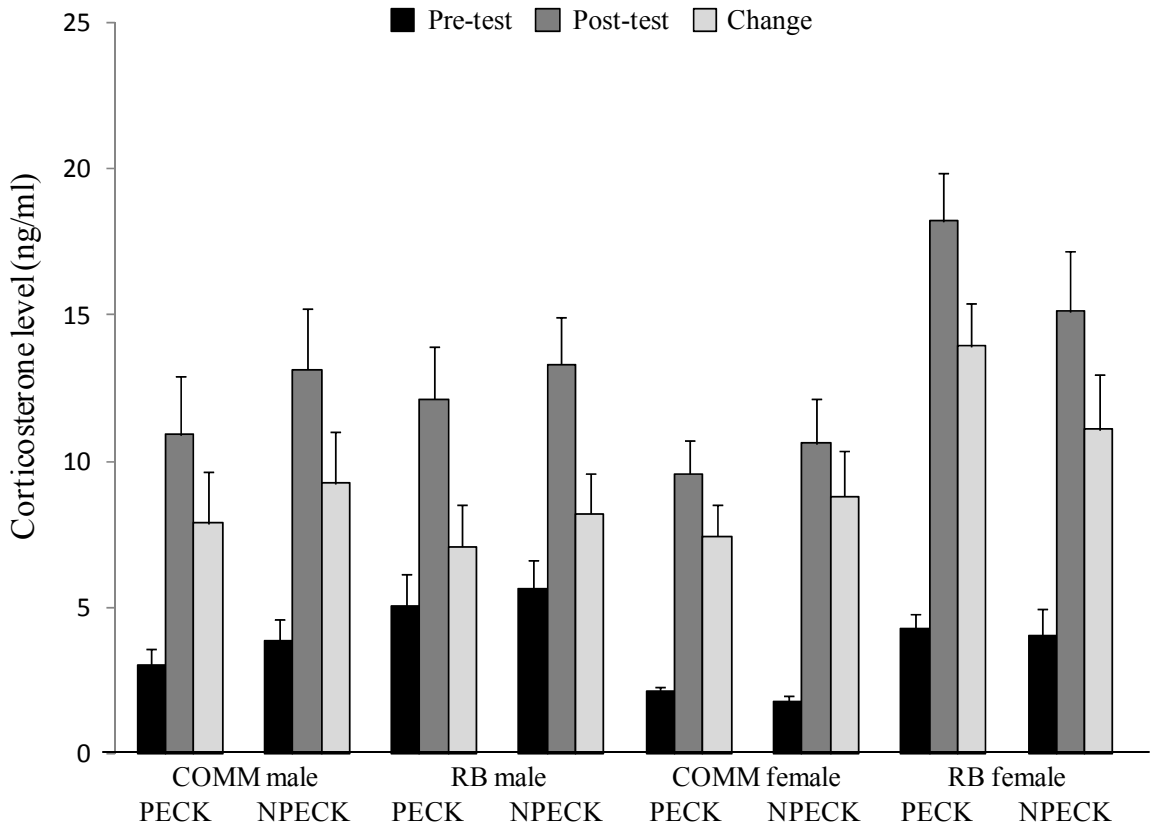
Note: Lat. = latency, amb. = ambulate, voc. = vocalize, L = Log-rank, W = Wilcoxon. Numbers of steps and squares entered, vocalizations and defecations are presented as lsmean ± SE. Latencies to ambulate and vocalize are presented as median (25-75% quartiles). [a-c]For each test, values within columns lacking common superscripts tended to differ significantly ($P < 0.10$).
[T]$P < 0.10$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$.

Figure. 4.1. Plasma corticosterone levels (pre-test[a], post-test[a] and change from pre- to post-test) of male and female turkeys of a commercial (COMM) and randombred (RB) strain tested in an open field test at 11 (males) and 12 (females) weeks of age. Raw data (mean ± SE) are presented. Statistical analyses were conducted on log-transformed data.



[a]Pre-test blood samples were collected one day prior to testing at the same time that post-test samples were to be collected the following day. Post-test samples were collected 15 min following the start of the test.
Different letters (a, b) indicate significant differences between pre- and post-test corticosterone levels of males and females within strains ($P < 0.05$).
*Asterisks indicate significant differences between sexes of different strains ($P < 0.05$).

Figure. 4.2. Plasma corticosterone levels (pre-test[a], post-test[a] and change from pre- to post-test) of male and female turkeys of a commercial (COMM) and randombred (RB) strain that were observed to perform feather pecking (PECK) and that did not perform feather pecking (NPECK). Turkeys were tested in an open field test at 11 (males) and 12 (females) wk of age.
Raw data (mean ± SE) are presented. Statistical analyses were conducted on log-transformed data. No significant differences were found.



[a]Pre-test blood samples were collected one day prior to testing at the same time that post-test samples were to be collected the following day. Post-test samples were collected 15 min following the start of the test.

REFERENCES

REFERENCES

Armario, A., N. Daviu, C. Muñoz-Abellán, C. Rabasa, S. Fuentes, X. Belda, H. Gagliano, and R. Nadal. 2012. What can we know from pituitary–adrenal hormones about the nature and consequences of exposure to emotional stressors? *Cell. Mol. Neurobiol.* 32: 749–758.

Bilčík, B., and L. J. Keeling. 1999. Changes in feather condition in relation to feather pecking and aggressive behaviour in laying hens. *Br. Poult. Sci.* 40: 444-451.

Blokhuis, H. J., L. J. Keeling, A. Gavinelli, and J. Serratosa. 2008. Animal welfare's impact on the food chain. *Trends Food Sci. Tech.* 19: S79-S87.

Busayi, R. M., C. E. Channing, and P. M. Hocking. 2006. Comparisons of damaging feather pecking and time budgets in male and female turkeys of a traditional breed and a genetically selected male line. *Appl. Anim. Behav. Sci.* 96: 281-292.

Cockrem, J. F. 2007. Stress, corticosterone responses and avian personalities. *J Ornithol*. 148: S169–S178.

Dalton, H. A., B. J. Wood, and S. Torrey. 2013. Injurious pecking in domestic turkeys: development, causes and potential solutions. *World's Poult. Sci. J.* 69: 865-875.

de Haas, E. N., J. E. Bolhuis, B. Kemp, T. G. G. Groothuis, and T. B. Rodenburg. 2014. Parents and early life environment affect behavioral development of laying hen chickens. *PLoS ONE*. 9: 1-12.

de Haas, E. N., B. L. Nielsen, A. J. Buitenhuis, and T. B. Rodenburg. 2010. Selection on feather pecking affects response to novelty and foraging behaviour in laying hens. *Appl. Anim. Behav. Sci.* 124:90-96.

Duggan, G., T. Widowski, M. Quinton, and S. Torrey. The development of injurious pecking in a COMM turkey facility. 2014. *J. Appl. Poult. Res.* 23: 1-11.

Erasmus, M. A. 2009. Examining physical methods for on-farm killing of turkeys. MS Thesis, University of Guelph, Guelph, ON, Canada.

Erasmus, M. A., and J. C. Swanson. 2014. Temperamental turkeys: reliability of behavioral responses to four tests of fear. *Appl. Anim. Behav. Sci.* 157: 100-108.

Federation of Animal Science Societies, 2012. Farm Animal Integrated Research. Available Online: http://www.fass.org/docs/FAIR2012_Summary.pdf

Forkman, B., A. Boissy, M. –C. Maunier-Salaün, E.Canali, and R. B. Jones. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiol. Behav.* 92: 340-374.

Hocking, P. 1993. Welfare of turkeys. Page 125-138 in Proceedings of the Fourth European Symposium on Poultry Welfare. Savory C J and B. O. Hughes, eds. Universities Federation for Animal Welfare, Potters Bar, UK.

Hazard, D., S. Leclaire, M. Couty, and D. Guémené. 2008. Genetic differences in coping strategies in response to prolonged and repeated restraint in Japanese quail divergently selected for long or short tonic immobility. *Horm. Behav.* 54: 645-653.

Huber-Eicher, B., and F. Sebö. 2001. The prevalence of feather pecking and development in COMM flocks of laying hens. *Appl. Anim. Behav. Sci.* 74: 223-231.

Huff, G., W. Huff, N. Rath, A. Donoghue, N. Anthony, and K. Nestor. 2007. Differential effects of sex and genetics on behavior and stress response of turkeys. Poult. Sci. 86:1294-1303.

Jensen, P., L. Keeling, K. Schutz, L. Andersson, P. Mormede, H. Brandstrom, B. Forkman, S. Kerje, R. Fredriksson, C. Ohlsson, S. Larsson, H. Mallmin, and A. Kindmark. 2005. Feather pecking in chickens is genetically related to behavioral and developmental traits. *Physiol. & Behav.* 86:52-60.

Jones, R. B., H. J. Blokhuis, and G. Beuving. 1995. Open-field and tonic immobility responses in domestic chicks of 2 genetic lines differing in their propensity to feather peck. *Br. Poult. Sci.* 36:525-530.

Jones, B., and A. Boissy. 2011. Fear and other negative emotions. Page 81 in Animal Welfare. 2nd rev. ed. Appleby, M. C, J. A. Mench, I. A. S. Olsson, and B. O. Hughes, eds. CABI International, Cambridge, MA., USA.

Koolhaas, J. M., S. M. Korte, S. F. De Boer, B. J. Van Der Vegt, C. G. Van Reenen, H. Hopster, I. C. De Jong, M. A. W. Ruis, and H. J. Blokhuis. 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neurosci. Biobehav. Rev.* 23: 925–935.

Korte, S. M., G. Beuving, W. Ruesink, and H. J. Blokhuis. 1997. Plasma catecholamine and corticosterone levels during manual restraint in chicks from a high and low feather pecking line of laying hens. *Physiol. & Behav.* 62: 437-441.

Martin, P., and P. Bateson. 1993. The reliability and validity of measures. Pages 114-116 in Measuring behaviour: an introductory guide. 2nd rev. ed. Cambridge University Press, New York, NY., USA.

Nestor, K. E., M. G. McCartney, and N. Bachev. 1969. Relative contributions of genetics and environment to turkey improvement. *Poult. Sci.* 48:1944-1949.

Rodenburg, T. B., A. J. Buitenhuis, B. Ask, K. A. Uitdehaag, P. Koene, J. J van der Poel, J. A. M. van Arendonk, and H. Bovenhuis. 2004. Genetic and phenotypic correlations between feather pecking and open-field response in laying hens at two different ages. *Behav. Genet.* 34: 407-415.

Rodenburg, T. B., Komen, H., Ellen, E. D., Uitdehaag, K. A., and J. A. M. van Arendonk. 2008. Selection method and early life-history affect behavioral development, feather pecking and cannibalism in laying hens: a review. *Appl. Anim. Behav. Sci.* 110: 217-228.

Rodenburg, T. B, M. M. van Krimpen, I. C. De Jong, E. N. de Haas, M. S. Kops, B. J. Riedstra, R. E. Nordquist, J. P. Wagenaar, M. Bestman and C. J. Nicol. 2013. The prevention and control of feather pecking in laying hens: identifying the underlying principles. *World's Poult. Sci. J.* 69: 361-374.

Suarez, S. S., and G. G. Gallup. 1983. Social reinstatement and open-field testing in chickens. *Anim. Learn. Behav.* 11: 119-126.

van Hierden, Y. M., S. M. Korte, W. Ruesink, C. G. van Reenen, B. Engel, J. M. Koolhaas, and H. J. Blokhuis. 2002. The development of feather pecking behavior and targeting of pecking in chicks from a high and low feather pecking line of laying hens. *Appl. Anim. Behav. Sci.* 77: 183-196.

Vestergaard, K., J. P. Kruijt, and J. A. Hogan. 1993. Feather pecking and chronic fear in groups of red junglefowl: their relations to dustbathing, rearing environment and social status. *Anim. Behav.* 45: 1127-1140.

CHAPTER 5: INDIVIDUAL AND STRAIN DIFFERENCES IN OPEN FIELD BEHAVIOR AND ITS RELATIONSHIP WITH MEAT QUALITY OF TURKEYS

ABSTRACT

The occurrence of pale, soft, exudative (PSE) turkey meat has become increasingly problematic, necessitating research examining factors that influence meat quality. Differences in meat quality are known to exist among pigs and cattle differing in temperament characteristics such as fearfulness. However, associations between temperament and meat quality of turkeys have not been examined. The objectives of this research were to 1) examine differences in meat quality characteristics (pH, R-value, and *L**, *a** and *b** color values) between a commercial (COMM) and randombred (RB) turkey strain, and 2) examine the relationship between meat quality and behavioral and physiological responses in an open field (OF) test, which is used to assess fear in poultry. Male COMM and RB turkeys were each housed in groups of 16 in 4 pens. Turkeys were individually tested in an OF (2.74 x 2.74 m, divided into 81 squares) at 11 weeks and birds were grouped into two clusters based on cluster analysis of OF behavior. Plasma corticosterone levels were assessed in subsamples of birds (n = 20/strain) before and after OF testing. Turkeys were processed and meat quality characteristics were evaluated in 10 to 11 birds per pen (n = 44/strain) at 15 to 17 weeks (COMM) and 20 to 21 weeks (RB). Results were analyzed using a mixed model (SAS 9.3). The R-value tended to be lower (COMM: 1.18 ± 0.02, RB: 1.24 ± 0.02, $P = 0.07$) for the COMM vs. the RB strain, but there were no differences in pH, or *L**, *a** or *b** color values. Corticosterone levels did not differ between clusters of either strain, and no differences in any meat quality characteristics were found between clusters. Results indicate that plasma corticosterone responses of COMM and RB turkeys to OF testing are similar, and there is little difference in meat quality characteristics between RB and COMM male

turkeys. Within strains, individual differences in behavioral OF responses are not associated with differences in meat quality.

INTRODUCTION

Annual turkey production in the United States has increased from 90 million in 1962 to over 250 million in 2012 (USDA, 2013). Increased turkey production was facilitated by changes in production, such as intensive housing, and selective breeding methods that reduced costs associated with raising turkeys (Martin, 2009). Concurrent with increased production, problems with meat quality have also increased (Petracci and Cavani, 2012). For example, the incidence of pale, soft, exudative (PSE) turkey meat which is associated with intensive selection for increased muscle growth (Strasburg and Chiang, 2009) has become increasingly problematic "due to the demand for processed, value-added poultry meat" (Sosnicki et al., 1998).

Meat quality, including PSE, may be affected by genetic strain, but results are contradictory. Research by Werner et al. (2008) failed to find a difference in meat quality characteristics (pH and color) between four turkey strains, whereas Sarica et al. (2011) reported differences in color parameters and pH between a fast growing, commercial strain and a slower growing strain. Another study comparing differences in breast muscle function among four turkey strains reported that functional characteristics that are ultimately related to meat quality (e.g. shear force and water holding capacity) are altered in faster growing strains, such that there is a "decrease in post-mortem functionality" relative to slower growing strains (Updike et al., 2005).

Meat quality may also be affected by differences in individual characteristics of animals such as stress reactivity, fearfulness, and other temperament characteristics. Research with pigs

has demonstrated that some muscle characteristics associated with meat quality (e.g. pH and muscle lactate levels) are associated with the animal's reactivity to stressors occurring months prior to slaughter (see Terlouw, 2005). For example, pigs displaying reduced fear of humans are less reactive to slaughter procedures, and pigs displaying reduced fear of novel objects have higher muscle pH levels (Terlouw and Rybarczyk, 2008).

Similarly, steers with poor temperaments (measured using flight speed, defined as the time taken to move a specified distance after being released from a weighing crate) are more susceptible to stress before slaughter; and have lower live weights, poorer feed conversion efficiency, lower carcass dressing percentages and lower initial muscle pH levels (Petherick et al., 2002; reviewed in Norris et al., 2014). Temperament is also associated with production measures such as bodyweight and time spent eating (Cafe et al., 2011; reviewed in Norris et al., 2014) and other meat quality characteristics, including cooking loss, shear force, color (Cafe et al., 2011) and dark cutting (Voisinet et al, 1997).

There is scant research examining temperament and meat quality in poultry, and results are contradictory. Research with broiler chickens failed to find a relationship between fear (measured using tonic immobility) and meat quality (Debut et al., 2003). However, meat quality was lower in quail selected for increased fear (increased tonic immobility duration) that were exposed to acute stress before slaughter, as indicated by higher pH following slaughter and increased drip loss (Remignon et al., 1998). The relationship between meat quality and fear responses has not been examined in turkeys, but results from studies with other species suggest that animals that differ in stress reactivity and fear responses may differ in characteristics associated with meat quality.

Fearfulness is one of the most widely studied aspects of poultry temperament. In addition to the tonic immobility test, the open field test is a widely used test for assessing fear responses of poultry. Recently, it was demonstrated that commercial male turkeys' responses in an open field test are repeatable between weeks (Erasmus and Swanson, 2014), and some open field responses differ between commercial and randombred turkeys (Chapter 4). This study tested the hypothesis that differences in behavioral and physiological open field responses are associated with differences in meat quality characteristics of turkeys. Additionally, strain differences in meat quality characteristics were also examined.

MATERIALS AND METHODS

All procedures were approved by the Institutional Animal Care and Use Committee of Michigan State University. Materials and methods were the same as for Chapter 4.

Animals and housing

This study was conducted in conjunction with another study examining the effects of different chilling treatments on meat quality of male randombred (RB) and male commercial (COMM) turkeys. Results reported here only concern the relationship between OF behavior and meat quality. Results concerning chilling treatments will be reported elsewhere.

For a full description of animals and housing, refer to Chapter 4 materials and methods. In total, 62 male RB turkeys were obtained. Commercial poults were placed 1 d prior to placement of RB poults. Male COMM (Hybrid Converter, Hybrid Turkeys, Kitchener, ON, Canada) and male RB (RBC2 line, The Ohio Agricultural Research and Development Center of The Ohio State University, Wooster, OH) turkeys were brooded in same-strain groups in eight

littered (wood shavings) pens (2.40 m x 3.05 m) in the same barn from 1 d - 2 wk of age. At 2 wk, birds were distributed throughout the barn and randomly assigned to one of 16 littered pens (2.4 m x 3.05 m) in the same barn and housed in same-strain groups (14 to 20 birds per pen; 4 pens of males and females of each strain) where they remained until study completion at 14 wk. At 7 wk, some COMM turkeys were removed from their pens for use in another study and the number of COMM turkeys in each pen was reduced to, and maintained at 16. Turkeys were individually identified using colored non-toxic livestock marker as described in Chapter 4 (materials and methods section).

Open field (OF) testing

Birds were tested in the OF test at 1 wk (Test 1), 4 wk (Test 2) and 11 wk (Test 3), but only results from Test 3 were used here to examine the relationship between OF behavior and meat quality. Data from Test 3 was used because reliability of OF behavior is higher at later ages (Erasmus and Swanson, 2014; Chapter 4). Furthermore, physiological responses to OF testing could only be analyzed for Test 3 due to time and cost constraints associated with running the assays.

The OF test arena and testing procedures are the same as described in Erasmus and Swanson (2014). Briefly, each bird was placed in the center of the OF test arena for 10 min. and behaviour was recorded in real time (60i fields per second) from two overhead high definition camcorders (VIXIA HF M41, Canon USA, Inc. Melville, NY, USA). Behavioural data, including the latency to ambulate, number of steps taken, number of squares entered, latency to vocalize, number of vocalizations and number of defecations, were collected from video recordings.

OF responses and meat quality

To examine the relationship between meat quality and OF responses, birds were assigned to clusters based on hierarchical cluster analysis of responses during OF Test 3. Test responses that were used in the cluster analysis included the latency to ambulate, number of steps and number of defecations. These test measures were selected because ambulatory behavior is believed to be an indicator of the level of fearfulness in the OF, whereas vocalizations are believed to be associated with social reinstatement motivation (Forkman et al., 2007). Except for the number of defecations of RB males ($r_S = 0.33$) and the latency of COMM males to ambulate ($r_S = 0.35$), all of the aforementioned OF test measures had moderate reliability ($r_S > 0.4$) when turkeys were tested between 4 and 11 weeks of age (Chapter 4).

Blood sample collection and corticosterone assays

Blood samples were collected from 6 birds in each pen ($n = 24$ birds/strain) during the second and third test periods. Only blood samples from wk 11 were assayed due to time and cost constraints. Furthermore, Test 3 rather than Test 2 was chosen for blood sample analysis because Test 3 was the last test that was conducted before slaughter. Blood samples were collected following the procedures outlined in Chapter 4. Briefly, samples were collected one day prior to OF testing in order to obtain pre-test corticosterone levels, and again the following day 15 min. after OF testing started. The same competitive enzyme immunoassay kit (Corticosterone HS, IDS Inc., Fountain Hills, AZ) used by Huff et al. (2007) was used to determine plasma corticosterone levels.

Processing

Males from both strains were slaughtered in three groups of up to 20 birds each. Five birds from each pen were randomly assigned to a slaughter group. To mimic current turkey production, male COMM turkeys (14.52 ± 0.29 kg average liveweight) were reared to 15, 16, and 17 wk in three replicates, respectively. The first and second replicate were slaughtered one week apart and the third replicate was slaughtered four days after the second. Considering their small body size, RB males (11.04 ± 0.10 kg average liveweight) were reared up to 20 and 21 wk and processed in three replicates with three to four days between processing.

After feed withdrawal for 12 h, 20 turkeys from each strain (5 turkeys per pen, randomly selected) were individually tagged in the wing and transported for 10-15 min. to the Michigan State University Meat Laboratory. The outside temperature at the time of transportation ranged between 5 and 14 °C. Upon arrival at the Meat Laboratory, the turkeys were shackled and electrically stunned for 3 s (80 mA, 60 Hz, 110 V), bled for 90 s by severing both the carotid artery and jugular vein on one side of the neck, and scalded for 120 s at 59 ºC. After mechanical defeathering and manual evisceration, the breast muscle (*Pectoralis major*) was obtained 20-25 min postmortem for meat quality analysis.

Muscle pH, R-value, and color measurements

The pH was measured following procedures of Jeong et al. (2011). Briefly, 5 g of muscle tissue was taken from the cranial area of the left breast fillet and homogenized in 25 ml of distilled/deionized water. The pH was then measured using a pH electrode (model 13-620-631, Fisher Scientific Inc., Houston, TX) attached to a pH meter (Accumet AR15, Fisher Scientific Inc., Pittsburgh, PA). The R-value (ratio of inosine:adenosine) was measured as an indicator of

adenosine triphosphate (ATP) depletion in the muscle using the method of Thompson et al. (1987).

Color measurements were taken after carcasses were chilled, following the procedure described by Jeong et al. (2011). Briefly, *L** (lightness), *a** (redness), and *b** (yellowness) values were measured on the skinless surface of each breast using a chromameter (CR-400, 8-mm aperture, illuminant C; Konika Minolta Sensing Inc., Osaka, Japan) that had been calibrated with a white plate (*L**, 97.28; *a**, −0.23; *b**, 2.43). Six readings of CIE *L**, *a**, and *b** were obtained from the area that was "free of any obvious blood-related defects, such as bruises, hemorrhages, or full blood vessels" (Fletcher et al., 2000).

Statistical analysis

All statistical analyses were performed using SAS version 9.3. Similar to a study examining individual differences in behaviour of goats (Miranda-de la Lama et al., 2014), hierarchical cluster analysis (PROC CLUSTER) was used to assign turkeys to a particular cluster. Prior to cluster analysis, variables (latency to ambulate, number of steps, number of squares and number of defecations) were standardized (mean = 0, sd = 1). Ward's method of clustering was specified.

The GLIMMIX procedure for a negative binomial distribution with the log link function was used to verify that the numbers of steps and defecations differed significantly between clusters. The LIFETEST procedure was used to verify that the latency to ambulate differed between clusters. The variable corresponding to cluster was specified in the STRATA statement.

Corticosterone levels, carcass weights and meat quality parameters (pH, R-value, and *L**, *a** and *b** color values) were compared among strains and clusters using a mixed model with

Tukey's adjustment for multiple comparisons. The pH, R-value and carcass weights were assessed before chilling treatments. Therefore, the slaughter group was included as a random effect and the lsmeans of the strain and cluster and their interaction were analyzed. Because turkeys reared within the same pen may have been more similar to turkeys reared in other pens, the variable corresponding to pen was also included as a random effect.

Color measurements (L^*, a^* and b^* values) were assessed after carcasses had been chilled. Therefore, the chilling treatment was included as a random effect nested within the slaughter group. Pen was also included as a random effect. Pre- and post-test corticosterone levels, carcass weight and a^* values were log-transformed. The change in corticosterone level (post-test – pre-test) was transformed using log + 1, and b^* values were transformed using log + 10 to meet normality assumptions.

RESULTS

Cluster analysis

Two clusters were identified for COMM (Fig. 5.1) and RB (Fig. 5.2) turkeys. In general, turkeys in cluster 1 did not ambulate whereas turkeys in cluster 2 ambulated. The median (25-75% range) latency to ambulate (s) differed significantly between clusters of both strains (COMM cluster 1 = 600 (600-600) vs. COMM cluster 2 = 89 (53-133), log-rank = 57.8, $P < 0.0001$, Wilcoxon = 52.2, $P < 0.0001$; RB cluster 1 = 600 (600-600) vs. RB cluster 2 = 86 (47-210), log-rank = 53.2, , $P < 0.0001$, Wilcoxon = 42.4, $P < 0.0001$).

Similarly, the number of steps (lsmean ± SE) differed significantly between clusters within strains (COMM cluster 1 = 0 (0-2) vs. COMM cluster 2 = 44 (36-76), $F_{42} = 28.7$, $P < 0.0001$; RB cluster 1 = 0 (0-0) vs. RB cluster 2 = 50 (14-119), $F_{46} = 119.2$, $P < 0.0001$).

However, the number of defecations differed between RB clusters but not between COMM clusters (COMM cluster 1 = 1 (0-2) vs. COMM cluster 2 = 2 (1-2), $F_{42} = 1.5$, $P = 0.2$; RB cluster 1 = 1 (0-1) vs. RB cluster 2 = 3 (1-4), $F_{46} = 18.6$, $P < 0.0001$). Pre- and post-test corticosterone levels did not differ between clusters within strains (Fig. 5.3).

Strain and cluster differences in meat quality and carcass weight

A total of four breast fillet samples from each strain were classified as PSE-like, based on pH ≤ 5.8 (Rathgeber et al., 1999).

Results of the main effects of strain, cluster and their interaction on meat quality and carcass weight are presented in Table 5.1. The main effects of strain and cluster were significant for carcass weight, but carcass weight did not differ between clusters within strains. No significant differences were found for *L** or pH. The interaction between strain and cluster was significant for *a**, but no post-hoc differences were found. There was a tendency for *b** to differ between clusters, and the R-value tended to differ between strains.

DISCUSSION

The only differences in meat quality characteristics that were found between strains was a difference in carcass weight, which is to be expected because COMM turkeys have been selected for increased body weight, and a tendency for R-values of RB turkey fillets to be higher on

average than R-values of COMM turkey fillets. The higher R-value in RB turkey fillets indicates that the onset of rigor mortis may be accelerated in RB turkeys compared to COMM turkeys. Accelerated rigor mortis along with low pH (< 5.8) and muscle temperatures above 35 °C are associated with reduced meat quality; specifically, leading to the development of PSE meat (Sosnicki et al., 1998). Although R-values tended to differ, *L** values did not differ between strains, and the percentage of PSE-like breast fillets was similar for both strains (COMM = 9%, RB = 8%). Overall, results demonstrated that there is little difference in meat quality characteristics between the slower growing RB strain and the COMM strain.

Our results are in agreement with the majority of studies comparing meat quality between different turkey strains. Results support those of Updike et al. (2005) who compared the same RB line that was used here with a commercial strain at 16, 18 and 22 weeks. Although strains differed in shear force and water holding capacity, pH values measured at 2 or 20h post-slaughter did not differ between the RB and COMM strains (Updike et al., 2005). Results from a study comparing lightweight and heavyweight lines of British United Turkeys (BUT) similarly reported no genotype differences in the rate of pH decline (Yost et al., 2002). Our results are also in agreement with those of Werner et al. (2008) comparing meat quality among a faster growing, commercial turkey strain (BUT Big 6) and three lines of slower growing broad breasted bronze turkeys. The faster growing commercial strain did not differ from the broad breasted bronze strains in color characteristics or in pH levels measured 20 min post-slaughter, but the pH measured at 4 h post-slaughter was significantly higher in the commercial strain compared to two of the three broad breasted bronze strains (Werner et al., 2008). In contrast, Sarica et al. (2011) reported that pH and *a** values were higher, and *b** values were lower in breast muscle from

commercial Hybrid Converter turkeys (the same strain used in our study) compared to slower growing American Bronze turkeys when birds were slaughtered at 17 and 21 weeks.

Discrepancies among studies may be due to genotype and methodological differences in animal housing and husbandry, and conditions before and at slaughter. Sarica et al. (2011) report genotype differences in meat quality results that were pooled across sexes and across different housing environments, whereas other studies used only males (Updike et al., 2005; Werner et al., 2008) or reported data separately for males and females (Yost et al., 2002). Furthermore, some turkeys in the study by Sarica et al. (2011) had outdoor access, whereas those in the other studies were reared indoors (Updike et al., 2005; Werner et al., 2008; Yost et al., 2002). Indeed, Sarica et al. (2011) found differences in a^* values, and protein, ash and fat content of breast meat from birds with and without outdoor access. We did not examine strain differences in water holding capacity or shear force, which have been reported to differ between strains (Updike et al., 2005). Further research is needed to examine how genetics and environmental conditions influence water holding capacity and shear force.

The majority of research concerning temperament and meat quality has been conducted with cattle and pigs, and several studies have provided evidence of a relationship between meat quality and behavioral temperament characteristics such as responses to fear tests, and physiological characteristics such as stress reactivity (e.g. Petherick et al., 2002; Terlouw and Rybarczyk, 2008). Individual animals are consistent in their stress reactivity, and the slaughter process is associated with a number of events and conditions that may induce stress. Indeed, the effects of pre-slaughter stress on meat quality have been well documented. Ante-mortem stressors such as transportation, heat stress and struggling of birds before slaughter promote faster post-mortem glycolysis which is associated with the development of PSE meat (McCurdy

et al., 1996; McKee and Sams, 1997) and increased ultimate pH levels (Terlouw and Rybarczyk, 2008). Therefore, it can be hypothesized that if individual animals are consistent in their stress reactivity and the slaughter process induces a stress response, then animals that differ in stress reactivity should differ in meat quality characteristics regardless of whether stress reactivity is assessed immediately prior to slaughter or weeks or months before slaughter.

Most research comparing meat quality among individuals differing in stress reactivity have done so by manipulating conditions immediately prior to slaughter. Turkeys in our study were likely exposed to acute stress before slaughter, because preslaughter and slaughter procedures are associated with a number of events that may elicit stress and fear responses, including catching and crating, transportation, separation from conspecifics, and temperature fluctuations (Terlouw et al., 2008). However, the conditions associated with our study may not have elicited large enough stress responses for differences in meat quality between clusters to become apparent. Indeed, turkeys were transported a distance of 5.7 km, with a duration of 10 to 15 min, and research by Owens and Sams (2000) found that transporting turkeys for as long as 3 h did not adversely affect meat quality. Further research is needed to examine the relationships between acute stress, temperament and meat quality of turkeys. Additional research is needed to examine whether turkeys' physiological stress responses during various fear tests are consistent over time.

In order to examine the relationship between turkey temperament (measured during rearing) and meat quality, we recorded both behavioral and physiological (plasma corticosterone levels) responses of turkeys in an OF test at 11 weeks of age. Post-test corticosterone levels were significantly higher vs. pre-test levels in both strains, indicating that turkeys had in fact experienced increased stress levels resulting from OF testing procedures (Chapter 4). We used

cluster analysis to group COMM and RB turkeys based on their behavior in an OF test, and verified that ambulatory behavior was indeed different between clusters within strains. However, there were no differences between clusters within strains in pH, R-value, or color characteristics. In addition, corticosterone levels did not differ between clusters within strains, indicating that differing behavioral responses to OF testing were not reflective of corresponding differences in physiological responses. Results therefore do not support our hypothesis that turkeys differing in open field behavior differ in meat quality characteristics. Because plasma corticosterone levels could only be assessed in 8-14 turkeys of each group, the sample size may have been too low to detect differences in corticosterone levels between clusters.

There do not appear to be any other studies examining the relationship between temperament and meat quality of turkeys. Nonetheless, our results are in agreement with those of Debut et al. (2003), who examined the association between meat quality and tonic immobility responses in female broiler chickens of a fast-growing and slow-growing strain. No association was found between fearfulness measured using the tonic immobility test and meat pH or color values (Debut et al., 2003). Similarly, a study with quail from genetic lines selected for long (increased fear) and short (decreased fear) durations of tonic immobility, respectively, found no differences in meat color characteristics between short and long tonic immobility lines, but the pH measured 24 h post-slaughter was significantly higher in quail from the long tonic immobility line when quail were exposed to an acute stressor before slaughter (Remignon et al., 1998).

An explanation for the discrepancies between our results and those of Remignon et al. (1998) and between those of Debut et al. (2003) and Remignon et al. (1998) may be that meat quality differences between birds differing in temperament characteristics are only observed when birds are exposed to acute stress before slaughter. Although Debut et al. (2003) examined

the effects of different pre-slaughter stressors on meat quality, they did not do so within tonic immobility classifications. Therefore, they were not able to test whether exposing chickens to a particular stressor before slaughter affected meat quality differently depending on the chicken's tonic immobility reaction. Furthermore, only TI was examined and it is possible that the relationship between meat quality and fear responses may depend on the type of fear test used.

In conclusion, our results are generally in agreement with most other studies that there is little difference in meat quality characteristics between randombred and commercial male turkeys. Although studies with pigs have found differences in meat quality between pigs differing in fear responses during rearing (Terlouw and Porcher, 2005), our results did not support the hypothesis that there is an association between meat quality and open field responses of turkeys.

ACKNOWLEDGEMENTS

We thank Hybrid Turkeys and Cuddy Farms for supplying the poults, and Dr. Darrin Karcher and Angelo Napolitano and the Michigan State University Poultry Research and Teaching Center for assistance. Furthermore, we thank Janet Ireland for conducting corticosterone assays, and Thanikarn Sansawat, Pranjal Singh and the student members of the Michigan State University Animal Behavior and Welfare Group for assistance with data collection. M.E. was supported by a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship and this work was supported by Michigan State University AgBioResearch.

APPENDIX

Table 5.1. Differences in live weight, carcass weight, pH, R-value, and *L**, *a**, and *b** values between male turkeys of commercial (COMM) and randombred (RB) strains.
Raw data are presented (mean ± SE). Analyses were conducted with transformed data (live weight, carcass weight, and *a** and *b** values).

	Carcass weight (kg)	*L**	*a**	*b**	pH	R-value
COMM Cluster 1 (n = 31)	14.0 ± 0.4^{a}	52.5 ± 0.9	4.4 ± 0.3	-1.1 ± 0.4	6.0 ± 0.03	1.2 ± 0.02
COMM Cluster 2 (n = 13)	14.4 ± 0.6^{a}	52.3 ± 1.4	4.1 ± 0.5	-1.7 ±0.7	6.0 ± 0.06	1.2 ± 0.04
RB Cluster 1 (n = 25)	10.8 ± 0.1^{b}	51.3 ± 0.8	4.6 ± 0.3	-2.1 ± 0.6	6.0 ± 0.03	1.2 ± 0.03
RB Cluster 2 (n = 22)	11.4 ± 0.2^{b}	50.8 ± 0.9	4.5 ± 0.3	-2.8 ± 0.4	5.9 ± 0.03	1.3 ± 0.03
Main effects						
Strain	$F_{50} = 6.3$*	$F_{53} = 0.4$	$F_{53} = 0.4$	$F_{53} = 2.1$	$F_{77} = 1.1$	$F_{77} = 3.2^{T}$
Cluster	$F_{50} = 6.3$*	$F_{53} = 1.1$	$F_{53} = 0.2$	$F_{53} = 3.2^{T}$	$F_{77} = 1.2$	$F_{77} = 0.4$
Strain x Cluster	$F_{50} = 0.3$	$F_{53} = 1.2$	$F_{53} = 7.7$**	$F_{53} = 2.2$	$F_{77} = 0.05$	$F_{77} = 0.4$

$^{a, b}$Means within columns lacking common superscripts differ ($P < 0.10$).
^{T}P < 0.10, *P < 0.05, **P < 0.01.

Figure. 5.1. Dendogram (graphical representation of cluster analysis) of clusters of commercial (COMM) male turkeys (n = 44) based on their latency to ambulate, number of steps and number of defecations in an open field test at 11 weeks of age. Two clusters were identified (dotted line).

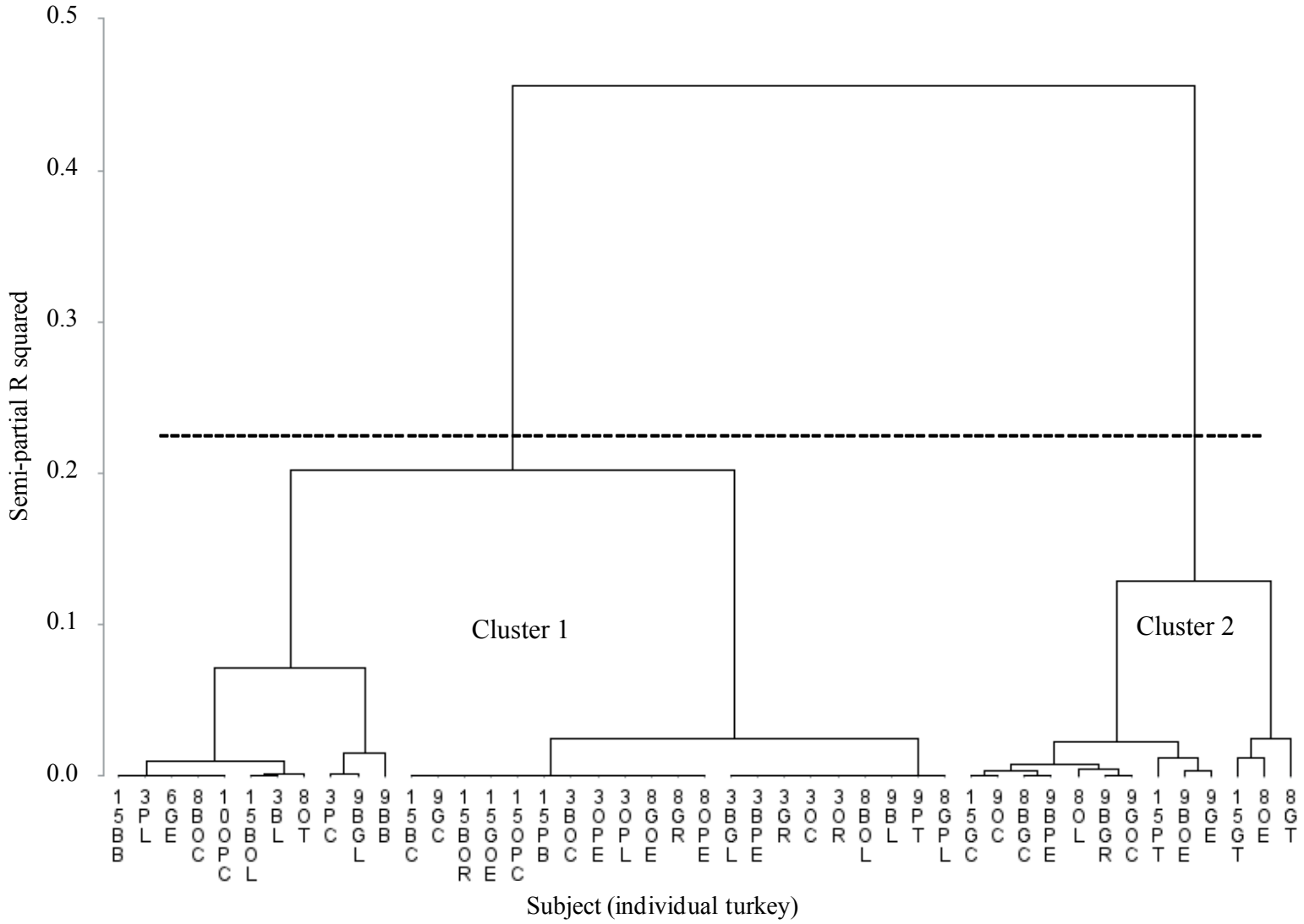


Figure. 5.2. Dendogram (graphical representation of cluster analysis) of clusters of randombred (RB) male turkeys (n = 47) based on their latency to ambulate, number of steps and number of defecations in an open field test at 11 weeks of age. Two clusters were identified (dotted line).

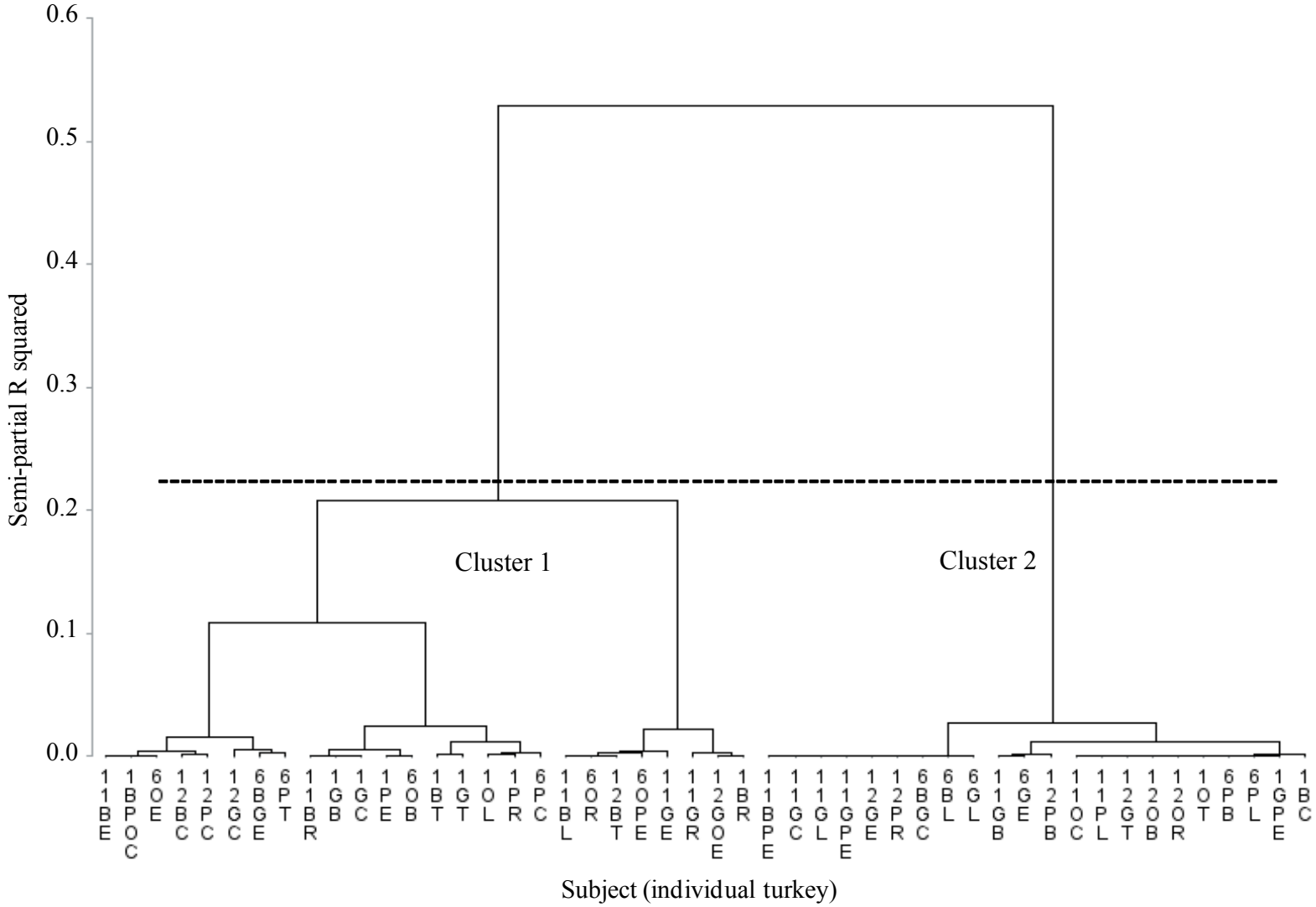
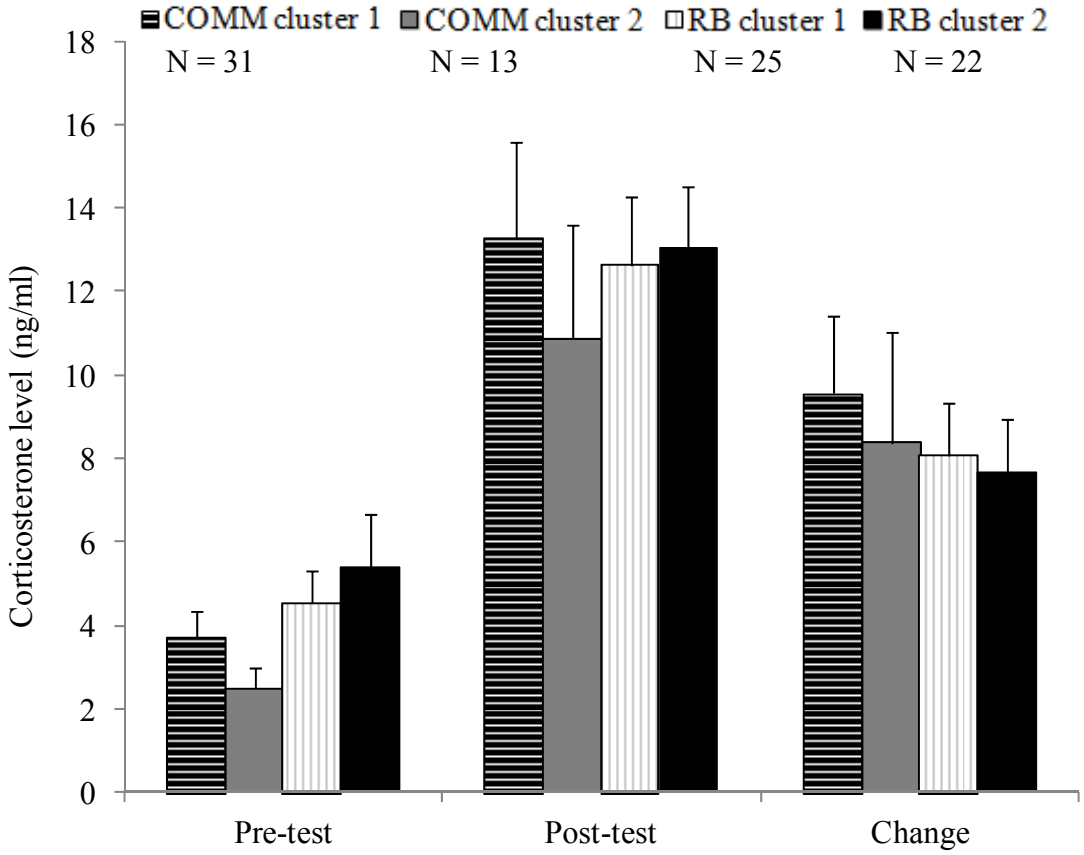


Figure. 5.3. Plasma corticosterone levels (pre-test, post-test and change from pre- to post-test) of male turkeys of a commercial (COMM) and randombred (RB) strain tested in an open field test at 11 wk of age. Turkeys were grouped into two clusters based on their responses in the open field test. Raw data (mean ± SE) are presented. Statistical analyses were conducted on log-transformed data.



REFERENCES

REFERENCES

Cafe, L. M., D. L. Robinson, D. M. Ferguson, G. H. Geesink, and P. L. Greenwood. 2011. Temperament and hypothalamic-pituitary-adrenal axis function are related and combine to affect growth, efficiency, carcass, and meat quality traits in Brahman steers. Dom. Anim. Endocrinol. 40:230-240.

Debut, M., C. Berri, E. Baeza, N. Sellier, C. Arnould, D. Guemene, N. Jehl, B. Boutten, Y. Jego, C. Beaumont, and E. Le Bihan-Duval. 2003. Variation of chicken technological meat quality in relation to genotype and preslaughter stress conditions. Poult. Sci. 82:1829–1838.

Erasmus, M. A., & Swanson, J. C. 2014. Temperamental turkeys: reliability of behavioral responses to four tests of fear. *Appl. Anim. Behav. Sci.* 157: 100-108.

Fletcher, D. L., M. Qiao, and D. P. Smith. 2000. The relationship of raw broiler breast meat color and pH to cooked meat color and pH. Poult. Sci. 79:784–788.

Forkman, B., A. Boissy, M. –C. Maunier-Salaün, E.Canali, and R. B. Jones. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol. Behav. 92: 340-374.

Huff, G., W. Huff, N. Rath, A. Donoghue, N. Anthony, and K. Nestor. 2007. Differential effects of sex and genetics on behavior and stress response of turkeys. Poult. Sci. 86:1294-1303.

Jeong, J. Y., K. K. Janardhanan, A. M. Booren, J. B. Harte, and I. Kang. 2011. Breast meat quality and consumer sensory properties of broiler carcasses chilled by water, air, or evaporative air. Poult. Sci. 90:694-700.

Martin, J., 2009. The commercialisation of british turkey production. Rural Hist. 20:209-228.

McCurdy, R. D., S. Barbut, and M. Quinton. 1996. Seasonal effect on pale soft exudative (PSE) occurrence in young turkey breast meat. Food Res. Int. 29:363–366.

McKee, S. R., and A. R. Sams. 1997. The effect of seasonal heat stress on rigor development and the incidence of pale, exudative turkey meat. Poult. Sci. 76:1616–1620.

Nestor, K. E., M. G. McCartney, and N. Bachev. 1969. Relative contributions of genetics and environment to turkey improvement. Poult. Sci. 48:1944-1949.

Owens, C. M., and A. R. Sams. 2000. The influence of transportation on turkey meat quality. Poult. Sci. 79:1204-1207.

Petherick, J. C., R. G. Holroyd, V. J. Doogan, and B. K. Venus. 2002. Productivity, carcass and meat quality of lot-fed Bos indicus cross steers grouped according to temperament. Aust. J. Exp. Agric. 42:389–98.

Petracci, M. and C. Cavani. 2012. Muscle growth and poultry meat quality issues. Nutrients. 4:1-12.

Rathgeber, B. M., M. D. Pato, J. A., Boles, and P. J. Shand. 1999. Rapid post-mortem glycolysis and delay chilling of turkey carcasses cause alterations to protein extractability and degradation of breast muscle proteins. J. Agric. Food Chem. 47:2529–2536.

Remignon, H., A. D. Mills, D. Guemene, V. Desrosiers, M. Garreau-Mills, M. Marche, and G. Marche. 1998. Meat quality traits and muscle characteristicsin high or low fear lines of Japanese quails (*Coturnix japonica*) subjected to acute stress. Br. Poult. Sci. 39:372-378.

Sarica, M., N. Ocak, S. Turhan, C. Kop, and U. S. Yamak. 2011. Evaluation of meat quality from 3 turkey genotypes reared with or without outdoor access. Poult. Sci. 90:1313-1323.

Sosnicki, A. A., M. L Greaser, M. Pietrzak, E. Pospiech, and V. Sante.1998. PSE-like syndrome in breast muscle of domestic turkeys: a review. J. Muscle Foods. 9:13-23.

Strasburg, G. M. and W. Chiang. 2009. Pale, soft, exudative turkey-the role of ryanodine receptor variation in meat quality. Poult. Sci. 88:1497-1505.

Terlouw E. M. C., and P. Rybarczyk. 2008. Explaining and predicting differences in meat quality through stress reactions at slaughter: the case of Large White and Duroc pigs. Meat Sci. 79:795–805.

Terlouw, E. M. C., and J. Porcher. 2005. Repeated handling of pigs during rearing. I. Refusal of contact by the handler and reactivity to familiar and unfamiliar humans. J. Anim. Sci. 83:1653-1663.

Terlouw, E. M. C., C. Arnould, B. Auperin, C. Berri, E. Le Bihan-Duval, V. Deiss, F. Lefèvre, B. J. Lensink, and L. Mounier. 2008. Pre-slaughter conditions, animal stress and welfare: current status and possible future research. Animal. 2:1501-1517.

Thompson, L. D., D. M. Janky, and S. A. Woodward. 1987. Tenderness and physical characteristics of broiler breast fillets harvested at various times from post-mortem electrically stimulated carcasses. Poult. Sci. 66:1158–1167.

Updike, M. S., H. N. Zerby, M. S. Sawdy, G. Kaletunc, and M. P. Wick. 2005. Turkey breast meat functionality differences among turkeys selected for body weight and/or breast yield. Meat Sci. 71:706–712.

USDA-NASS. 2014. Quick Stats. Available online: http://quickstats.nass.usda.gov/data/printable/6F84B583-B4A8-3B68-89CF-D3F8CD9BFE7B

Voisinet, B. D., T. Grandin, S. F. O'Connor, J. D. Tatuma, and M. J. Deesing. 1997. *Bos Indicus*-cross feedlot cattle with excitable temperaments have tougher meat and a higher incidence of borderline dark cutters. Meat Sci. 4:367-377.

Werner, C., J. Riegel, and M. Wicke. 2008. Slaughter performance of four different turkey strains, with special focus on the muscle fiber structure and the meat quality of the breast muscle. Poult. Sci. 87:1849-1859.

Yost, J. K., P. B. Kenney, S. D. Slider, R. W. Russell, and J. Killefer. 2002. Influence of selection for breast muscle mass on myosin isoform composition and metabolism of deep pectoralis muscles of male and female turkeys. Poult. Sci. 81:911-917.

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS

Injurious pecking is an important problem affecting the well-being of commercially farmed turkeys, yet there has been little research on this subject. The large amount of scientific literature regarding feather pecking in chickens indicates that feather pecking is a multifactorial problem, and the propensity to develop feather pecking differs between individual chickens and is related to fearfulness and the ability of the hens to cope with stress (Rodenburg et al., 2013). However, individual differences in fearfulness had not previously been studied in turkeys. Furthermore, tests used to assess fear in hens and quail have been used to draw conclusions regarding the well-being of turkeys (e.g. Noble et al., 1996 beak trimming), in spite of there being no research into turkeys' fear behavior in general, or on the reliability and validity of these tests for turkeys. Individual differences in behavior and stress reactivity are also associated with meat quality of other species (e.g. Terlouw, 2005). Research into meat quality has gained importance in the turkey industry due to the increased incidence of PSE meat (Sosnicki et al., 1998). Four studies were conducted to examine the reliability and validity of fear tests for turkeys, genetic and sex differences in feather pecking and whether feather pecking and meat quality are associated with differences in fear behavior.

Chapter 2 examined whether the same tests of fear that are commonly used for chickens and quail are reliable when used to assess fear in male commercial turkeys. The changes in turkeys' responses over time were also discussed. Responses to TI, OF and VA tests were reliable, as indicated by moderately high Spearman rank correlation coefficients. For the NO test, the only test measure that was reliable was the latency to peck the object. Reliability of test responses may have been affected by the body conformation and growth rates of the commercial domestic turkey. Indeed, behavioral responses of the commercial turkey to fear tests may be

difficult to interpret due to the changes in body conformation resulting from genetic selection for increased feed efficiency and higher body weights. For example, the larger breast muscle size of the commercial domestic turkey may impact the turkey's ability to right itself during a TI test when the turkey is tested at an older age. Therefore, there may be changes in TI responses resulting from age, development and also changes in body conformation. An important area of future research is to examine how changes in body conformation and age affect behavior.

The same data that was used in Chapter 2 was used in Chapter 3 to examine whether turkeys' responses to TI, OF, VA and NO tests were correlated at two ages (4 to 6 weeks and 8 to10 weeks). Turkeys showing extreme behavioral responses were compared across test situations to assess whether they maintained differences in their responses relative to one another. Results were rather complicated, revealing differences in test correspondence depending on the age at which turkeys were tested. At 4 to 6 weeks, ambulatory activity in the OF test tended to be correlated with the latency to peck the observer during the VA test, indicating that OF ambulation tended to increase as VA peck latency increased. At 8 to10 weeks, ambulatory activity in the OF test increased as vocalization during TI decreased, and tended to decrease as the latency to peck the object during the NO test increased. Overall, results suggested that behavioral responses to fear tests appear to be specific to the test, and therefore context-dependent because results from Chapter 3 indicated little correspondence between test measures. Fear tests may be testing different aspects or dimensions of fear and fearfulness, or they may be testing behavior that is not necessarily associated with fear per se.

A limitation of this research is that only some aspects of reliability and validity were assessed. Reliability in terms of consistency over repeated applications was assessed, but the sensitivity of test measures (whether test measures detect small changes in the true value of the

level of fear) was not examined because additional experiments, turkeys and resources would have been required. Additionally, the external validity (whether results are applicable to situations outside of the experimental environment) was not tested because this would require research to be conducted in commercial facilities. Further research is therefore needed to validate the four fear tests. The finding that the OF test was reliable in females and in randombred turkeys (Chapter 4) demonstrates that the OF test is generalizable across sexes and strains when these sexes and strains are housed and tested under the same standard conditions.

Fear tests are difficult to interpret because they may have different floor and ceiling effects. Moreover, fear tests do not necessarily measure the same motivations even though they are believed to assess some aspects of fear (discussed in Chapter 3). For example, the OF test assesses neophobia as well as social reinstatement motivation (Forkman et al., 2007). The TI test on the other hand, is an unlearned response. Although TI duration is influenced by the level of fear experienced before TI is induced, responses to OF and TI tests may differ because unlike TI, behavior in the OF test is not the terminal response to capture by a predator. The OF test is therefore not a "life or death" situation and may not be associated with the same degree of fear compared to fear induced by a predatory attack. In the VA test, turkeys were repeatedly exposed to the same observer and the latency to approach and peck the observer decreased with each test period, indicating reduced fear and/or habituation over time. In fact, most turkeys approached and pecked the observer; therefore, it is unlikely that the VA test was measuring fear per se. Arguably, the NO test was most likely to induce a fear response each time the turkeys were tested (i.e. turkeys were less likely to become habituated to NO testing) because turkeys were exposed to a different unfamiliar object each time they were tested, and the object was placed in

their home pen. However, the NO test was also the least repeatable of the four fear tests, as discussed in Chapter 2.

Caution is needed when using fear tests to draw conclusions about animal welfare. Fear tests should not be used indiscriminately because behavior appears to be situation-dependent. Within the scientific literature, there is debate as to whether fearfulness can be considered a unitary construct, or whether fearfulness has different dimensions (Boissy, 1995). If each fear test measures a different dimension of fearfulness, then it is expected that there would be less correspondence between tests than if they were all measuring the same underlying state. A lack of correlation between fear tests have led some researchers to conclude that different fear tests used for poultry are multidimensional and caution is needed when interpreting fear responses (e.g. Miller et al., 2006; Richard et al., 2008). Overall, results are in agreement with the notion that fear responses are context-specific and that fear is multidimensional. Indeed, fear tests do not all measure fear in the same way, some fear tests impose additional stress or situations that are associated with other behavior and not just fear, and fear tests differ in the degree of fear that is elicited.

Based on results from Chapters 2 and 3, the OF test was chosen to further examine the relationship between fear-related behavior, feather pecking and meat quality. The OF test was chosen for several reasons: turkeys' responses were repeatable (Chapter 2), the OF test was associated and tended to be associated with behavior in the TI and VA tests, respectively (Chapter 3), and the OF test enables birds to be tested individually and is easier to use and implement compared to the TI test.

In Chapters 4 and 5, statistical analysis confirmed that turkeys' behavioral OF responses are significantly different whether turkeys at the extreme ends of the behavior are compared

(Chapter 4) or whether turkeys are grouped together using cluster analysis (Chapter 5). These statistically significant differences between extreme behavioral phenotypes and between clusters may suggest that turkeys can be categorized into distinct coping styles. However, despite the behavioral differences, there did not appear to be an association between behavior and physiological responses to the OF test because corticosterone levels measured before and after OF testing did not differ between turkeys of different clusters (Chapter 5). Similar results have been reported when behavioral and physiological responses were compared between HFP and LFP chicken lines (reviewed in Groothuis and Carere, 2005). This research therefore could not determine whether coping styles exist in turkeys as they do in other species, and results raise more questions about the link between behavior and physiology. Furthermore, the relationship with other behavior such as aggression needs to be assessed in order to develop a clear picture of turkey temperament.

In contrast to research with laying hens (Rodenburg et al., 2004; reviewed in Rodenburg et al., 2013), a clear relationship between feather pecking and OF responses was not found. Most studies examining fearfulness and feather pecking in laying hens uses the HFP and LFP selection lines. Therefore, it may be more likely that relationships will be found between feather pecking and fearfulness in selection lines, because variability is reduced in lines that are genetically similar. Although feather pecking did not appear to be associated with OF behavior, Chapter 4 provides new information regarding feather pecking of turkeys. Specifically, it was interesting to learn how many individuals were observed to perform feather pecking, and that the percentage of birds that pecked did not differ between sexes or strains. However, no distinction was made between turkeys that performed gentle vs. severe pecking, and this is an important avenue for

future research because the two forms of feather pecking may not be related (see Rodenburg et al., 2013).

Chapter 5 examined whether OF behavior was related to meat quality in commercial and randombred male turkeys. Strains tended to differ in the rate at which rigor mortis developed, but no other meat quality differences were found. In contrast to research with other species (e.g. Terlouw, 2005), there was no relationship between physiological or behavioral fear responses and meat quality. However, turkeys were not specifically subjected to an intense stressor before slaughter, and it is likely that pre-slaughter stress may result in observable differences in meat quality. Further research is needed to examine whether physiological responses to fear testing are comparable to physiological responses resulting from other stressors, and whether pre-slaughter stress is higher, and meat quality poorer, in turkeys showing increased physiological stress responses.

The studies presented in this dissertation provide new insights into turkeys' behavioral and physiological fear responses and feather pecking behavior and a number of avenues for future research have been identified. There appear to be distinct individual differences in behavior that are correlated over time, but the behavior depends on the situation because it is not strongly correlated to behavior in other tests. Very little is known about turkey temperament and individual differences in behavioral and physiological stress responses, which presents an interesting area for future research. Furthermore, the absence of a relationship between physiological and behavioral responses indicates that more research is needed to understand how individual animals respond to stress and whether stress reactivity is more likely to be associated with meat quality if animals are stressed immediately prior to slaughter. Feather pecking remains a complex problem that is poorly understood, but results reveal that there are strain differences in

the severity of feather pecking damage. Further investigation of genetic differences in the behavior of the recipient of feather pecking, and of the possible genetic differences in feather cover and skin integrity may provide a means to reduce feather pecking damage.

REFERENCES

REFERENCES

Boissy, A. Fear and fearfulness in animals. 1995. Q. Rev. Biol. 70: 165-191.

Groothuis, T. G. G., and C. Carere. 2005. Avian personalities: characterization and epigenesis. Neurosci. Biobehav. Rev. 29: 137-150.

Forkman, B., A. Boissy, M. –C. Maunier-Salaün, E. Canali, and R. B. Jones, R. B. 2007. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol. Behav. 92: 340-374.

Miller, K. A., J. P. Garner, and J. A. Mench. 2006. Is fearfulness a trait that can be measured with behavioural tests? A validation of four fear tests for Japanese quail. Anim. Behav. 71:1323-1334.

Noble, D. O., K. E. Nestor, and K. K. Krueger. 1996b. The effect of beak trimming on two strains of commercial turkeys. 2. Behavior traits. Poult. Sci. 75: 1468-1471.

Norris, D., J.W. Ngambi, M. Mabelebele, O.J. Alabi and K. Benyi. 2014. Genetic selection for docility: a review. J. Anim. Plant Sci. 24: 374-379.

Richard, S., N. Wacrenier-Ceré, D. Hazard, H. Saint-Dizier, C. Arnould, and J. M. Faure. 2008. Behavioral and endocrine fear responses in Japanese quail upon presentation of a novel object in the home cage. Behav. Process. 77: 313-319.

Rodenburg, T. B., A. J. Buitenhuis, B. Ask, K. A. Uitdehaag, P. Koene, J. J van der Poel, J. A. M. van Arendonk, and H. Bovenhuis. 2004. Genetic and phenotypic correlations between feather pecking and open-field response in laying hens at two different ages. *Behav. Genet.* 34: 407-415.

Rodenburg, T. B., M. M. van Krimpen, I. C. de Jong, E. N. de Haas, M. S. Kops, B. J. Riedstra, R. E. Nordquist, J. P. Wagenaar, M. Bestman, and C. J. Nicol. 2013. The prevention and control of feather pecking in laying hens: identifying the underlying principles. Worlds Poult. Sci. J. 69: 361-373.

Sosnicki, A. A., M. L Greaser, M. Pietrzak, E. Pospiech, and V. Sante.1998. PSE-like syndrome in breast muscle of domestic turkeys: a review. J. Muscle Foods. 9: 13-23.

Terlouw, C. 2005. Stress reactions at slaughter and meat quality in pigs: genetic background and prior experience - A brief review of recent. Livest. Prod. Sci. 94: 125-135.