Mood swings in cows

Cognitive and
physiological assessments

Louise Kremer

Propositions

1. The judgement bias task is biased.
   (this thesis)

2. Because every cow has a unique personality, it is impossible to make all cows happy.
   (this thesis)

3. Eating insects is ethically worse than eating beef.

4. Scientists are experts in objective writing with hidden subjective meaning.

5. Even the most inscrutable individuals reveal their true self during boardgames.

6. In a system that rewards overwork, offering a massage is a plaster on the symptoms rather than a cure for the disease.

Propositions belonging to the thesis, entitled

Mood swings in cows: cognitive and physiological assessments

Louise Kremer
Wageningen, 25 April 2022

Mood swings in cows

Cognitive and
physiological assessments

Louise Kremer

Thesis committee

Promotor

Dr E.A.M. Bokkers
Associate Professor, Animal Production Systems group
Wageningen University & Research

Co-promotors

Dr L.E. Webb
Assistant Professor, Animal Production Systems group
Wageningen University & Research

Dr C.G. van Reenen
Senior researcher, Animal Production Systems group & Wageningen Livestock Research
Wageningen University & Research

Other members

Dr L. Gygax, Humboldt-Universität zu Berlin, Germany
Prof. Dr B. Kemp, Wageningen University & Research
Prof. Dr M.T. Mendl, University of Bristol, United Kingdom
Dr R.E. Nordquist, Utrecht University

This research was conducted under the auspices of the Graduate School of Wageningen Institute of Animal Science (WIAS)

Mood swings in cows

Cognitive and physiological assessments

Louise Kremer

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 25 April 2022
at 11 a.m. in the Aula.

Louise Kremer
Mood swings in cows: cognitive and physiological assessments, 250 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summaries in English and French

ISBN 978-94-6447-084-0
DOI https://doi.org/10.18174/562024

To my husband,
who knows as much as I do about this thesis

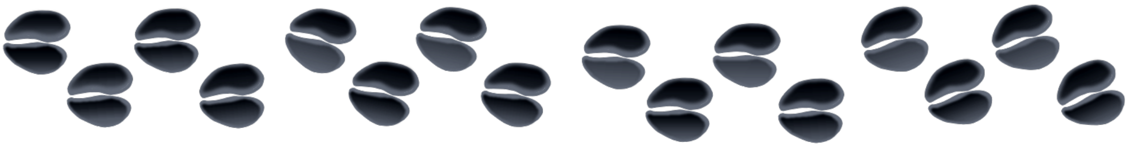
Abstract

In Europe, the welfare issues inherent to dairy cow farming are acknowledged both by society and the public authority. Valid assessments of dairy cow welfare are required, above all by the determination of mood valence. For sentient beings like cows, mood valence constitutes a fundamental determinant of welfare as it codes information on the intrinsic pleasantness or aversiveness of long-term, integrated animal subjective experiences. Recently, cognitive indicators of mood valence have been proposed but their methods of evaluation are too time-consuming to allow for on-farm assessments of dairy cow welfare. In contrast, the potential of physiological measures as valid indicators of mood valence has been overlooked, while physiological assessments of mood may be compatible with the practical monitoring of cow welfare. The aim of this thesis was to identify physiological indicators of mood valence that would be compatible with practical assessments of welfare in commercial dairy farms. To achieve this goal, we first conducted an experiment that aimed at identifying a valid and relatively feasible Judgement Bias Task (JBT) for dairy cows. The JBT is a cognitive tool commonly used to evaluate judgement bias, a known indicator of animal mood valence that can serve as a validation measure for the identification of more practical physiological indicators of mood valence. Among three JBTs that differed in terms of punishers (i.e. either an absence of reward, an air-puff or an electric shock), we identified the JBT associated with the air-puff as the most feasible and sensitive method of judgement bias assessment. In a second experiment, we attempted to develop a model of mood valence based on successive housing manipulations to induce positive and negative mood in cows. To validate this model, we compared judgement biases before and after manipulating the housing conditions. Against expectations, housing did not influence judgement biases – which may suggest that the model failed at inducing the expected mood shifts. Alongside judgement biases, non-invasive physiological measures – hair cortisol, heart rate variability and milk-derived measures – were also assessed. Their validity as indicators of mood valence was examined by exploring their correlations with judgement biases. None of the physiological measures, however, was found to consistently correlate with judgement biases – suggesting that they do not constitute valid indicators of mood valence. Nonetheless, daily milk fluctuations increased in cows characterized as fearful when they were exposed to the aversive housing manipulations. Milk fluctuations may hence constitute a promising physiological indicator of negative mood in dairy cows that is compatible with practical assessments of welfare.

Table of contents

Chapter 1

General introduction



Background

In western countries, farm animal welfare has long been under public scrutiny (Alonso et al. 2020). In the eye of society, current welfare standards are unsatisfactory, and consumers express the need for products derived from more humanely raised (non-human) animals (Broom 2017). Farm animals should no longer have a life ‘worth living’, but instead experience a ‘good life’ (Green and Mellor 2011). However, European governments have remained deaf to the societal demand for the promotion of positive experiences in farm animals. Current directives just lay down the minimum animal welfare requirements and only forbid the most unethical practices (Bonafos et al. 2010). Asides from being deemed insufficient (Pejman et al. 2019), such directives do not even exist for all livestock groups. Dairy cow welfare, for instance, is not regulated by any specific legislation (Nalon and Stevenson 2019), whereas it is considered as the second greatest animal welfare problem by the European Parliament (Broom, 2017). An assessment tool of dairy cow welfare is, consequently, required to objectively establish the seriousness of the situation. By allowing to monitor dairy cow welfare, such a tool would furthermore permit the identification of farm practices effective at improving dairy cow welfare – thereby potentially leading to better dairy cow welfare (Leliveld and Provolo 2020).

For a valid evaluation of dairy cow welfare, welfare assessments must align with the current definition of animal welfare. Although disputed (Hewson 2003; Robbins et al. 2018), most researchers now support the hedonistic view of animal welfare (Robbins et al. 2018), which exclusively focuses on affective states. The term affective state describes the multifaceted psychological phenomenon that includes changes in individual subjective experience, cognition, behaviour and physiology (Paul et al. 2005). Typically, affective states are conceptualised as phenomena that are necessarily valenced (i.e. pleasant or unpleasant) and which can vary in arousal/activation as well as in duration (Mendl et al. 2010b). Following the theory of welfare hedonism, animal welfare is considered to be optimal when the balance between positive and negative affective states is overall significantly positive (Green and Mellor 2011). Adequate assessment of dairy cow welfare must hence allow for valid evaluation of affective states.

Various factors influence affective states and must be taken into consideration when assessing farm animal welfare. Such factors can be classified into two distinct categories: the exogenous factors, which most often represent manipulable features of the environment, and the endogenous factors, which relate to the individual intrinsic characteristics. In farm species, the housing conditions in which animals are raised constitute a prominent exogenous factor that influences affective states since previous studies in calves and pigs have shown that improved housing conditions are associated with more positive affective states (Bučková et al. 2019; Douglas et al. 2012), while worsened housing conditions in sheep are associated with more negative affective states (Destrez et al. 2013). In contrast, personality, that is a correlated set of individual behavioural and physiological traits that are consistent over time and context (van Reenen, 2012), constitutes a crucial endogenous factor that influences affective states as it modulates the individual perception of its environment (Asher et al. 2016; Roelofs et al. 2016; Lecorps et al. 2018). Housing and personality are, thus, two examples of major determinants of animal affective states – the influence of which must be carefully considered when studying dairy cow affective states.

Among the different psychological phenomena encompassed by the term 'affective states' (Paul et al. 2005; Quigley et al. 2014), the study of moods over emotions may be of primary importance in the context of animal welfare. Whereas emotions represent intense and short-term affective states that consist of adaptive responses to specific stimuli, moods represent less intense but relatively longer-lasting affective states, which are thought to result from an accumulation of emotions (Mendl et al. 2010b; Nettle and Bateson 2012). Moods, therefore, consist of affective states that persist over time, which are not directed at a specific stimulus – but instead represent the cumulative impact of several stimuli on the animals' psychological experiences (Eldar et al. 2016; Raoult et al. 2017; Bliss-Moreau and Rudebeck 2021). Timewise, moods can also be self-sustaining as they may influence individuals' subsequent emotions (Raoult et al. 2017). This idea is supported by previous studies demonstrating that positive moods may buffer negative emotions – and even amplify positive emotions on certain occasions – (Neumann et al. 2001; Reefmann et al. 2012; van Steenbergen et al. 2021), while negative moods may worsen negative emotions (Mendl et al. 2010b; Groenewold et al. 2013; Clarkson et al. 2020). In the context of animal welfare studies, moods may hence be of greater relevance than emotions as they provide an integrated picture of the affective experiences over prolonged periods – thereby echoing the notion of animal welfare from a temporal perspective. Moods may also be of greater practical relevance than emotions. While farmers cannot control for every emotion experienced by their animals, they still have the opportunity to create an environment that can favour positive emotions and minimise negative ones. Consequently, dairy farmers have the possibility of enhancing their animals' mood in practice – even though cows will inevitably experience negative emotions every now and then. Based on these considerations, assessments of dairy cow affective states should thus primarily focus on evaluating mood over emotion.

Furthermore, assessments of dairy cow mood should aim at detecting shifts in valence. As previously mentioned, moods (and other affective states) are described in a two-dimensional framework according to their valence and their arousal (Mendl et al. 2010b). Valence characterises the pleasantness/positivity or the aversiveness/negativity of mood, while arousal represents its degree of activation (Mendl et al. 2010b). From a welfare standpoint, the valence axis is hence of particular interest as it conveys information relative to the hedonic quality of individuals' subjective experience. To ensure the adequate assessment of dairy cow mood, the identification of indicators of valence is, hence, crucial.

In practice, the subjective experience of mood cannot be directly assessed in animals that are incapable of verbal speech, but it can be inferred from the measurable components of mood – that is from the behavioural, cognitive and physiological components (Mendl et al. 2010b). Although research on the behavioural facet of mood has considerably advanced our understanding of animal welfare (e.g. Murphy et al. 2014), the reliability of behavioural assessments of mood has been criticised (Paul et al. 2005). In consequence, this thesis will primarily focus on the cognitive and physiological indicators of mood.

To date, considerable research has been conducted with the aim of identifying cognitive indicators of mood valence. Traditionally, mood valence is inferred from measures of cognitive biases – which reflect the influence of affective processes on cognitive mechanisms. In animal welfare studies, two types of cognitive biases have been given special consideration, namely the judgement and attention biases (Paul et al. 2005).

Judgement bias represents the most commonly used indicator of mood valence in affective sciences – even though it is unsuitable for on-farm assessments. Judgement bias reflects affect-driven shifts in the interpretation of ambiguous stimuli (Eysenck et al. 1991; Mendl et al. 2009; Roelofs et al. 2016). Like humans, animals in a positive mood are expected to interpret ambiguous information positively, hence to be more optimistic and to display more positive judgement biases, than animals in more a negative mood

– and vice versa (Harding et al. 2004). In practice, judgement bias is assessed using a Judgement Bias Task (JBT), during which animals are trained to discriminate between two stimuli that are associated with a relatively positive and a relatively negative reinforcer, respectively. Once trained, animals are then exposed to ambiguous stimuli and their behaviour is recorded to measure judgement biases – behaviours associated with the anticipation of a relatively positive outcome reflecting more positive judgement bias (i.e. better affective states) than behaviours associated with the anticipation of a relatively negative outcome. Judgement bias is commonly accepted as a valid measure of mood valence, since positive and negative manipulations of mood have repeatedly been associated with the expected shift in judgement biases in a wide range of species (Lagisz et al. 2020; Neville et al. 2020). Although considered valid, the suitability of the judgement bias as an on-farm indicator of mood valence has, however, been criticised – mainly because JBT usually necessitates extensive periods of training (Roelofs et al. 2016). Judgement biases, hence, constitute a valid – yet unpractical – indicator of mood valence, which can serve as a reference measure to further investigate the validity of more practical candidate indicators of mood valence.

Recently, attention bias has been proposed as a more practical alternative to the judgement bias for measuring animal mood valence (Lee et al. 2016). Attention bias reflects affect-driven shifts in the individuals' allocation of attention towards several salient stimuli (MacLeod et al. 1986; Paul et al. 2005). Like humans, animals in a negative mood are expected to display heightened attention toward threatening stimuli compared to individuals in relatively better affective states (Mogg et al. 1995; Roy et al. 2009). In practice, attention bias is assessed using an Attention Bias Task (ABT), during which animals are often simultaneously exposed to a rewarding and a threatening stimuli (Lee et al. 2016, 2017; Monk et al. 2018b, 2019b). For attention bias assessments, several measures are typically recorded such as the time spent by the animals looking at – or interacting with – either stimulus and their vigilance behaviours (e.g. Lee et al. 2016, 2017). In ruminants, attention bias is a well-established measure of anxiety states since anxiety has repeatedly been associated with increased time spent looking at the threat and sustained vigilance (Lee et al. 2016, 2017; Monk et al. 2018a). Provided that the attention bias is sensitive to negative experiences other than anxiety and allows to discriminate positive from negative moods, attention bias could, hence, effectively constitute a more practical alternative to the judgement bias for researchers – considering that its assessment does not require training. In practice, however, the assessment of attention bias may also be too time-consuming for ABT to be used as an on-farm tool, since its implementation requires dedicated facilities and its analysis relies on extensive behavioural observations. Indicators of mood valence that are compatible with on-farm assessments must, hence, be identified to ensure their successful implementation in commercial dairy farms.

Although the potential of physiological measures as indicators of mood valence has been given less attention compared to cognitive and behavioural measures, physiology might be key to developing valid on-farm assessments of mood valence. To date, the investigation of physiological indicators of mood valence may have been hampered by the idea that physiological measures primarily reflect arousal over valence. However, the fact that certain physiological measures, particularly cardiovascular ones, effectively constitute indicators of arousal (e.g. heart rate: Wascher 2021) does not necessarily imply that the same holds for each and every physiological measure. Beyond the cardiovascular sphere, physiological measures also appertain to the neuroendocrine, immune and metabolic systems – the regulation of which appears to be influenced by long-term processes such as negative and positive moods (McEwen 2003; van Steenbergen et al. 2021). Certain physiological measures may, thus, constitute valid indicators of mood valence. In addition to potentially being valid indicators of mood valence, these physiological measures may also be appropriate for on-farm assessments. In the same way that myriads of physiological measures exist, various methods of physiological assessments, which vary in terms of implementation complexity depending on the measure of interest, also exist. Nowadays, numerous physiological parameters can be easily measured through the non-invasive collection of specific biofluids (e.g. hair, milk) or using sensor technology (e.g. heart rate sensor) without necessarily disturbing individuals' routine activities. Such physiological measures could, hence, allow for the repeated measurements of mood valence – thereby allowing farmers to conveniently monitor the welfare of their dairy cows.

Based on these considerations, the main goal of my PhD thesis was to identify valid and practical physiological indicators of mood valence in dairy cows. To achieve this goal, an incremental approach based on six steps was developed (**Figure 1**).

Outline of the thesis

The first step consisted of the synthesis of existing methodologies to assess animal mood valence, with as aim the identification of the most promising indicators of affective states with their pros and cons (**Chapter 2**). The findings from this review stimulated the research conducted in the subsequent chapters.

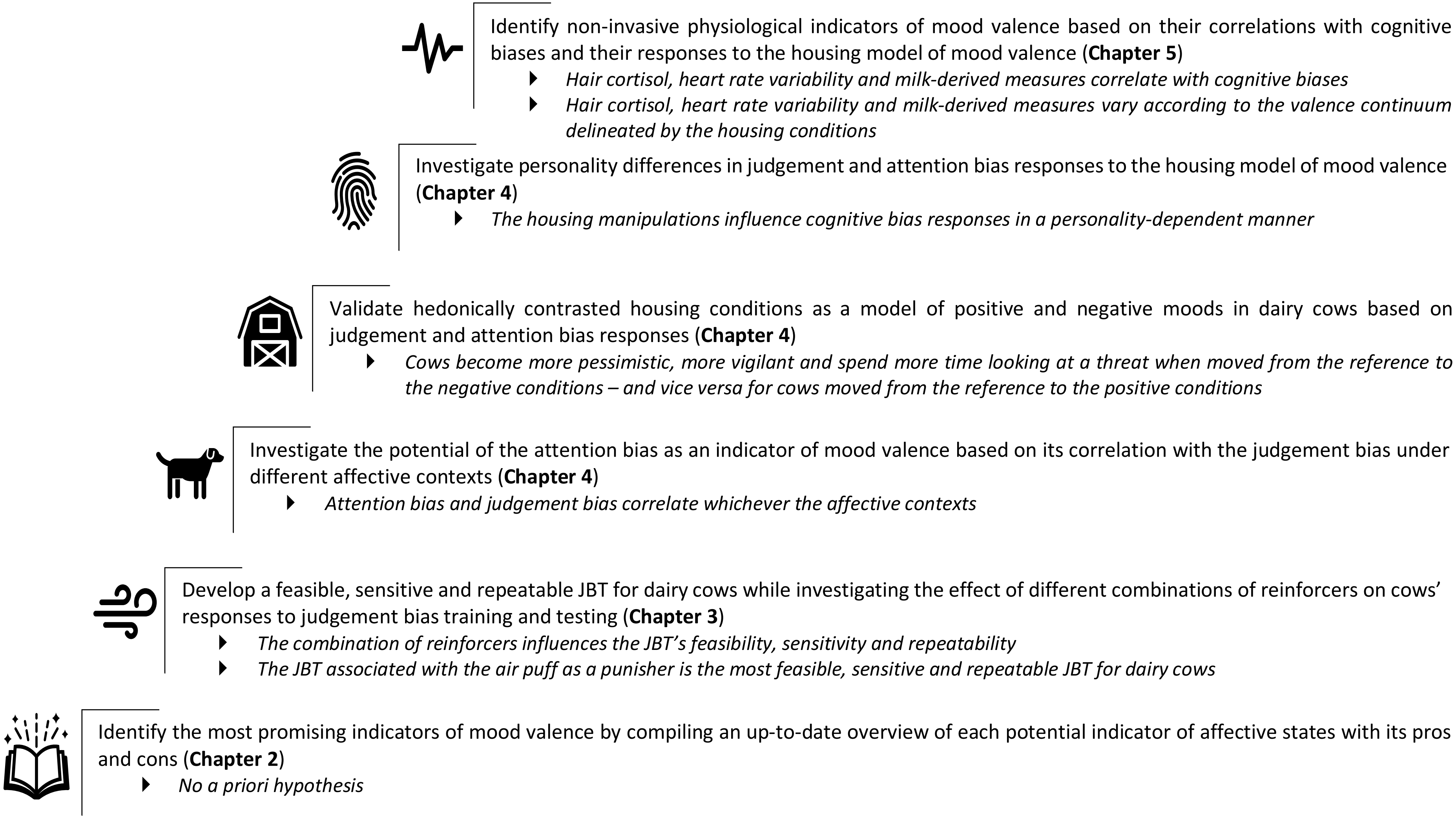
The second step consisted of the development of a JBT specifically designed for dairy cows (**Chapter 3**) to ensure that their mood valence could be appropriately assessed. At the beginning of this research project, no JBT had ever been designed for adult cows and a reference methodology allowing to further validate indicators of mood valence that are compatible with on-farm assessments was consequently lacking. The first experiment was, thus, dedicated to the elaboration of a relatively feasible and sensitive JBT suitable for adult dairy cows. To this end, three JBTs varying in terms of combination of reinforcers were designed – based on the idea that the combination of reinforcers could potentially influence both the animals' discrimination training (Avarguès-Weber et al. 2010) and the JBT's sensitivity to mood shifts (Mendl et al. 2009; Roelofs et al. 2016). The objective was to identify the most feasible and sensitive JBT among the three, to then use it as a reference methodology to assess mood valence in dairy cows.

The third step consisted of the development of an ABT suitable for dairy cows with the aim to provide researchers with a methodological alternative to the JBT that would allow for a more practical (but still potentially not on-farm compatible) assessment of mood valence in dairy cows (**Chapter 4**). To this end, an ABT suitable for dairy cows was developed based on existing set-ups in sheep and beef cattle (Lee et al. 2016, 2017). The objective was to assess the validity of the ABT outcomes as indicators of mood valence by investigating the correlations between measures of attention bias and judgement bias across different affective contexts.

The fourth step consisted of the validation of a putative model of mood valence based on housing manipulations (hereafter referred to as the 'housing model') to ensure that positive and negative mood shifts could effectively be induced in dairy cows (**Chapter 4**). Although successfully developed in other livestock species (Bučková et al. 2019; Douglas et al. 2012), no housing model of mood had yet been validated in cows – despite abundant literature linking housing conditions to cow welfare. In a second experiment, a promising model of mood valence for cows was established based on three main housing elements that are suggested to influence cow welfare: the stocking density (Fregonesi et al. 2007; Schütz et al. 2015; Winckler et al. 2015), the social stability (Schirmann et al. 2011; Wilcox et al. 2013; Gutmann et al. 2015) and the level of enrichment (Devries and Keyserlingk 2006; Huzzey et al. 2006; McConnachie et al. 2018). In practice, the experiment consisted of a longitudinal study where cows were first housed under stable reference conditions before being housed under supposedly weekly-improved or weekly-worsened conditions to elicit the intended mood shifts. The objective was to validate the housing-induced mood shifts by investigating cows' responses to the newly developed JBT and ABT both before and after the application of the housing treatment.

The fifth step consisted of investigating personality effects on cows' responses to the JBT and to the ABT, when individuals were housed under different affective contexts (**Chapter 4**). The objective of this additional step was two-folds. First, we wanted to explore the possibility that personality influences cow perception of the cognitive bias tasks *themselves*, since a previous study in calves demonstrated that individuals characterised as fearful were more pessimistic than less fearful individuals – even though all calves were housed under similar housing conditions. Second, we wanted to determine whether personality-based differences in cow perception of their housing conditions existed, considering that dairy cows have a multifaceted personality that can mediate their subjective experiences of events (Koolhaas et al. 1999; Finkemeier et al. 2018; Roelofs et al. 2016).

Finally, the sixth step consisted of identifying non-invasive physiological indicators of mood valence with the hope of providing farmers with a practical tool allowing them to monitor dairy cow welfare (**Chapter 5**). Here, we focused on three categories of non-invasive physiological measures that had previously been identified as promising indicators of affective valence based on human and animal literature, i.e. hair cortisol, heart rate variability indices, and milk-derived measures. Our final goal was to validate these measures as indicators of affective states, by investigating their correlations with measures of judgement and attention biases and the interactive effects of contrasted housing conditions and dairy cow personality on the different physiological measures.



Identify non-invasive physiological indicators of mood valence based on their correlations with cognitive biases and their responses to the housing model of mood valence (**Chapter 5**)

- *Hair cortisol, heart rate variability and milk-derived measures correlate with cognitive biases*
- *Hair cortisol, heart rate variability and milk-derived measures vary according to the valence continuum delineated by the housing conditions*

Investigate personality differences in judgement and attention bias responses to the housing model of mood valence (**Chapter 4**)

- *The housing manipulations influence cognitive bias responses in a personality-dependent manner*

Validate hedonically contrasted housing conditions as a model of positive and negative moods in dairy cows based on judgement and attention bias responses (**Chapter 4**)

- *Cows become more pessimistic, more vigilant and spend more time looking at a threat when moved from the reference to the negative conditions – and vice versa for cows moved from the reference to the positive conditions*

Investigate the potential of the attention bias as an indicator of mood valence based on its correlation with the judgement bias under different affective contexts (**Chapter 4**)

- *Attention bias and judgement bias correlate whichever the affective contexts*

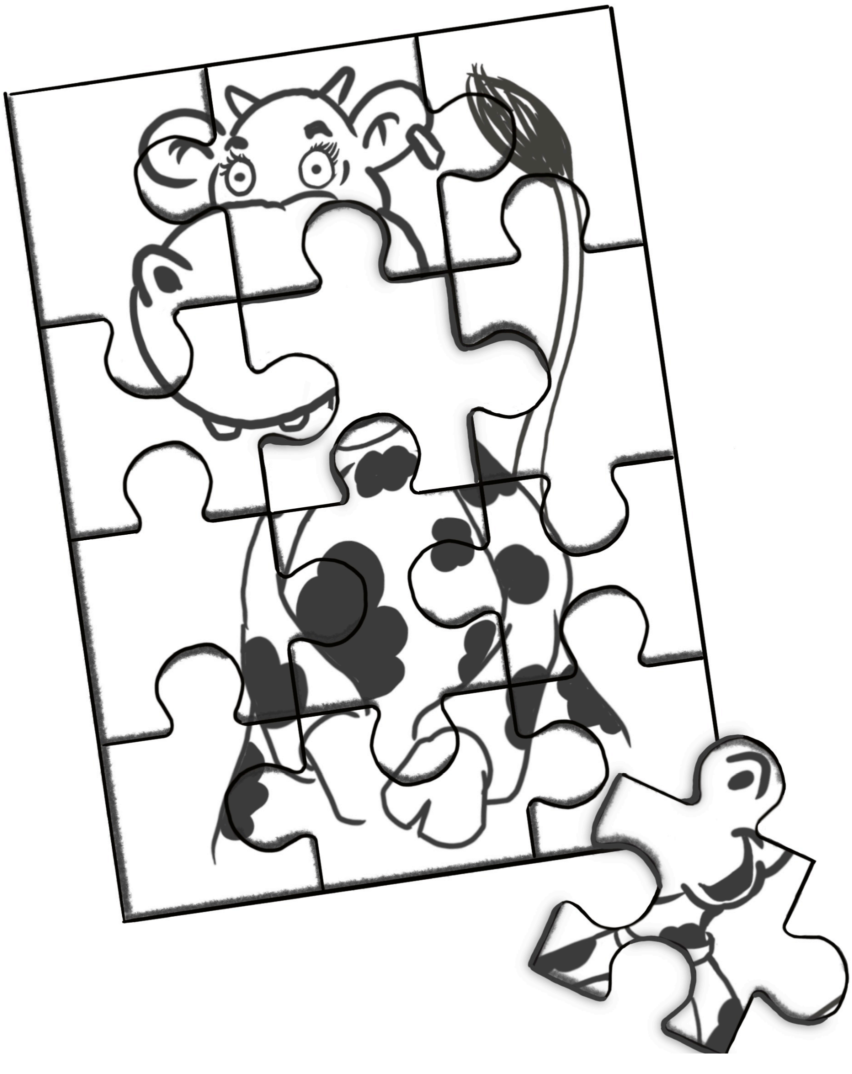
Develop a feasible, sensitive and repeatable JBT for dairy cows while investigating the effect of different combinations of reinforcers on cows' responses to judgement bias training and testing (**Chapter 3**)

- *The combination of reinforcers influences the JBT's feasibility, sensitivity and repeatability*
- *The JBT associated with the air puff as a punisher is the most feasible, sensitive and repeatable JBT for dairy cows*

Identify the most promising indicators of mood valence by compiling an up-to-date overview of each potential indicator of affective states with its pros and cons (**Chapter 2**)

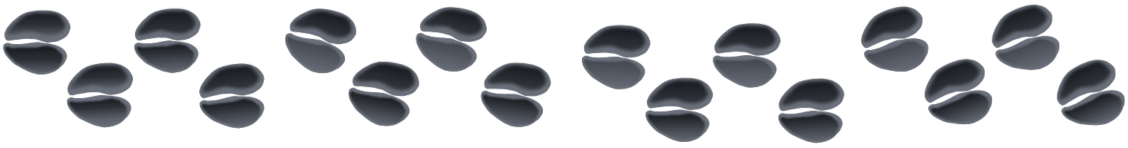
- *No a priori hypothesis*

Figure 1. Stepwise approach developed in this PhD thesis to identify valid and practical physiological indicators of mood valence in dairy cows. For each step, the main objectives are written in roman style, while the hypotheses are specified in italic. JBT: Judgement Bias Task. ABT: Attention Bias Task.



Chapter 2

The nuts and bolts of animal emotion



Neuroscience and Biobehavioural Reviews (2020) 113, 273-286

https://doi.org/10.1016/j.neubiorev.2020.01.028

L. Kremer[1], S.E.J. Klein Holkenborg[1], I. Reimert[2], J.E. Bolhuis[2], L.E. Webb[2]

[1] Animal Production Systems Group, Department of Animal Sciences, Wageningen University & Research, PO Box 338, 6700AH Wageningen, the Netherlands

[2] Adaptation Physiology Group, Department of Animal Sciences, Wageningen University & Research, PO Box 338, 6700AH Wageningen, the Netherlands

Abstract

The study of animal emotion, as with its human equivalent, can be confusing due to the complicated and inconsistent use of terminology, and the number of interlinked fields and topics it encompasses. With this review, we aim to provide an up-to-date and, to the best of our knowledge, complete overview of the field of animal emotion, especially intended for new-comers to the field who wish to get a grasp of this field. We start by tackling the terminology and proposing definitions of commonly used terms, and present the different frameworks used for the study of animal emotion. Here, we heavily draw from human literature, as the definitions of animal emotion are derived originally from human research. We follow-up with an overview of current methodologies for the study of animal emotion, in particular the valence dimension of emotion, and include some of the associated limitations linked to these methodologies. We end by pointing out key areas for future research.

Keywords

Emotion – Animal – Affect – Valence – Welfare

Introduction

Interest in the emotional lives of non-human animals (hereafter animals) has grown in the past three decades. Some of this interest stems from the increasing public concern for the welfare of captive and domesticated animals (Cornish et al. 2016). Many animal welfare scientists today agree that the 'feelings' of animals, or animal emotions, are at the core of animal welfare science (Dawkins 1988, Duncan 1996, Fraser 2008). Interest in animal emotion has also emerged in many other fields ranging from (evolutionary) zoology (e.g. De Waal 2008, Soltis et al. 2009), to affective and social neuroscience (e.g. Panksepp 2004) and psychopharmacology (e.g. Hinchcliffe et al. 2017). Understanding how emotional experiences manifest themselves, how these experiences can be indirectly assessed in non-verbal beings and which animal species are likely to experience emotion is crucial to our understanding of animals, as well as humans. Several reviews on the topic of animal emotion already exist (Bekoff 2000, Désiré et al. 2002, Mendl et al. 2010b, De Waal 2011, Makowska and Weary 2013, Anderson and Adolphs 2014, Murphy et al. 2014, de Vere and Kuczaj 2016, Perry and Baciadonna 2017), but most of them focus on specific topics within the field. The aim of the present review is to provide an up-to-date and as complete as possible overview of the current knowledge in the field of animal emotion. With this review we aim to support scientists starting in this field to grasp the basics and efficiently acquire a complete overview of current developments in this area. We start by exploring the terminology and concepts linked to emotion: what do emotional terms mean, what is the function of emotion, and how can emotion be conceptualised and categorised into various frameworks? Here we draw heavily from human literature because definitions in humans and animals are similar, and animal scientists typically make use of human research. Second, we present promising methodologies for the assessment of animal emotion, based on either behaviour, cognition or physiology. We end by pointing out potential gaps in animal emotion research which warrant future attention.

Part 1 – understanding emotion

Definitions of the terms emotion, mood and affect

Use of the terms 'emotion', 'mood', and 'affect' in both human and animal literature is inconsistent. Definitions are not systematically given by authors and when definitions are provided, these often differ between authors (de Vere and Kuczaj 2016). Absence of agreement on what these terms refer to has obstructed progress in this field (Paul and Mendl 2018) and is one reason for the debate regarding which, if any, animal species can be said to experience emotion (de Vere and Kuczaj 2016).

Hebb (1946) defined emotions as ‘certain neurophysiological states, inferred from behaviour, about which little is known except that by definition they predispose toward certain specific kinds of action’. Other more recent definitions of emotion(s) include ‘states elicited by rewards and punishments, including changes in rewards and punishments’ (Rolls 2000), ‘psychological phenomena that help in behavioural management and control’ (Bekoff 2000), ‘an intense but short-lived affective response to an event that is materialised in specific body changes’ (Désiré et al. 2002), ‘a process that facilitates appropriate responses to a wide range of both internal and environmental situations’ (Parr and Waller 2006), ‘something that moves one’s body and mind’ (Veissier et al. 2009), and ‘a temporary state brought about by biologically relevant external stimuli, marked by specific changes in the organism’s body and mind’ (De Waal 2011). Frequently, emotion is simply considered as an internal state which intervenes between perceived stimuli and subsequent responses, resulting in tendencies to engage in certain behaviours (e.g. Hinde 1985, Waller and Micheletta 2013). For example, De Waal (2011) states that ‘emotions potentiate action’. In human research, the notions of emotion and motivation are intertwined (Berridge 2018), and some consider emotion as, in part, a goal-achieving motivation that explains the drive for certain behaviours (Fanselow 2018). However, there is in fact no consensus on the causal direction of the link between emotion and behaviour (see section 2.4). Nevertheless, these definitions do include some common features, as proposed for example by Paul and Mendl (2018) for animals: an emotion is a multicomponent (subjective, physiological, behavioural and cognitive) response to a stimulus or event that is typically of importance to the individual, it is always valenced (pleasant or unpleasant) and can vary in activation/arousal and duration/persistence. The subjective, physiological, behavioural and cognitive components of emotion are described below in section 2.4. The duration of an emotion is somewhat understudied in both humans and animals, but based on human literature may span from seconds to days (Wallbott 1986, Fitness and Fletcher 1993, Gilboa and Revelle 1994).

Berkowitz (2000) wrote of mood: ‘It is an affective state that typically is fairly long-lasting, often at a relatively low or moderate level of intensity, and generally objectless and free-floating’. In contrast to shorter-term emotion, mood has indeed been described as occurring without being directed at a particular object, stimulus or event (Russell 2003), which explains its frequent characterisation as 'free floating' (Trimmer et al. 2013). According to this view, emotion likely involves more information processing (e.g. appraisal of an object) than mood (Frijda 1986). Mood, in both humans and animals, has been described as the outcome of the accumulation of short-term emotional experiences, resulting in a 'running mean' of positions occupied across scales of valence and arousal over time (Mendl et al. 2010b, Nettle and Bateson 2012, Trimmer et al. 2013). If mood is an accumulation of environmentally-triggered emotions, it is also, albeit indirectly, affected by the environment. Mood has furthermore been conceptualised as a background baseline to which individuals fall back in the absence of acute emotion (Nettle and Bateson 2012), and as a predisposition to act in certain ways (Trimmer et al. 2013). Being in a particular mood can influence cognitive processes and facilitate appropriate behaviour and decision-making, which may in turn influence short-term emotional responses (Russell 2003). This suggests a bidirectional causal relationship between short-term emotions and longer-term moods (Mendl et al. 2010b). Regarding the duration of mood, authors are generally cautious in providing clear durations of mood, choosing terms such as ‘long-lasting’ (Berkowitz 2000) or ‘prolonged’ (Russell and Barrett 1999).

The term 'affect' is frequently used synonymously with emotion or mood in animal literature (Paul et al. 2005), yet sometimes these three terms are given distinct meanings. In human research, affective states often refer to mood states (Russell 2003). Some scientists consider affect to be the basic ability to approach positive stimuli and avoid negative stimuli, which is a behavioural skill many simple organisms such as fruit flies, bacteria and even plants, are capable of (Bliss-Moreau 2017). Others define affect as the overarching, umbrella term to encompass both emotions and moods (Paul et al. 2005, Quigley et al. 2014, Bethell 2015). Affect is also sometimes used to refer solely to the subjective experience of emotion (Panksepp 2005) or to a persistent trait characteristic, i.e. a personality trait (Hinde 1985). In humans, affect and personality are intertwined, leading individuals towards a tendency to experience certain emotions and moods more often and intensely (Watson and Tellegen 1985). For instance, the human personality dimensions of extraversion and neuroticism have been associated with a tendency towards more frequent positive and negative affect, respectively (Winter and Kuiper 1997). Note that the term ‘emotionality’ (sometimes referred to as ‘fearfulness’) is generally used in human and animal research to indicate a personality trait: propensity to experience (strong) positive or negative emotions (Archer 1973, Lecorps et al. 2018). The possible occurrence of trait affect or the link between personality and emotion in animals could explain variation in individual affective responses to the same stimuli (de Vere and Kuczaj 2016). Recently, personality was shown to influence affect in pigs (Asher et al. 2016), calves (Lecorps et al. 2018), dogs (Barnard et al. 2018) and ants (d’Ettorre et al. 2017). In calves and dogs, higher levels of emotionality/fearfulness were linked to negative mood, while in ants and pigs, negative mood was linked to less active personality types in barren environments (here negative mood refers to lower optimism on a judgement bias test; see section 3.2). Following the lead of animal literature on emotional processes, we will hereafter use the term affect as an umbrella term for emotion and mood – unless specified otherwise (**Figure 1**). The issue of whether affect requires awareness is addressed in section 2.4.

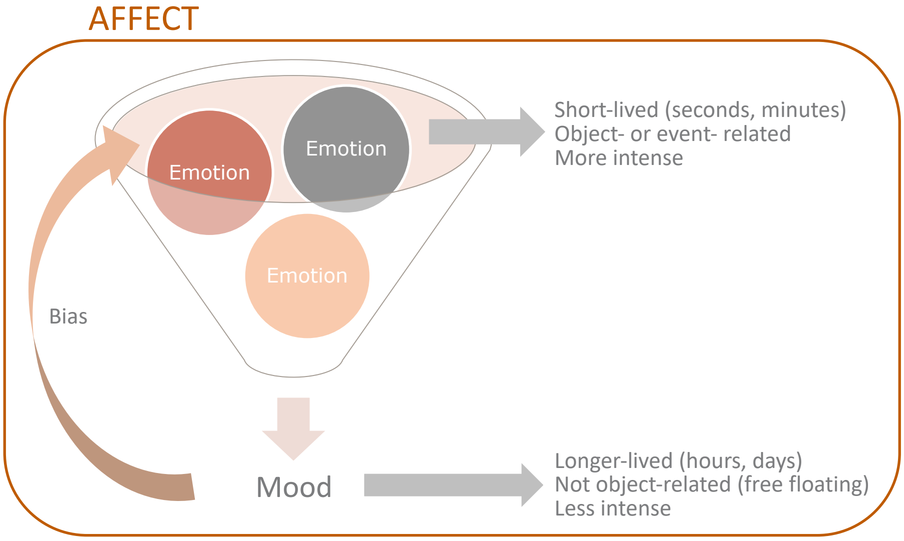


Figure 1. Illustration of the relationship between emotion, mood and affect.

Functions of affect

Affect has been described as a super-ordinate mechanism promoting approach towards valuable resources and reward, and avoidance of harm and punishment, thereby facilitating survival and reproductive success (Bethell 2015). Affect directly enhances fitness by motivating and ‘directing’ (or ‘being linked to’ depending on the theory, see section 2.4) relevant behaviour and recruiting appropriate physiological resources (Dawkins 1990, Duncan 1996), and is hence commonly perceived as a functional adaptation (Darwin 1872, Öhman and Mineka 2001, LeDoux 2012, Nettle and Bateson 2012, Trimmer et al. 2013). Affect assists with generalisation from one situation to the next, by providing a form of predictive judgement or expectation in situations that have not been encountered before, or by biasing attention to certain types of stimuli. Negative affect may place a negative veil over future subjective experiences (Grippo and Johnson 2009), for example leading animals to perceive challenges as more challenging than they are, whereas positive affect may cause an animal to perceive its environment more favourably (Mendl et al. 2010b), creating a buffer against negative events (Van der Harst and Spruijt 2007, Reefmann et al. 2012).

Affective experiences may also have evolved to facilitate group living in social species by improving the bonds, cooperation and communication between the different group members (Spoor and Kelly 2004, De Waal 2008, Špinka 2012). Affect is indeed not confined to an individual, but instead can spread from one individual to another (Hatfield et al. 1994, Špinka 2012, Briefer 2018). In this way, animals may receive signals from conspecifics which are in a negative affective state due to, for instance, the presence of a predator, and via the process of *emotional contagion* (Hatfield et al. 1994), a simple form of empathy (De Waal 2008), become negatively inclined themselves (chickens: Edgar et al. 2011, rats: Reimert et al. 2015, rats: Saito et al. 2016, ravens: Adriaense et al. 2019). Animals may also become positively excited by signals from others that are in a positive affective state (Held and Špinka 2011, Reimert et al. 2013, Saito et al. 2016). The presence of conspecifics (and even the presence of members of other species of social animals such as humans) may moreover 'buffer' the negative affective state of an animal and thereby relieve that animal from its negative state, i.e. a process (in psychology) known as *social facilitation* or more specifically termed *social support* (Rault 2012, Reimert et al. 2014, Edgar et al. 2015). Social facilitation represents the effect the presence of an individual has on the behaviour of another individual (Zajonc 1965, Nicol 1995) and may, for example, reduce fear responses in an individual in the presence of a peer displaying no or little fear (Nicol 1995). Buffering of negative affective states via the presence of another individual is a process found in a wide range of species including mammals, birds, fish and even invertebrates (Ditzen and Heinrichs 2014, Oliveira and Faustino 2017, Kiyokawa and Hennessy 2018).

Conceptual frameworks to study affect

Affective experiences in both animals and humans have been studied following different approaches. First, affective processes can be studied using two main conceptual frameworks: basic/modular/discrete or continuous/dimensional. In both animals and humans, emotions (and moods, depending on the definition) have, on the one hand, been described as modular or discrete (Darwin 1872), and this suggests that different types of emotions are processed by different areas of the brain. The discrete emotions approach is challenged because scientists disagree on the number and labels of emotions and because a single label can refer to a number of different states (LeDoux 2012, Weidman et al. 2017). A recent survey conducted among human emotion scientists showed that there was consensus for the existence (i.e. empirically tested) of five discrete emotions at best; these were anger, fear, disgust, sadness and joy (Ekman 2016). On the other hand, emotions and moods can be conceptualised as dimensional, and hence classified and quantified along two or more continuous, dimensional scales such as valence (pleasant/unpleasant or positive/negative) and arousal (activation) (Russell 1980, Wundt, 1896, in Ekman 2016). For example, whereas people adopting the discrete approach would speak of 'fear', those using the dimensional approach would speak of a negative, high arousal emotion. This two-dimensional framework is also referred to as 'core affect space' (Trimmer et al. 2013), with 'core affect' referring to any experience that varies across the axes of valence and arousal (Russell 2003)[1]. Dimensions other than valence and arousal have also been proposed to categorise core affect, for example 'persistence in time' (duration) (Anderson and Adolphs 2014). Others categorise affect in an altogether different space, with for example separate axes for positive and negative affect (Watson and Tellegen 1985). In human literature, it is still unclear whether positive and negative affect are the two extremes of a single dimension, or whether they can co-occur (Gill et al. 2017). Gill et al. (2017) found negative correlations between positive and negative affect within and between subjects, and therefore support the idea of affective valence bipolarity, that is, a single axis ranging from negative affect to positive affect. Conversely, others consider positive and negative affect as not necessarily mutually exclusive – a claim supported by the existence of mixed-feelings, i.e. two emotions from opposed valences such as joy and sadness occurring simultaneously (Larsen et al. 2001). In their conclusion, Larsen et al. (2001) nonetheless acknowledge that Russel's valence-arousal model of affect in which

---

[1] Russel (2003) defines core affect as 'a neurophysiological state that is consciously accessible as a simple, nonreflective feeling that is an integral blend of hedonic (pleasure-displeasure) and arousal (sleepy-activated) values'.

positive and negative affect are part of one single dimension (Russell 1980) holds for main emotional experiences, but should be refined with regard to bittersweet experiences. Mendl et al. (2010b) suggested a new framework that combines both discrete and dimensional views, where 'core affect' is continuously experienced (mood) and combined with evaluation of the environment to generate discrete emotions. Mendl's framework is presented in **Figure 2**.

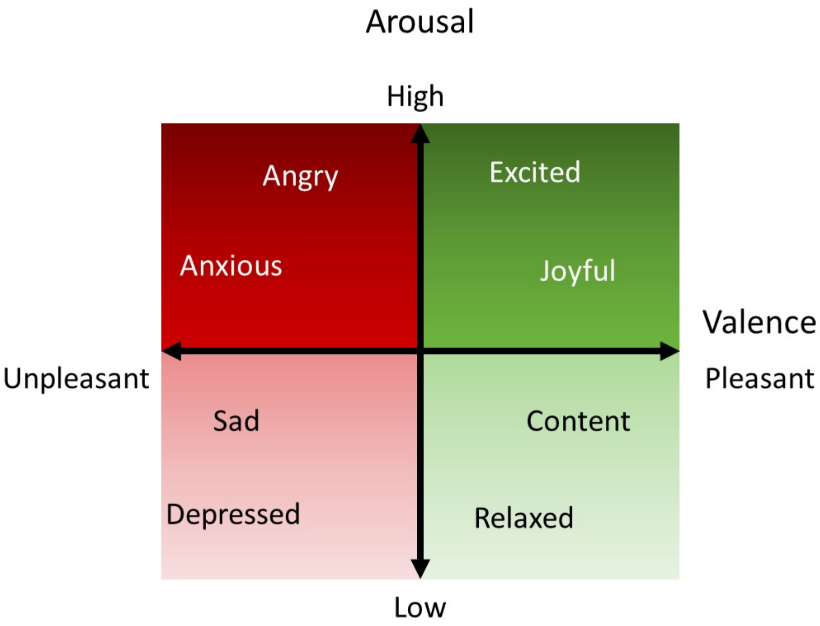


Figure 2. Core affect represented in two-dimensional space. The words in the four quadrants refer to discrete emotions. Recreated based on Mendl et al. (2010b).

Second, theories of emotion (note not mood) generally include different levels of complexity. In human literature, several levels of emotion have been proposed. Some speak of the existence of a number of 'primary' or basic emotions (Izard et al. 1993), the interaction of which can result in more complex, or 'secondary' emotions (Gray 1990, Ekman 1992, Izard et al. 1993).

Damasio et al. (2010)[2] discriminate between three kinds of emotion, given here in the order from least to most complex: *background emotion*, *universal emotion* and *social emotion*: *Background emotion* results from signals from a combination of regulatory systems, such as metabolic and homeostatic processes, and represents one's ongoing 'state of being'; *Universal emotion* refers to commonly expressed discrete emotions such as 'fear', 'anger', 'joy'; and finally *social emotion* emerges from a combination of universal affect influenced by a social context and encompasses more

---

[2] Damasio et al. (2010) define emotions as evolved 'automated programs of actions', that are not necessarily felt.

complex emotion such as ‘guilt’, ‘shame’, and ‘pride’. Another example of emotion classification is proposed by Panksepp (2010)[3], who discriminates between *primary-process emotion*, *secondary-process emotion*, and *tertiary-process emotion* – based on a neurobiological approach. The primary-process (or basic primordial) emotion, which encompasses unconditioned and ‘instinctual’ emotion action systems, homeostatic emotion and sensory-related emotion with a low level of emotional control, involves sub-neocortical activation. Panksepp distinguishes between at least seven primary-process emotional systems: ‘seeking’, ‘fear’, ‘rage’, ‘lust’, ‘care’, ‘panic’, and ‘play’ (Panksepp 2005). S*econdary-process* emotion results from memory and learning processes in the basal ganglia. ‘Wanting’ – as defined by (Berridge 1999) – for instance has been proposed as secondary-process emotion derived from the primary-process emotion ‘seeking’ (Panksepp et al. 2017). Finally, *tertiary-process emotional systems*, such as rumination and contemplation, requires higher cognitive abilities, allows an individual to act with intention, and emerges from the activation of neo-cortical structures (Panksepp 2010). Of note, Panksepp (2010) supports the theory that the ‘feeling’, or subjective component of emotion arises from ancient subcortical structures of the brain. Panksepp hence defends the idea that at least all mammals are able to experience primary emotions. This is in contrast with other authors who differentiate/separate the behavioural and neurophysiological changes from the feelings (Kringelbach and Berridge 2017). Berridge and Kringelbach (2008) for example discriminate ‘wanting’, an incentive salience emerging mostly from subcortical activation that can occur without conscious awareness, from wanting – a conscious desire emerging from cortical structures.

Regardless of the classification adopted, authors tend to distinguish between a set of ‘raw’, possibly innate, emotional states, from which more complex, learned emotion is derived. In animals, differentiating between different levels of emotion is less common than in humans, although it is clear that different species will likely be capable of different levels of complexity in their emotional experiences.

Third, there are several distinct theories regarding the causation and processing of emotion (Moors 2009, Scherer 2009a,b)[4]. Note that we here focus on emotion, which is a response to specific internal or external stimuli, as opposed to mood which is

---

[3] Panksepp (2005) defines emotions as the ‘umbrella concept that includes affective, cognitive, behavioral expressive, and a host of physiological changes’.

[4] The term emotion is used here in reference to an ‘emotional episode’ (Moors et al. 2013). An emotion episode is pluri-component, and involves changes at the appraisal, motivational, somatic, motor and subjective levels. According to Moors et al. (2013), appraisal can be seen as *one* determinant or *the* core determinant of the emotional feeling.

possibly free-floating. Appraisal theories, for example, suggest that a cognitive process (not necessarily conscious) that evaluates the importance of events for fitness, referred to as *appraisal*, precedes and causes an emotion (Scherer 1999, Moors et al. 2013). Appraisal might additionally be an ingredient for emotion rather than only a cause per se (Ellsworth 2013). Thus, appraisal theories suggest that the emotion elicited by a specific stimulus or event does not depend on the situation itself, but is based on the appraisal of the situation and hence may differ between individuals, or within individuals over time (Ellsworth 2013). Appraisal theories suggest several appraisal variables, also called criteria, that are important in differentiating between emotions (valence, goal relevance, coping potential, agency and novelty). The combination of appraisal variables can lead to a wide range of potential emotions (Ellsworth 2013). Appraisal and emotion, moreover, constantly evolve in time in a dynamic process (Scherer 2009b, Ellsworth 2013).

Another popular theory regarding causation and processing of emotion is Barrett's theory of constructed emotions (Barrett 2017), following the conceptual act theory (Barrett 2006) and building on Russell's (2003) theory of core affect. In Barrett's theory, emotions are seen as constructions of the world based on experience, not as reactions to the world, and different emotions are suggested to be labelled as a result of knowledge and experience. In other words, emotions are not direct, more or less automated responses to stimuli, but rather the outcomes of learning through experience. Thus, here the categorisation of emotion is seen as something that helps to shape experience (Moors 2009).

The causal link between emotion and behaviour is discussed in section 2.4 below.

Components of affect

In most theories of affect, different components of an affective episode are distinguished, but the specific components and their labels may differ from one author to the next. Most authors, however, refer to the following components of affect: feeling/subjective, motor/behavioural, cognitive, and somatic/physiological (e.g. Scherer 2001, Désiré et al. 2002). We focus here on emotion, as opposed to affect, as more is known about acute emotion in this context than longer term mood, except for the cognitive component which is often linked to background affective states, hence mood-like states, and except for the composite-indicators in the physiological part, which most likely also reflect mood rather than short-term emotion. This means that the term emotion is deliberately used here throughout, except in the section about the cognitive component of affect.

Feeling or subjective component

Several terminologies are used to refer to the feeling component, including ‘the subjective component’, ‘the conscious component’, ‘the experiential component’ as well as simply the term ‘feeling’. Some researchers argue that emotion inherently involves subjective experiences (Clore et al. 1994, Dawkins 2006), consciousness having emerged from primordial emotion itself (Denton et al. 2009). Others question the feeling component of emotion in both animals and humans (Winkielman et al. 2007). However, even if empirical evidence of unconscious elicitation of emotions exists (Öhman and Soares 1994), the presence of non-conscious emotion itself is still debatable (Winkielman et al. 2007). One study nonetheless showed that subliminal positive and negative emotional visual stimuli could elicit distinct behavioural responses among participants without them reporting any difference in terms of emotional valence and arousal before and after the subliminal exposure (Winkielman et al. 2005). These authors concluded that the emotional stimuli were able to alter participants’ behaviour and emotional state, without them being aware of it (Winkielman et al. 2005).

As explained by LeDoux and Hofmann (2018), who focus on fear, some simply consider the subjective component of emotion as a psychological construct, and hence not an inherent part of emotional experience. Adherents to this theory consider individual subjective reports of emotion as an invalid indicator of emotion based on the lack of correlation between individual brain activation and subjective reports (LeDoux and Hofmann 2018).

Others propose that the ‘richness’ of a species’ subjective experience depends on its level of consciousness (Damasio et al. 2010), and that different levels of animal consciousness are associated with different levels of complexity in the emotional repertoire (Le Neindre et al. 2017), although no consensus has yet been reached on the different existing levels of consciousness (de Vere and Kuczaj 2016).

Addressing all positions on the subject, some authors carefully conclude that animal ‘emotional processes [...] may or may not have subjective components, depending on the species and circumstances involved’ (Paul et al. 2005). Where in phylogeny the subjective experience of emotion emerges is not known. Although subjective life is often thought to be restricted to species with high levels of brain organization, neuro-ethological data show that animals can also have forms of subjectivity, emerging from activity in evolutionarily 'old' brain areas (Panksepp 2004, Fabbro et al. 2015, Panksepp 2016). Subjective experience of emotion in invertebrates, but also many vertebrates, is subject to intense debate (Duncan 2006). In any case, whether it exists or not in particular species, the experiential component of emotion remains difficult to assess in all animals as they cannot verbalise it. Most animal researchers nonetheless assume that it can be inferred from other components of emotion, i.e. the behavioural, cognitive and physiological ones (Mend et al. 2010), which are described later on in this review.

Behavioural component

Emotion involves behavioural changes, but there is no consensus on the causal direction for this link: while most state that emotion causes behaviour, others hold that behaviour is part of emotion, and others yet believe that behaviour in fact causes emotion (see Anderson and Adolphs 2014 for a detailed review). In addition, behavioural responses themselves may feed back to the brain and lead to an adjustment in the current emotional state, a principle termed ‘interoception’ (Anderson and Adolphs 2014). Some researchers, therefore, prefer to use the term ‘emotional episode’ (Moors 2009, Scherer 2009b) or ‘emotional process’ (Ellsworth 2013) to refer to anything from the stimulus to the consequences of an emotion. An emotional episode or process is thus broader than the emotion per se. Despite this lack in consensus regarding the direction or nature of the relationship between emotion and behaviour, behavioural changes, such as facial expressions, can be used as an important indicator of emotion in non-verbal species (see section 3.1).

Cognitive component

Affect and cognition are closely interwoven (Hinde 1985). It is suggested that affect evolved earlier than several intellectual abilities, arising from ancient subcortical brain structures, in turn suggesting that affect is more widespread across species than intellect is (Dawkins 2000, Panksepp et al. 2017). Supporters of this theory nonetheless acknowledge the fact that higher cortical structures are involved in affect regulation and inhibition processes (e.g. Damasio and Carvalho 2013). Some scientists view affective processes as dependent upon cognitive processes and vice versa (Lazarus 1999), while others see affective and cognitive processes as independent systems. For instance, Panksepp (2003) views the affective system as subcortical and the cognitive system as cortical and more hence recent in evolutionary terms. In humans, cognitive processes can trigger, or be affected by, particular emotions and moods (Hinde 1985, Mathews and MacLeod 2002), and this is most likely true in animals too. This bidirectional link between affect and cognition can hence be used to indirectly assess animal affect (see section 3.2). In humans, emotions and moods are known to cause *cognitive biases*: manipulation of information processing by the brain, affecting judgement, attention and memory (Mathews and MacLeod 1994, Mineka et al. 1998, Lerner and Keltner 2000, Schwarz 2000, Mathews and MacLeod 2002). Recently cognitive bias testing has been applied to a variety of animal species, including invertebrates (see section 3.2.1). Emotions and moods seem to also influence sensitivity to reward loss (Burman et al. 2008) (see section 3.2.2 below for further details), which can more or less be seen as a type of cognitive bias in terms of evaluation of current/past negative events (as opposed to ambiguous future events in judgement bias).

Physiological component

Traditionally, psychophysiology focused on univariate physiological parameters as measures of emotional arousal (Cacioppo et al. 2007). Subsequent work on multivariate analysis of autonomic measures later revealed emotion-specific physiological responses (Kragel and LaBar 2013). Human research also established promising links between emotion and physiological changes at the neuroendocrine and immune levels (Steptoe et al. 2005). An approach integrating the different physiological systems (neuroendocrine, immune and autonomic) has also been promoted to further clarify psychophysiological relationships (Cacioppo et al. 2007). These findings open up an avenue of research for animal welfare scientists (see section 3.3), especially with mammalian species – which share sub-neocortical limbic systems structures with humans (Panksepp 2005).

A brief note on animal satisfaction with life

Satisfaction with life in humans, also referred to as happiness, is by many defined as the ‘subjective enjoyment of one’s life as a whole’ (Veenhoven 2000). When evaluating how happy they are, humans draw from two sources of information: how well they feel most of the time (affective/hedonic happiness) and how their life-as-it-is compares with standards of how they believe their life should be (cognitive happiness) (Diener 2000, Kringelbach and Berridge 2009). Overall, little attention has specifically been given to the topic of animal happiness but Boissy et al. (2007) proposed that frequent positive affective experiences could lead to frequent positive moods, which could then lead to an overall state of satisfaction with life or 'happiness': a persistent, positive background state.

Affective happiness as described in humans is likely to define animal happiness, since there is currently no evidence that animals are capable of cognitive happiness. Affective happiness, and thus animal happiness, is a separate concept from the transient affective experiences that are emotions and moods, as it represents the balance in the frequency of all positive and negative affective experiences over time. This balance tends to be stable under stable conditions and represents how one feels *most of the time* (**Figure 3**). Similarly to humans, happiness in animals could be assessed using the frequency of positive and negative affect over a set period of time and computing the ratio of positive to negative affective experiences, that is, the affect balance (Webb et al. 2018). Affect balance is an indicator used in human happiness research since the 1960s (Glatzer and Gulyas 2014) which correlates well with self-reports of happiness but to our knowledge has not yet been applied in animals. Indicators of short-term emotions, which could be used to compute affect balance, are described below in part 2 of this review.

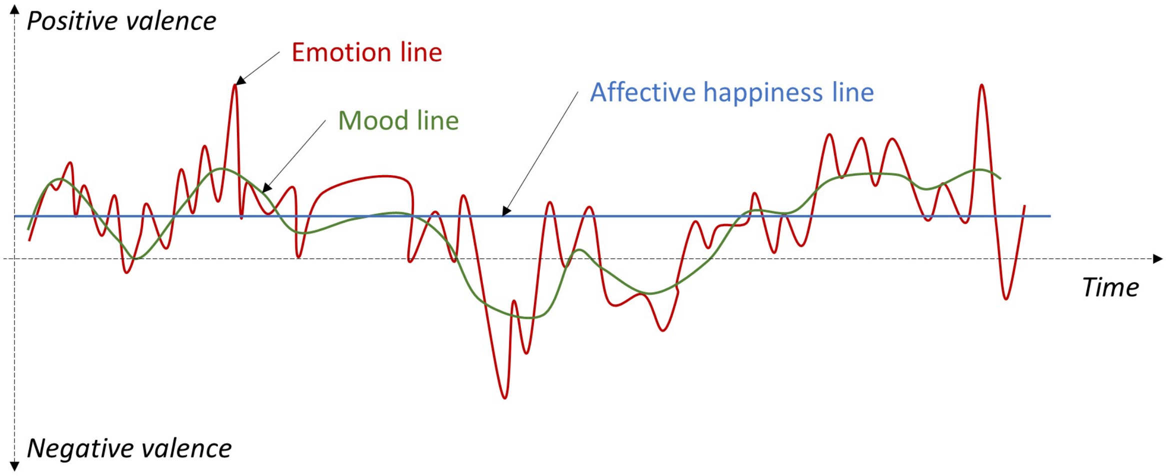


Figure 3. Rough illustration of the relationship between short-term emotions, longer-term moods and affective happiness; source: Webb et al. (2018).

Part 2 - assessing affect in non-verbal beings

How exactly subjective experiences arise and what we should be looking for when searching for evidence of subjectivity is called 'the hard problem' by Chalmers (1995). It is even called the ‘hardest problem in the whole of biology’ by Dawkins (2006).

Although many people are confident in their belief of affective capacities in at least some animal species, it remains a mere assumption. We simply cannot know what it is like to be a different individual than ourselves, and essentially, this is also true for fellow human beings (De Waal 2011). Human verbal reports of emotions and moods are accepted as the 'gold standard' indicator of affect (Paul et al. 2005), assuming humans can accurately perceive, interpret and report their own affective experiences. Since animals are not capable of verbal speech to communicate how they feel, we must rely on other methods to indirectly assess their affective experiences. These different methods, which focus on the behavioural, cognitive, and physiological components of affect, aim to identify similarities with humans and get some indication of affective experiences in animals. The currently known methods will be discussed below, grouped by the component they are based on. We focus here on valence, because: 1) affect is always valenced (Paul and Mendl 2018), 2) arousal seems less relevant to overall wellbeing (based on findings in human happiness research, Diener et al. 2009, Webb et al. 2018), and 3) discussing indicators of arousal in addition to indicators of valence would make this review impossibly long. As this review aims to cover all topics in the field of animal affect, we introduce below all known methods to assess affect in animals but we do not cover these methods in-depth. Instead we provide an extensive literature list for further in-depth reading. Finally, we do not specifically address whether these methods assess short-term emotions or moods as this is not always clear/known and would again make this review too long, but rather combine these into the umbrella term of affect.

Methods based on the behavioural component of affect

The valence of affective experiences in animals can be studied by observing their behaviour in their home environments (Fureix and Meagher 2015, Reimert et al. 2017) or in specific test situations (Forkman et al. 2007, Murphy et al. 2014). The particular behaviour studied can range from 'whole animal behaviour' to only observing specific body parts of the animal. Examples of 'whole animal behaviour' are approach and avoidance behaviour (Paul et al. 2018), freezing and play behaviour (Paul et al. 2005, Boissy et al. 2007, Held and Špinka 2011). Freezing and play behaviours are thought to be markers of negative and positive affective experience, respectively (Paul et al. 2005, Held and Špinka 2011), but see Ahloy-Dallaire et al. (2018) for a critical review of play as an indicator of positive affect. Other behaviours that have been associated with affect and more or less involve the whole body are: 1. anticipatory behaviours – i.e. behaviours displayed in anticipation of a reward or punishment, e.g. increased locomotion and frequent behavioural transitions (Spruijt et al. 2001); 2. consumptive behaviours – i.e. eating or drinking behaviours (Paul et al. 2005); 3. affiliative behaviour – i.e. social behaviour such as allogrooming (Boissy et al. 2007, Proctor and Carder 2015); 4. aggressive behaviours such as attacking or threatening another individual (Kudryavtseva 2000, Wingfield et al. 2006) and defensive behaviour such as seen in horses starting to kick upon being approached (Sankey et al. 2010) and 5. displacement behaviours indicative of (heightened) anxiety such as destructive behaviour and hyperactivity in dogs (Ohl et al. 2008) or scratching in primates (Baker and Aureli 1997, Coleman and Pierre 2014). The behaviour of the 'whole animal' is also observed in the Qualitative Behaviour Assessment (QBA) method (Wemelsfelder et al. 2001, Temple et al. 2011). In this method, the way an animal behaves is evaluated via scores on, amongst others, affective words such as 'nervous' and 'calm'.

Specific body parts of the animal that have received attention in relation to affect are head and tail postures and movements (Briefer et al. 2015, Reimert et al. 2015). Facial expressions – recently reviewed by Descovich et al. (2017) – have been studied as indicators of affective experiences for a while in primates (Andrew 1963, Van Hooff 1967) and are now also under investigation in other animals. An elegant and objective system has been developed to aid the analysis of facial expressions, called the Facial Action Coding System (FACS) (e.g. primates: Vick et al. 2007, Parr et al. 2010, dogs: Waller and Micheletta 2013, horses: Wathan et al. 2015). Based on the FACS, another system has been developed to specifically study pain in the faces of animals, called the Grimace Scale (e.g. mice: Langford et al. 2010, horses: Dalla Costa et al. 2014, pigs: Di Giminiani et al. 2016, sheep: McLennan et al. 2016, ferrets: Reijgwart et al. 2017). In these aforementioned studies the entire face is taken into account. There are also studies that only looked at one aspect of the face in relation to affect, such as the percentage of visible eye white (Sandem et al. 2006, Lambert and Carder 2017) and type of ear posture and frequency of change between different ear postures (Reefmann et al. 2009a, Boissy et al. 2011, Reimert et al. 2013, Goumon and Špinka 2016). Frequent ear posture changes in sheep, for example, seem associated with negative affect, whereas passive ear postures (ears hanging down loosely) seem associated with low arousal positive affect (Reefmann et al. 2009a).

Furthermore, specific vocalizations have also been related to affect (Briefer 2012, Leliveld et al. 2017). Rats, for example, emit a specific type of 22 kHz ultrasonic vocalizations in putatively negative affective contexts (Portfors 2007, Burgdorf et al. 2008) and a group of high frequency vocalisations referred to as '50 kHz' ultrasonic vocalisations in putatively positive contexts (Panksepp and Burgdorf 2000, Brudzynski and Pniak 2002, Burgdorf et al. 2011, Rygula et al. 2012).

Lateralized behaviours associated with brain function asymmetries have been shown in numerous species (Vallortigara 2000): the right hemisphere appears dominant in processing negative affect while the left hemisphere appears dominant in processing positive affect (Leliveld et al. 2013). As a consequence, interest in the affective meaning of specific lateralized behaviours has recently increased (Leliveld et al. 2013). In summary, it seems that animals exhibit a behavioural lateralization according to their perception of the pleasantness/averseness of an external stimulus (Siniscalchi et al. 2018). For example, dogs have been shown to display a right-biased head orientating response to recordings of joyful humans, probably as a result of left-brain-hemisphere activation (Siniscalchi et al. 2018). In addition, mares were found to mainly use their left eye to explore a negatively valenced stimulus, but their right eye to investigate a neutral stimulus, while no difference in eye lateralization was found for a positively regarded stimulus (Des Roches et al. 2008). However, the influence of specific types of affect (e.g. fear versus anger) on lateralised behaviour remains to be elucidated. To our knowledge, only one study in dairy cattle looked into this matter, without conclusive results (Kappel et al. 2017).

Behaviour, hence, is an important component of affective experiences in animals. Caution is however required as interpretation of behaviour is not always straightforward and the relationship between affect and behaviour is complex (Hinde 1985, Faragó et al. 2017). For instance, a specific behaviour can be observed in various of affective contexts and vice versa (Hebb 1946, Paul et al. 2005). Species-specific behaviour as well as subject-related and environment-related contexts need to be taken into account (Hebb 1946, Hinde 1985, Paul et al. 2005). It is moreover likely that high arousal behavioural indicators are easier to detect than low arousal ones. And finally, it is not always easy to distinguish affective behaviour from non-affective behaviour (Hinde 1985, Maestripieri et al. 1992). Therefore, we would like to end this section by emphasising that making inferences about affective experiences based on the behavioural component should always be done with care.

Methods based on the cognitive component of affect

Cognitive biases

Affect-congruent cognitive biases are inclinations to process information in particular ways due to affective states. These cognitive biases include judgement, attention and memory biases, and are described in more detail below.

Judgement bias. Affect-congruent judgement bias is the propensity to judge ambiguous cues or situations more or less optimistically. This type of cognitive bias is usually tested in animals by training them to respond in a certain way to a positively- and a negatively-associated cue (Harding et al. 2004). For example, animals are trained to go to a positive cue to receive a reward, and not go to a negative cue to avoid a punisher – this type of test is referred to as a *go/no-go* paradigm. Once trained, animals are exposed to intermediate, ambiguous, novel cues. The hypothesis is that, as in humans, negative affective states incline animals to respond to ambiguous cues as if they predict a negative event and vice versa (Harding et al. 2004). This test was first designed in animals for rats (Harding et al. 2004), and has since then been adjusted and applied to numerous other species – from chimpanzees (Bateson and Nettle 2015) to bumblebees (Perry et al. 2016). Generally, the test varies to suit the needs of the particular species being studied. Recently, however, a promising (spatial) test design that can suit two very different mammalian groups – horses and rodents – was proposed (Hintze et al. 2018).

Judgement bias tests have been criticised on several accounts: many studies find opposite results to those initially expected (Doyle et al. 2010a, Burman et al. 2011, Baciadonna and McElligott 2015); the activity of an individual animal may influence results (Mendl et al. 2009, Mendl et al. 2010a); ambiguity can be lost after repetitive presentations of the ambiguous cues (Doyle et al. 2010b, Roelofs et al. 2016), training is time-consuming (Roelofs et al. 2016) and judgement bias testing might provide cognitive enrichment which may impact affect in itself (Roelofs et al. 2016). Several strategies to counteract the undesirable loss in ambiguity following repeated exposure to ambiguous cues have already been suggested, i.e. partial reinforcement of ambiguous cues (Neave et al. 2013) or rewarding ambiguous cues according to expectations (Hintze et al. 2018). Another important consideration for this cognitive test is that the ‘reward’ and ‘punisher’ in go/no-go paradigms should be perceived as having equal strength on the motivation of the test animal. As mentioned by Mendl et al. (2009), if this is not the case, and the punisher is for example stronger than the reward, the animal is likely to choose not to respond (no-go) to all ambiguous cues because the cost of making a mistake in this case is simply too high. A final important consideration linked to judgement bias testing is that motivation level across successive sessions may vary, for example with animals reaching satiety where food rewards are used, and this is generally only tested *a posteriori* (e.g. Henry et al. 2017).

Attention bias. Affect-congruent attention bias is the propensity for heightened awareness of, or attention towards, novel or negative aspects of the environment in individuals experiencing a negative affective state, such as fear or depression. Attention biases can hence be assessed in animals by testing how attentive individuals are to a visual cue that is presented in a location associated with a threatening stimulus (Paul et al. 2005) or to novel auditory cues (Rochais et al. 2016). Attention bias can also be assessed by testing how attention-demanding tasks are interrupted by threatening stimuli, hypothesising that animals in negative affective states would be distracted more (Paul et al. 2005). Attention biases linked with anxiety have been studied in various species, including rhesus macaques (Bethell et al. 2012), starlings (Brilot and Bateson 2012), sheep (Lee et al. 2016, Monk et al. 2018b), and cattle (Lee et al. 2017). In horses, decreased attention towards novel auditory stimuli was linked to depressive-like states (Rochais et al. 2016) and in pigs they have been studied to test the impact of housing conditions (Luo et al. 2019).

Attention bias studies in animals are relatively recent and the validity of the developed methodologies is still under investigation. Compared to judgement bias tests, attention bias tests require shorter (Monk et al. 2018b) to no training time (Brilot and Bateson 2012). Attention bias, similarly to other cognitive biases, is influenced by personality (Cussen and Mench 2014, Luo et al. 2019) and social rank (Bethell et al. 2012), which is not surprising since personality and social rank are likely to influence affective state. Validation of attention bias as a potential marker of positive affect – by opposition to negative affect – is still required (Lee et al. 2017, Monk et al. 2018b). Furthermore, the influence of arousal on attention bias needs to be investigated (Monk et al. 2018b). Monk et al. (2018b) suggest that sensors or physiological measurements could help disentangle the impacts of valence and arousal on attention bias.

Memory bias. Affect-congruent memory bias is the propensity to remember specific objects or events that are in line with one’s current affective state (Klaassen et al. 2002, Burman and Mendl 2018). Depressed humans, for example, have been shown to recall negative experiences better than non-depressed people (e.g. Mineka and Nugent 1995).

To the authors’ knowledge, animal studies on this subject have only been conducted in rodents (mice: Takatsu-Coleman et al. 2013, rats: Burman and Mendl 2018). An example of such a test was training rats with different social statuses – that is, rats assumed to be in different affective states – to receive one pellet of food in each arm of a radial arm maze (Burman and Mendl 2018). Following this, rats were exposed to either a positive, neutral or negative event: letting them walk down one arm of the maze with either 12 food pellets, 1 food pellet or quinine-soaked pellets, respectively. It is hypothesised that rats in a more positive affective state, i.e. high social status, will remember arms associated to positive events better than arms associated to negative events, and vice versa. Memory is based on approach and avoidance behaviours. In this example with rats, social status did not seem to create a memory bias (Burman and Mendl 2018). Takatsu-Coleman et al. (2013), however, observed that mice exposed to 12 h-social isolation displayed a memory bias for the arms paired with the aversive event (higher avoidance), compared to control mice which had experienced no social isolation. Memory bias may also be influenced by affective arousal (Paul et al. 2005), and remains a time-consuming tool to investigate animal affective states because animals must first be trained (Burman and Mendl 2018).

Affective bias is slightly different from memory bias yet related as it refers to a bias in preference, which is linked with memory and learning processes (Stuart et al. 2013): in brief, one’s preference for particular resources are based on one’s affective state at the time of first encounter with the resources. Affective bias tests have been applied to rats (Stuart et al. 2013, Stuart et al. 2015, Hinchcliffe et al. 2017) and mice (Graulich et al. 2016). In practice, the rodents are exposed to two rewards of equal value. The first reward is presented with no manipulation of affect, while the second is presented following or directly preceding either a positive treatment (e.g. social play) to induce positive affect or a negative treatment (e.g. social isolation) to induce negative affect. Once trained, the preference of the rodents for the two rewards is tested. It is hypothesised that rodents will prefer the second reward when it is associated with the positive treatment, and avoid the second reward when it is associated with the negative treatment (Stuart et al. 2013, Hinchcliffe et al. 2017). Affective bias tests would thus offer the possibility to discriminate between affective states of different valence; but some suggest that their implementation should be restricted to studies of short-term affective manipulations (Graulich et al. 2016).

Reward loss sensitivity

Another indicator of affect that more or less relies on cognition, and may be viewed as a bias in evaluation, is one's sensitivity to reward and punishment (or reward loss). Typically, individuals are more sensitive to reward losses than gains, but when in a negative affective state, individuals show an increased sensitivity to such losses (Burman et al. 2008). To assess sensitivity to reward loss, a successive negative contrast technique can be used (Flaherty 1999). With this technique, reward loss can, for instance, be simulated by unexpectedly decreasing the size of a food reward in an operant or runway paradigm (Rosas et al. 2007, Burman et al. 2008). The expectation is that, following an unexpected and maintained decrease in reward, animals will work less, or run slower for this decreased reward in comparison to animals that have been trained to work or run for the same reward size from the beginning. Furthermore, animals in a negative affective state are expected to work even less or run even slower for this new smaller reward for an extended period of time (Flaherty 1999, Burman et al. 2008). The latter is thought to reflect higher sensitivity to reward loss, hence higher and/or more persistent negative affect following a loss. As with some other cognitive bias tests, sensitivity to reward loss requires training, may depend on individual differences and its interpretation may be complicated by differences in rewards sensitivity per se (Luo et al. 2018).

Methods based on the physiological component of affect

The papers referred to below are restricted to those in literature investigating a direct link between affective valence (hence not arousal) and physiological changes in animals.

Neuroendocrine biomarkers of affect

Animal researchers have shown increased interest towards neuroendocrine markers as potential indicators of positive or negative affect. Interestingly, these biomarker candidates have often been investigated in parallel with a judgement bias paradigm.

Dopamine. Central dopamine (DA) is known to assign a motivational value to rewarding behaviours, i.e. to be involved in positive (anticipatory) behaviours (Berridge and Robinson 1998). Recently, its role in reward processes has been investigated using the judgement bias test: bumblebees treated with a DA antagonist were more pessimistic than their control counterparts (Perry et al. 2016). Furthermore, in rats, knock-out of DA transporter in the nucleus accumbens has been found to increase anxiety- and depression-like behaviour (Bahi and Dreyer 2019). These results support the role of central DA in affective processes. Nonetheless, it is worth noting that in monkeys, only a subset of midbrain DA neurons were inhibited by aversive stimuli and excited by rewarding ones; while most neurons were excited by both aversive and rewarding stimuli (Matsumoto and Hikosaka 2009). Different groups of DA neurons would thus have distinct functions, e.g. motivational or action-oriented (Matsumoto and Hikosaka 2009). Animal researchers should thus focus on motivational-related dopaminergic system to investigate DA as a marker of affective valence. Peripheral measures of DA as markers of mood-disorders have also been recently investigated in humans, and DA levels in peripheral blood mononuclear cells have been shown to be lower in depressed individuals (Zheng et al. 2016).

Serotonin. Serotonin (5-HT: 5-hydroxytryptamine) depletion in humans has been linked to vulnerability to depression (Ruhé et al. 2007). Similarly, 5-HT depletion has been associated to a pessimistic bias in judgement bias tests – i.e. linked with negative affective states – in both sheep and pigs (Doyle et al. 2011, Stracke et al. 2017a), supporting the role of 5-HT in negative affective processes. However, tryptophan supplementation in pigs – which is a precursor of 5-HT – failed to induce the expected optimistic bias in a judgement bias test (Stracke et al. 2017b). More recently, the link between 5-HT and putative positive affect induced by enrichment in pigs was also investigated (Rius et al. 2018), without conclusive results. There are limitations to studying 5-HT functioning in the brain, and therefore peripheral (blood) 5-HT measurements have been explored. 5-HT uptake, storage and release in blood platelets shows some similarity with that in brain 5-HT (Stahl 1977, for review see Mück-Šeler and Pivac 2011) and were found to be altered by depression in humans (Barton et al. 2008). Ursinus et al. (2013) reported correlations between pigs' fear-related behaviours in an open field test and both brain and peripheral (platelet) 5-HT levels and platelet 5-HT uptake but concluded that these relationships might reflect personality differences rather than variations in affective state. Further research is hence needed to understand the potential of central and peripheral 5-HT as marker of negative affect.

Oxytocin. The potential of oxytocin (OT) as an indicator of animal affect has been investigated in several species, with a bias towards the study of positive affect – for a critical review, see Rault et al. (2017). Results remain inconsistent between species: neither straw provision in pigs or positive tactile stimuli in horses induced expected rises in plasma OT levels (Lansade et al. 2018, Rius et al. 2018).The role of the oxytocinergic system in affective processes remain poorly understood and different theories have been put forward. Rault et al. (2017) hypothesised that higher OT levels may reflect positive animal affect in stable social contexts, and negative affect in socially challenging situations. Kemp and Guastella (2011) proposed that, in humans at least, OT enhances approach related-behaviours while reducing withdrawal-related ones. OT might hence in social contexts facilitate positive affect linked to approach behaviour, such as trust, but also facilitate negative affect linked to approach behaviour, such as anger and jealousy (Kemp and Guastella 2011). Conversely, OT might inhibit affect like fear – a theory consistent with the anxiolytic properties of the hormone (Neumann and Landgraf 2012). The potential of OT as a biomarker of affect has been investigated in various matrices. In dogs, in particular, urinary and plasma OT have been showed to increase in response to positive stimuli (Handlin et al. 2011, Mitsui et al. 2011). Synchrony of OT levels may also be an indicator of social positive affect: strongly bonded dyads of marmosets have been shown to exhibit synchronised OT fluctuation (Finkenwirth et al. 2015). Note that the use of peripheral OT level has been questioned, as well as the validity of current assays (McCullough et al. 2013).

Opioids. In humans, the role of the opioid system in the regulation of positive and negative affect has been acknowledged (for a review see Nummenmaa and Tuominen 2017). However, in animals, research on the subject is scarce and appears to mainly focus on animal models of depression (Boissy et al. 2007, Robinson et al. 2017). Yet, one study by Kalbe and Puppe (2010) found that long-term food-rewarding cognitive enrichment, likely promoting positive affect, modified the opioid receptor mRNA expression in the brains of pigs. It has recently been suggested that opioids would facilitate approach orientated-affect (including positive and negative affect), while modulating withdrawal-oriented affect (Nummenmaa and Tuominen 2017). This suggests that opioids cannot be used to discriminate between affective states of different valence (Nummenmaa and Tuominen 2017), as their concentrations are affected by both positive and negative affect in the same direction.

Immune biomarkers of affect

In humans, several classes of immune biomarkers linked to affect have been brought to light: acute phase proteins (e.g. fibrinogen and positive affect: Steptoe et al. 2005, c-reactive protein and positive affect: Steptoe et al. 2007), cytokines (e.g. interleukin 6 and positive affect in women: Steptoe et al. 2007, seven peripheral cytokines and negative affect: Graham-Engeland et al. 2018), and immunoglobulins (e.g. salivary immunoglobin A and both positive and negative affect: Hucklebridge et al. 2000). Consequently, some researchers have encouraged animal scientists to conduct studies on the link between animal affect and the immune system (Tuchscherer et al. 1998, Boissy et al. 2007). Saliva protein composition, for instance, could potentially be used as an indicator of animal affect (Grigoriev et al. 2003). In particular, salivary alpha amylase and salivary immunoglobulin A have been suggested as promising indicators of positive affect in animals (Boissy et al. 2007). The latter hypothesis seems to be confirmed by a recent study conducted in calves that found higher salivary immunoglobulin A levels in calves exposed to a positive affective stimulus compared to those exposed to a negative one (Lv et al. 2018).

Autonomic biomarkers of affect

Animal researchers often investigate profiles of autonomic responses to affective stimuli rather than univariate indicators of affective response. For conceptual purposes, the indicators are nevertheless presented separately below.

Heart rate variability. Some indices of heart rate variability (HRV) are thought to reflect affective valence. HRV can be assessed by non-linear, frequency domain or time domain indices, for which the root mean square of successive inter heartbeat interval differences (rMSSD) reflecting the vagal cardiac influence is an example (Von Borell et al. 2007). HRV (e.g. rMMSD) is lower in depressed humans (van der Kooy et al. 2006, Patron et al. 2014, Schiweck et al. 2019), and HRV (SDNN: SD of normal-to-normal interval) has also been shown to increase in response to a pleasant tactile stimulus (Triscoli et al. 2017). In humans, six out of nine HRV indices (but not rMSSD) were found to differ between joy and sadness (Shi et al. 2017). In dogs, negative affective states following isolation have been associated with a decrease in rMSSD (Katayama et al. 2016). In horses, regular relaxing massages have been associated with higher HRV (rMMSD, among others) (Kowalik et al. 2017). Similarly, in sheep, rMSSD was higher when animals were in a putative positive affective state (being groomed) than when they were in a putative negative affective state (being isolated) (Reefmann et al. 2009b). Nonetheless, Briefer et al. (2015) found no relationship between rMMSD and affective valence when controlling for arousal in goats.

Respiratory rate. Respiratory rate has also been investigated as a potential indicator of affective valence in animals. In sheep, respiration rate increased in response to feed-related negative affective stimuli and decreased in response to feed-related positive affective stimuli (Reefmann et al. 2009a). In goats, however, respiration rates have only been linked to affective arousal, not valence (Briefer et al. 2015).

Peripheral temperature. Affect can be accompanied by a drop in peripheral temperatures, subsequently followed by a rise in core body temperatures – a phenomenon called *emotional fever* (Cabanac and Gosselin 1993, Proctor and Carder 2015). Vasoconstriction caused by stress causes blood to be diverted away from the periphery and towards centrally located, vital organs (Oka et al. 2001). Afterwards, post-stressor vasodilatation occurs to dissipate the accumulated heat, and peripheral temperatures consequently rise. Primates exposed to a threatening stimulus show a decrease in nasal temperature – which does not occur in response to a neutral stimulus (Kuraoka and Nakamura 2011). Conversely, the withers and nasal temperature of sheep increased after a putative pleasant experience, i.e. brushing (Tamioso et al. 2017). Another study suggests that a significant drop in nasal temperature in cows reflects a change in affect in terms of valence (from positive to negative or vice versa) regardless of arousal (Proctor and Carder 2016).

Other candidates for physiological indicators of affect

So far, despite promising research into the physiological facet of affect, no single indicators has yet been identified as a reliable indicator of affective valence. Research on the subject is still in its infancy and new indicators are currently being investigated and put forward, including composite indicators.

Telomere attrition. A telomere is a repetitive DNA segment at the extremities of a chromosome, which maintains the genome integrity and naturally shortens during mitosis (Stewart et al. 2012). In humans, affective experiences have been shown to influence telomere attrition: for instance, depression has been linked with shorter peripheral telomere length (Ridout et al. 2016), while a personality profile linked to higher optimism has been linked with longer telomeres (Schutte et al. 2016). Recently, telomere attrition has also been proposed as a marker of animal welfare (Bateson 2016). Consistent with this hypothesis, roe deer in poor environmental conditions have shorter peripheral telomeres than those experiencing better conditions (Wilbourn et al. 2017). In line with Webb et al. (2018), we encourage animal welfare researchers to investigate links between telomere length and animal welfare, with a particular focus on long-term affective states such as mood or happiness.

Allostatic load index. The allostatic load index is a composite indicator that reflects the overall physiological dysregulations occurring in response to cumulative, long-term, chronic stress at the neuroendocrine, immune, autonomic and metabolic levels (Juster et al. 2010). These physiological responses to stress are known to be sensitive to one's evaluation of a threat (i.e. one's subjective perception) (McEwen and Gianaros 2010). Consistent with this idea, the allostatic load index has recently been associated with human affective experiences: the allostatic load has been positively correlated with frequent negative affect (Dich et al. 2014) and inversely correlated with frequent positive affect (Schenk et al. 2018). We hence suggest that the allostatic load could also be a promising indicator of long-term affect in animals.

-Omics techniques. Exploratory research on physiological components of affect is expected to flourish in the future. In particular, the use of –omics techniques (i.e. metabolomics, proteomics, ...) may allow researchers to obtain a more comprehensive insight into the different physiological pathways related to affective states. Years ago, metabolomics were proposed as a valuable tool to study neuropsychiatric disorders in humans (Quinones and Kaddurah-Daouk 2009), but this application in animals is still limited (Goldansaz et al. 2017).

Comparative neuroscience

Other options to learn more about animal affect include neuroscientific approaches. As LeDoux (1996) suggested, the most reliable and objective way to measure an individual's current affective state is by directly looking at the ongoing processes in the brain. Techniques in affective neuroscience indeed make it possible to study animal (and human) brains. This has yielded several types of neuroscientific evidence for the existence of animal affect. Comparison of human and animal brain structures has revealed similarities in neural circuits and parts of the brain that are important for affective experience and processing. Primary affect is wired into subcortical structures that are anatomically and neurochemically homologous in all mammals, suggesting that at least these types of affect are fairly widespread among mammals (Gray 1987, LeDoux 1995, Panksepp 2011). Primates also share important cortical structures that are necessary for 'secondary affect', with humans (Damasio and Carvalho 2013).

Positron emission tomography (PET) scanning and functional magnetic resonance imaging (fMRI) procedures can be used to compare brain activity in humans and non-humans and to map affective systems (Andersen et al. 2002, Takamatsu et al.

2003). For example, by using PET imaging, the separation distress system in animal brains has been shown to be strikingly similar to human sadness systems (Damasio et al. 2000). By means of electrical and chemical stimulation of homologous subcortical regions of mammalian brains, evidence has also been found for other distinct affective systems such as lust, fear, rage, panic/grief and play (Panksepp 2010).

Limitations of physiological markers of affect

Using physiological measurements as indicators of affect entails some general methodological issues. First, the time course of responses in relation to the trigger should be accounted for, as well as circadian and individual effects on baseline levels (von Borell and Ladewig 1992, Ruis et al. 1997, Schrader and Ladewig 1999). Second, the collection of invasive samples may induce confounding reactions (Broom and Johnson 1993), though there is a growing number of assays for non-invasive biofluids, such as saliva and urine (Mormède et al. 2007). Third, some physiological changes may not necessarily accompany affective change, as is known in humans (Lane et al. 1997, Stone and Nielson 2001). Fourth, certain single biomarker measures may not be able to distinguish the subtleties of affect (Dawkins 2000) and are said to potentially mask the richness of animal affective experiences (Bekoff 2000). Fifth, limitations of measurements of brain activity as described here include the need for immobility, which involves either training animals or anesthetising them, or the need for more less intrusive wearable devices. Sixth, studies must also ensure that they take arousal into account, as arousal may have a confounding effect on physiology, possibly pointing to inexistent valence effects, in particular where negative and positive cues/events/contexts differ in arousal levels. And finally, it is important that researchers give attention to discriminating trait biomarkers linked to individual differences/personality and state biomarkers linked to affect.

General discussion, conclusions and future research

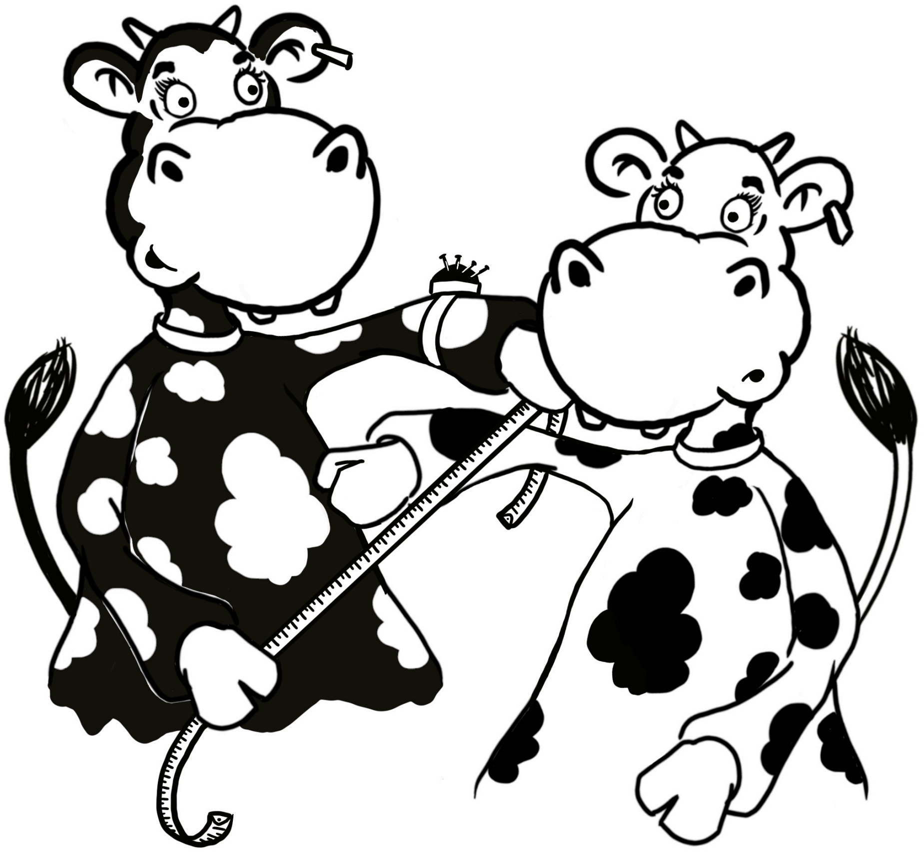
The aim of this review was to provide an overview of current knowledge in the area of animal affect, starting with an explanation of what affect is thought to be and following with a description of current methods to assess affect in non-verbal beings. Given the inconsistent use of the affective terminology in both animal and human literature, and ensuing confusion, it is essential that authors define the words they use and consistently use them throughout their articles. Recently, a definition of affect was provided by Paul and Mendl (2018), who discuss the benefits of descriptive versus prescriptive definitions, which will hopefully help bridge the gap between the different disciplines and researchers and thereby support this field of research in moving forwards more rapidly. This may also reconcile human and animal affect research, with both fields having much to learn from each other.

Another important question is how do methodologies to assess emotion and mood differ? Should they differ? And can these two affective processes be disentangled in practice? As an example, since optimism is likely impacted by both short-term emotion and baseline mood (as well as personality), it is not completely clear what the judgement bias test assesses exactly. So far it seems that this test assesses a combination of both mood and emotion, because long-term environmental conditions affect optimism in animals (Harding et al. 2004, Douglas et al. 2012), but acute events preceding the test also impact the outcome (Sanger et al. 2011). With these two affective systems being heavily dependent on one another, it may be difficult to attempt to disentangle them with certain methodologies. Moreover, affective systems depend heavily on individual differences, also referred to as trait affect, and it is crucial therefore that researchers take into account these personality variables when investigating state affect. For example, cognitive biases may be the result of individual personalities rather than transient affect (Paul et al. 2005, Mendl et al. 2009, Luo et al. 2019).

As previously mentioned by others (e.g. Boissy et al. 2007), positive animal affect has in the past received less attention than negative affect, and this gap in research is now encouragingly being addressed (e.g. Finlayson et al. 2016, Ahloy-Dallaire et al. 2018) and should hopefully continue to be in the future. Low arousal positive affect may however have been less studied due to practical limitations in assessment. Low arousal is sometimes incompatible with test situations, such as exposure to play pens (Reimert et al. 2013), which create high arousal states. Low arousal behavioural indicators may moreover be more subtle or more difficult to disentangle from low arousal neutral or negative states, e.g. inactivity (Fureix and Meagher 2015). The added complexity of separating low arousal positive emotion from low arousal positive mood, may also come into play here. However, if animal welfare is the focus of the study, the question of whether it matters what kind of positive affective state is under study, is also relevant. In human happiness for example, it is the frequency of positive affect, regardless of whether it is emotion or mood and regardless of intensity, that is of importance (Diener et al. 2009). So if the aim of a study is simply to assess animal welfare or long-term happiness, maybe the exact nature of the affective state that is recorded is less important than capturing the average frequency of positive affect over time.

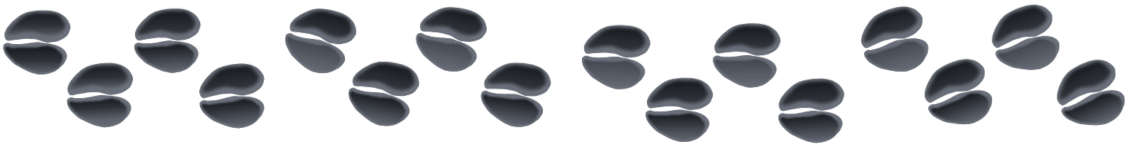
As seen above, none of the current methods to assess affect in animals is perfect; they each have their own advantages and disadvantages. Any one measure is still an indirect measure and cannot unequivocally prove subjective affective experience(s) in animals (Mendl et al. 2009). Since all methods have limited value on their own, it has been proposed – and we agree – that coupling results of different techniques will yield the best interpretation of animal affect (Broom and Johnson, 1993; Paul et al. 2005; de Vere and Kuczaj, 2016). Physiology may be key here. There have recently been several promising physiological markers of animal affect put forward. Though in the past physiology was used more often as an indication of activation/arousal, interesting links with affective valence are now apparent, especially for mood or happiness, hence long-term affect.

Finally, some researchers emphasize that combining 'hard' and 'soft' (anecdotal) research is needed to improve our understanding of animal affect (Bekoff, 2000; Morris et al. 2008). They argue we should not confine attention to quantitative data that is based on abstract measures, but also include qualitative, narrative data, for example QBA, or simply a detailed description of observations. It is often argued that this approach inherently involves unscientific anthropomorphic assumptions, but others would argue that critical anthropomorphism (Burghardt, 1991) could have valuable contributions to the study of animal affect in addition to conventional experimental approaches (Hebb 1946, De Waal 1999, Wemelsfelder 1999, Morris et al. 2008, Veissier et al. 2009, Konok et al. 2015, Spunt et al. 2017).



Chapter 3

Developing a feasible and sensitive judgement bias task in dairy cows



Animal Cognition (2021), 1-21

https://doi.org/10.1007/s10071-021-01563-8

L. Kremer[1,2], C.G. van Reenen[1,2], B. Engel[3], E.A.M. Bokkers[1], S. K. Schnabel[3], J.T.N. van der Werf[2], L.E. Webb[1]

[1] Animal Production Systems group, Wageningen University and Research, Wageningen, the Netherlands

[2] Wageningen Livestock Research, Wageningen University and Research, Wageningen, the Netherlands

[3] Biometris, Wageningen University, Wageningen University and Research, Wageningen, the Netherlands

Abstract

Judgement bias tasks (JBTs) are used to assess the influence of farm practices on livestock affective states. The tasks must be adjusted to the species and age group of focus. In cattle, most JBTs were designed for calves instead of adult cows. This study aimed to develop a JBT suitable for adult dairy cows, combining feasibility, validity, sensitivity and repeatability. Three JBTs were developed in which cows were trained to reach or avoid reaching a feeder, the location of which signalled a reward or punisher. The tasks differed in terms of punisher – cows being allocated either to 'no-reward', an air puff or an electric shock. Cows were then exposed twice to three ambiguous positions of the feeder, on two separate occasions. Speed of learning and proportions of correct responses to the conditioned locations were used to assess the feasibility of the task. Adjusted latencies to reach the ambiguous feeder positions were used to examine whether response patterns matched the linear and monotonic graded pattern expected in a valid and sensitive JBT at baseline. Latencies to reach the feeders in the two repeated testing sessions were compared to assess ambiguity loss over tasks' repetitions. The validity of using spatial JBTs for dairy cows was demonstrated. While the effect on JBT feasibility was nuanced, the punisher did influence JBT sensitivity. None of the three JBTs' repeatability could be supported. We conclude that using an air puff as punisher led to the most sensitive, yet non-repeatable, JBT for dairy cows.

Keywords

Judgement bias – Punisher – Affective state – Feasibility – Sensitivity – Cows

Introduction

During their life, dairy cows typically experience a number of potentially challenging events, such as overcrowding (Fustini et al. 2017) or regrouping (Phillips and Rind 2001), which are likely to influence their welfare. Animal welfare is defined here as a multidimensional concept that revolves around three major areas: the animal's ability to display natural behaviour, its physical condition and its affective state (Hemsworth et al. 2015; Blokhuis et al. 2019). An accurate assessment of cow welfare must, therefore, take into account affective states (Watanabe 2007). Affective states relate to individual positive and negative mental states (Duncan 2006; Fraser 2009) and encompass both emotions and moods. We define emotions here as 'states elicited by rewards and punishments' (Rolls 2000) – and moods as background states resulting from the accumulation of emotions (Mendl et al. 2010b). In animal welfare studies, the Judgement Bias Task (JBT), has commonly been used to objectively assess affective states (for meta-reviews see Lagisz et al. 2020 and Neville et al. 2020). The JBT is assumed to open a window into the affective states of animals by studying 'judgement biases', i.e. the influence of affective states on the interpretation of ambiguous stimuli (Eysenck et al. 1991; Mendl et al. 2009; Roelofs et al. 2016). Like humans (e.g. Blanchette and Richards 2010), animals in positive affective states are more likely to interpret ambiguous information more positively, hence to be more optimistic, than animals in more negative states – and vice versa (Harding et al. 2004). In livestock research, JBTs are generally used to investigate the impact of supposedly negative (e.g. shearing: Sanger et al. 2011) or positive husbandry practices (e.g. human grooming: Baciadonna et al. 2016) on the affective states of farm animals (Baciadonna and McElligott 2015). To this day, however, the application of JBT in dairy cows remains anecdotal (one study only: Crump et al. 2021), and our understanding of dairy cows' emotional life is consequently limited. By proposing and designing a JBT for dairy cows, dairy scientists may become more inclined to implement this unique tool in their research – an important step to further develop our knowledge of dairy cows' affective states, and eventually meet the dairy industry's ambition for improved animal welfare (Weary and Von Keyserlingk 2017).

In practice, judgement bias is assessed by investigating whether an individual displays a behaviour associated with the anticipation of a relatively positive or negative outcome in response to ambiguous situations. To measure judgement bias, researchers typically train animals to discriminate between two conditioned cues signalling either a reward, which represents the positive cue (P), or a less positive reward or punisher, which represents the negative cue (N). These conditioned cues typically differ according to a unique sensory continuum (auditory: e.g. Brilot et al. 2010, tactile: e.g. Brydges et al. 2011, visual: e.g. Bateson and Matheson 2007). Two main types of JBT exist – namely the active choice task and the Go/NoGo task. In active choice tasks, animals learn to display one active response to P (e.g. touch a circle symbol) and an alternate active response to N (e.g. touch a triangle symbol). In Go/NoGo tasks, animals are trained to perform one active response to P (i.e. 'Go') and to suppress this active response in response to N (i.e. 'NoGo'). Following the lead of previous studies conducted on herbivores (sheep: e.g. Doyle et al. 2010a,b, calves: e.g. Lecorps et al. 2018, goats: e.g. Baciadonna et al. 2016, horses: e.g. Briefer Freymond et al. 2014), this paper focused on a spatial discrimination task based on a Go/NoGo paradigm. Once trained, animals are generally exposed to three ambiguous cues, one at the midpoint of the sensory scale between P and N (A), one halfway between A and P (Ap) and one halfway between A and N (An) (Lagisz et al. 2020). Eventually, the judgement bias in Go/NoGo tasks is assessed based on the proportions of Go responses to, or latencies to reach, the ambiguous cues – relatively high proportions of Go-responses and short latencies reflecting more optimistic judgements, hence more positive affective states. Judgement bias, therefore, simply provides a relative measure of affective states (Bateson and Nettle 2015) and JBTs can only be used to make comparative inferences of affective states either between different populations or different treatments (Lagisz et al. 2020). JBTs remain, nonetheless, the only tool to date allowing researchers to investigate both positive and negative shifts in animal affective states – which explains its popularity and widespread use within the scientific community.

When designing a JBT, researchers must take various practical and theoretical considerations into account (Baciadonna and McElligott 2015; Bethell 2015; Roelofs et al. 2016; Hintze et al. 2018; Neville et al. 2020).

First, the *feasibility* of the task must be ensured to facilitate the adoption and the implementation of the JBT within different research groups. In practice, the feasibility of the JBTs is challenged by the duration of the training period and the number of successfully trained animals within this period (e.g. Roelofs et al. 2016; Hintze et al. 2018). Animals are typically considered trained once they reach a pre-determined training criterion, which may in some cases demand a high number of training sessions. Even with extensive training, some animals may still be excluded from the experiment for not meeting the training criterion rapidly enough (e.g. Jones et al. 2017). Consequently, the JBT results may be biased toward a population of 'learners', which may limit the generalisation of the findings (Roelofs et al. 2016). Strategies to optimise the training procedures are, therefore, warranted to enhance what we call here the tasks' feasibility – particularly in experimentations involving large animals like dairy cows, where handling is challenging (Douphrate et al. 2009).

Second, the *internal validity* of a JBT must be guaranteed to ensure a correct interpretation of the results (Mendl et al. 2009; Hintze et al. 2018). The internal validity of a tool is defined as the strength of causality between a treatment and a measured outcome (Slack and Draugalis 2001). In the JBT, this relates to the extent to which animal responses are caused by the exposure to ambiguous situations. In a valid JBT, the baseline responses (i.e. before application of any treatment, hence under reference conditions) should follow a monotonic graded pattern: latencies to reach the cues should increase as the ambiguous cues are further away from P on the sensory scale. This pattern of responses ensures that individuals respond to the ambiguous cues within the framework of the JBT and according to the learnt outcomes of the conditioned cues (Roelofs et al. 2016; Hintze et al. 2018). In other words, an erratic pattern of responses to the ambiguous cues suggests that the animals consider the middle cues as novel or meaningless rather than ambiguous, and therefore that the animals do not rely on the learnt positive and negative outcomes associated with P and N to make their decisions to approach or not the middle cues (Mendl et al. 2009; Gygax 2014; Jones et al. 2017; Hintze et al. 2018). The internal validity of a JBT should be ensured at baseline, before using the task to investigate the effects of certain treatments on animal affective states.

Third, the *sensitivity* of the task should be maximised to ensure the identification of treatment-induced shifts in animal affective states. The sensitivity of a tool is defined as the tool's ability to detect the effect it measures. In a JBT, sensitivity relates to the task's ability to detect both positive and negative treatment-induced judgement biases. In a sensitive JBT, the baseline response pattern of latencies across the ambiguous cues should ideally be linear. JBTs with baseline patterns biased toward N, for instance, are likely to be less sensitive to treatment-induced negative affect, because negative judgement bias may then only be detectable at the ambiguous cues positioned closest to P (**Figure 1**), as suggested elsewhere (Mendl et al. 2009; Lagisz et al. 2020). In practice, JBT sensitivity seems highly heterogeneous (Lagisz et al. 2020; Neville et al. 2020). Several factors inherent to the JBT set-up have recently been identified as sensitivity modulators (Lagisz et al. 2020) – including the training reinforcement combination (e.g. large reward/small reward, or reward/punisher) and the sensory continuum selected for the cues (e.g. spatial or auditory). Researchers should, therefore, carefully consider these methodological aspects when designing JBT. Failure to account for these modulators increases the risk of false negatives – i.e. the JBT fails to detect the effect of a treatment on animal affective states (e.g. Horváth et al. 2016). Such type II errors may lead researchers to erroneously claim that certain husbandry practices do not affect livestock welfare while these are, in fact, beneficial or detrimental to the animals. Designing a JBT with a set-up that maximises the task' sensitivity is also all the more valuable when considering a study population of females like dairy cows – as females appear to be less sensitive than males to judgement bias (Lagisz et al. 2020)

Fourth, the *repeatability* of the task should also be ensured to avoid erroneous interpretation of the results (Roelofs et al. 2016). Here, we define repeatability as the task's ability to ensure that repeated exposures to the ambiguous cues do not lead to ambiguity loss. If the animals associate a specific outcome with the ambiguous cues over several exposures, then the ambiguous cues become, by definition, no longer ambiguous (Roelofs et al. 2016; Hintze et al. 2018). In practice, repeated exposures to ambiguous cues have been associated with increased reluctance to approach the cues (Doyle et al. 2010b) – which could falsely be interpreted as a treatment-induced pessimistic bias within the context of a longitudinal study. Before using any newly developed JBT, researchers must hence assess the repeatability of their task at baseline to ensure its suitability for longitudinal designs. The necessity to develop repeatable JBTs arises from recent evidence demonstrating the importance of endogenous factors, such as personality traits, on animal responses to the JBT (e.g. fearfulness in calves, Lecorps et al. 2018). Unlike trans-sectional studies, longitudinal studies allow researchers to control for individual differences that would otherwise bias the outcomes of the JBT. Longitudinal studies, furthermore, allow for the introduction of extra training sessions between the first testing session (baseline before the application of the treatment) and the second testing session (after the application of the treatment). These additional sessions may serve as a 'wash-out' period – potentially reducing the likelihood of animals remembering their first encounter with the ambiguous cues (Doyle et al. 2010b). The ability of such wash-out period to potentiate ambiguous loss over repeated testing remains, nonetheless, to be proven.

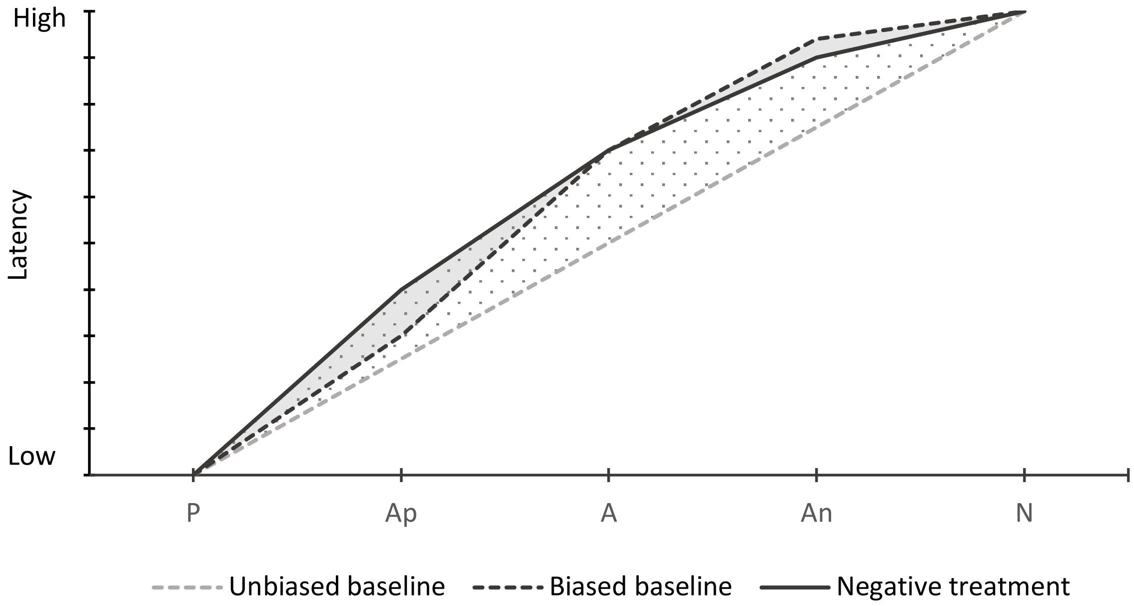


Figure 1. Example of expected latencies in response to the conditioned and ambiguous cues before (dotted lines) and after (solid line) the application of a negative treatment in the context of a judgement bias task (JBT). Treatments inducing negative shifts in animal affective states lead to more pessimistic responses to the ambiguous cues, i.e. the latencies to reach the ambiguous cues will be higher. The linear monotonic graded baseline represents the pattern of latencies obtained in response to the cues before the application of any affective treatment in the context of a valid and sensitive JBT (grey dotted line). In a less sensitive JBT (black dotted line), the profile of responses before the application of the treatment is not linear. The differences in responses obtained before and after the negative treatment are greater in the case of an unbiased baseline (dotted area) compared with a negatively biased baseline (grey area). In the negatively-biased scenario, the treatment-induced negative affective shift may not be detected – or only in response to the ambiguous positive Ap cue. P: positive, Ap: ambiguous positive, A: truly ambiguous, An: ambiguous negative; N: negative.

In light of the aforementioned considerations, our goal was to develop a feasible, valid, sensitive and repeatable JBT for dairy cows. In this paper, we focused on a specific methodological aspect of the JBT: the combination of reinforcers. The combination of reinforcers can modulate the feasibility of a discrimination learning task, by influencing an animal's ability to discriminate between two perceptually similar conditioned cues. For example, in a visual discrimination task using sucrose as a positive reinforcer, bees' visual discrimination of two shades of the same colour was enhanced by the use of a quinine solution instead of water as the negative reinforcer (Avarguès-Weber et al. 2010). The combination of reinforcers can also impact the sensitivity of JBTs (Mendl et al. 2009; Roelofs et al. 2016). Animal decision-making about whether or not to approach an ambiguous cue is thought to result from the interaction of the two generalisation gradients around the positive and the negative conditioned cues (**Figure 2**) (Roelofs et al. 2016). The generalisation gradient describes the phenomenon by which individuals transfers a learnt behavioural response from one conditioned cue to other perceptually similar cues (Guttman and Kalish 1956; Schechtman et al. 2010) and depends on the inherent properties of the reinforcer associated with the stimulus. In humans, for instance, threat-intensity has been shown to widen the generalisation gradient around N (Dunsmoor et al. 2017). Modifying the combination of reinforcers by replacing the type of punisher may, thus, influence JBTs in terms of both feasibility and sensitivity.

Consequently, we developed three JBTs differing solely in terms of negative reinforcer, hereafter called punishers – using either a 'no-reward' (i.e. the absence of feed delivery), an air puff or an electrical shock. The 'no-reward' and air puff punishers were selected based on previous studies that successfully developed JBTs using feed-reward/ 'no-reward' (e.g. Hintze et al. 2018; Crump et al. 2021) and feed-reward/air puff combinations as reinforcers in herbivores (e.g. Destrez et al. 2013; Lecorps et al. 2018). The electric shock was selected based on its proven efficacy to contain cattle on pasture via electric fencing (McDonald et al. 1981) and its common use as a punisher in rodents (e.g. Enkel et al. 2010). The overall validity of our spatial discrimination task was assessed to investigate whether cows effectively perceived the intermediary spatial cues as ambiguous. We assessed the repeatability of each JBT separately to investigate their potential for longitudinal studies. The repeatability of the tasks was evaluated while implementing a wash-out period of two weeks between two testing periods. Finally, we completed our assessment of each JBT by comparing the tasks in terms of feasibility and sensitivity. We hypothesised that, in dairy cows: 1) The reward/punisher combination influences the feasibility of the JBT; 2) Spatial JBT is overall valid in dairy cows; 3) The reward/punisher combination influences the sensitivity of the task; 4) All JBTs are repeatable when a wash-out period is implemented between two testing periods, regardless of the combination of reinforcers applied.

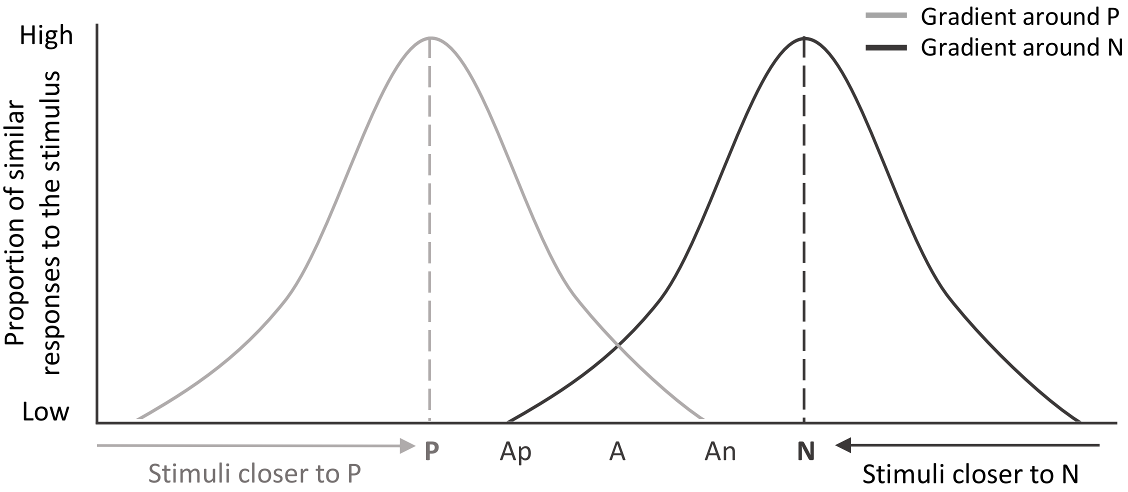


Figure 2. Generalization gradients represented as Gaussian distributions around the negative cue N in dark grey and around the positive cue P in light grey (adapted from Roelofs et al. 2016). Animals are more likely to display the same behaviour learnt in response to the conditioned cue (e.g. P or N) when faced with unconditioned cues similar to the conditioned one. Ap: ambiguous positive cue, A: truly ambiguous cue, An: ambiguous negative cue.

Materials and methods

The study was conducted between May and October 2018 at the experimental farm of Dairy Campus, Leeuwarden, the Netherlands. All procedures complied with the Dutch law for animal experiments and were approved by Wageningen University Committee on Animal Care and Use.

Animals and management conditions

Experimental animals were mid-lactating Friesian Holstein dairy cows (N=39; 3.7 ± 0.1 years old on average when enrolled in the study, 23.1 ± 0.8 kg of milk per day; 692.8 ± 11.1kg body weight two weeks before the start of the study) between their first and third lactation. The study was divided into two experimental batches of three months each (N1=21 and N2=18, respectively). Focal cows (i.e. cows used in the experiment) were housed in a solid floored free-stall barn, opened to the exterior. Focal cows were housed with 32 ± 3 (mean ± sd) companion cows (i.e. cows not used in the experiment) that were mixed and replaced according to the farm's regular schedule and need. Dim artificial lighting was provided 24 h/24 h. Cows had access to four automatic brushes and 54 flexible cubicles with gel mattresses (AgriProm) and covered with sawdust. Cows received a total mixed ration of grass silage (10.5 % of dry matter), maize silage (15.8 %), brewer's grains (4.5 %), grinded whole soy (7.7 %), grinded whole wheat (8.1 %), concentrates (14.5 %) and minerals (1.8 %) around 9:00 h – that was pushed towards the fences around 17:00 h. Additionally, cows had free access to four automatic concentrate dispensers delivering a pre-set daily amount of concentrates based on individual milk production and *ad libitum* access to four water troughs. Milking occurred twice a day between 08:00-09:00 h and 15:00-16:00 h. To facilitate handling, the barn was divided into two pens during workdays (Mon-Fri) and focal cows were separated from their companions. The three punisher treatments were balanced for focal cows' parity.

Experimental design

Focal cows were subjected to one of three judgement bias procedures that differed in terms of punishers. The punisher was either an inaccessible feed reward coupled with a 10 s time-out (NOTH), a 5 bar-air puff (AIR) (SPECAIR HI 275/35), or an electric shock (ELEC) (GARMIN Delta XC, 7/18). The feed reward, which consisted of 150 g of concentrates, remained out of cows' reach and sight by storing it inside the receptacle of a wood-crafted feeder on wheels. The air puff was delivered via an air pipe connected to the bottom of the feeder bowl – where cows would usually eat the feed when the latter was made accessible. The air puff experience also included hearing a loud noise, as a result of sudden air release. The electric shock was delivered from a neck collar. Punishers were assumed to initially induce frustration (NOTH), fear and frustration (AIR), and a combination of pain, fear and frustration (ELEC). Assumptions of punisher-induced affective states were based on the appraisal theories, which postulate that specific situations trigger specific affective states (Sander et al. 2005). Since frustration is thought to emerge from a situational inability to attain a goal, we hypothesised that all punishers elicited frustration because cows could not fulfil their desire to eat, a desire likely triggered by the smell of the concentrates emanating from the receptacle. Additionally, we hypothesised that the AIR- and ELEC-punishers elicited fear because these stimuli of low intrinsic pleasantness were sudden, unfamiliar and unpredictable (Désiré et al. 2004). Arguably, however, the release of the air puff and of the electric shock may have become more predictable across repeated exposures. Finally, we assume that the ELEC-punisher induced pain in cows.

A pilot-study was also conducted before the main experiment, during which cows were trained to reach a bucket filled with 150 g of concentrates during 7 consecutive trials per day for 2 weeks. The objectives of this pilot-study were 1) to ensure that cows were willing to eat 1.050 kg of concentrates on top of their daily ration of concentrates, 2) to ensure that cows were willing to participate in the task over an extended period and 3) to optimise cow handling inside the experimental facility. In this way, the experimenters ensured that cows would not stop responding to P across training sessions due to a lack of interest in the reward and they learnt to handle cows in a stress-free and efficient manner.

Two experimenters remained present during the study – experimenter 1 (L.K) being in charge of preparing the experimental facility and releasing the appropriate reward or punisher, and experimenter 2 being in charge of handling the cows. Experimenter 2 differed between the two batches, while experimenter 1 remained the same throughout. The judgement bias procedures consisted of several phases (**Figure 3**) described in detail below.

Judgement bias

Experimental facility

The judgement bias procedures took place in a dedicated 7×7 m and 3.5 m-high wooden-walled arena located in-between the barn and the milking parlour. The arena had a roof, a concrete floor and no window. Artificial light (4000 K cool white) was provided with six fluorescent tubes and two spotlights. Four cameras (CAMCOLBUL2, Velleman, Belgium) were installed inside the testing arena. A starting box adjacent to the arena allowed the cows to enter the testing arena. Cows were brought from their home pen to the arena in groups of three (test group) using a familiar milking route (**Figure 4**). The order of test groups was randomly-determined each day, except for replacement group 7, initially trained during the evenings.

Habituation

The habituation procedures took place in six incremental steps (**Figure 5**). The order of habituation (first, second and third cow) within a group was determined based on the punisher allocation: each of the six possible sequences of three punishers was randomly allocated to at least one test group per batch. This order was maintained during the entire experiment. During step 1, the habituation session consisted in a unique 5 min-trial during which cows were habituated to remain in groups of three inside the arena. Three buckets filled with 150 g of concentrates were initially interspersed in the arena – as an incentive for cows to explore the arena.

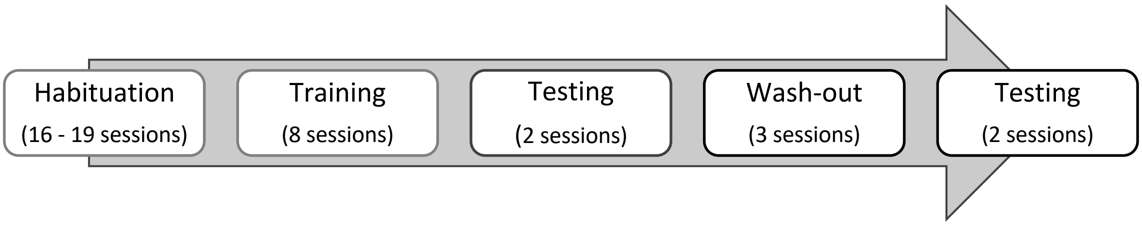


Figure 3. Timeline of the phases of the judgement bias task for one experimental batch. All sessions took place during the weekdays (Mon-Fri). All groups were habituated on the same days. One to two habituation sessions were conducted per day. Habituation sessions lasted 5 to 15 min per cow. Extra-habituation sessions were provided to cows who did not reach the habituation criterion after 16 sessions. Extra sessions were provided within 24h following the 16th habituation session. During training, three groups of 3 cows could be trained per day (i.e. two days were needed to train all cows in one batch to the same training session). In the first batch, a seventh group was also trained in the evenings after milking. Training sessions lasted 30 to 45 min per cow. Two consecutive days were required to complete one testing session with all cows. Testing sessions also lasted 30 to 45 min per cow. The same holds for wash-out sessions.

From step 2, cows were habituated to stay 90 s alone inside the arena, with two buckets filled with concentrates positioned in diagonal corners. From step 3, the buckets were replaced by two automatic feeders that were always positioned in a concentric fashion from the centre of the arena. An electric wire was connected between the feeder and a console located in the experimenter’s office so that experimenter 1 could release 150 g of concentrates each time a cow would successfully reach the feeder. If the cow did not reach the feeder within 90 s, concentrates were still released. If the cow still did not reach the feeder on her own 30 s after the release of concentrates, experimenter 1 entered the arena and gently orientated the cow toward the feeder while encouraging her vocally and petting her hips. From step 4, cows were habituated to being inside the starting box for 15 s before the start of each trial. From step 5, cows were habituated to wearing the electrical collar – that was gently put on the cow’s neck in the waiting area prior to the habituation session of the first test cow. All cows were habituated to wearing the collar – regardless of their allocated punisher. From step 6, each session consisted of 2 trials, in-between which cows were trained to turning around in the turning area to re-enter the arena. At this point, cows only had access to one feeder positioned in a pre-selected corner of the arena, balanced across groups. If the cow did not reach the feeder within 90 s, concentrates were still released. If the cow still did not reach the feeder on her own 30 s after the release of concentrates, experimenter 1 entered the arena and gently orientated the cow toward the feeder while encouraging her vocally and petting her hips. Cows were considered habituated to the judgement bias procedures once they had reached the feeder for two consecutive trials within 90s.

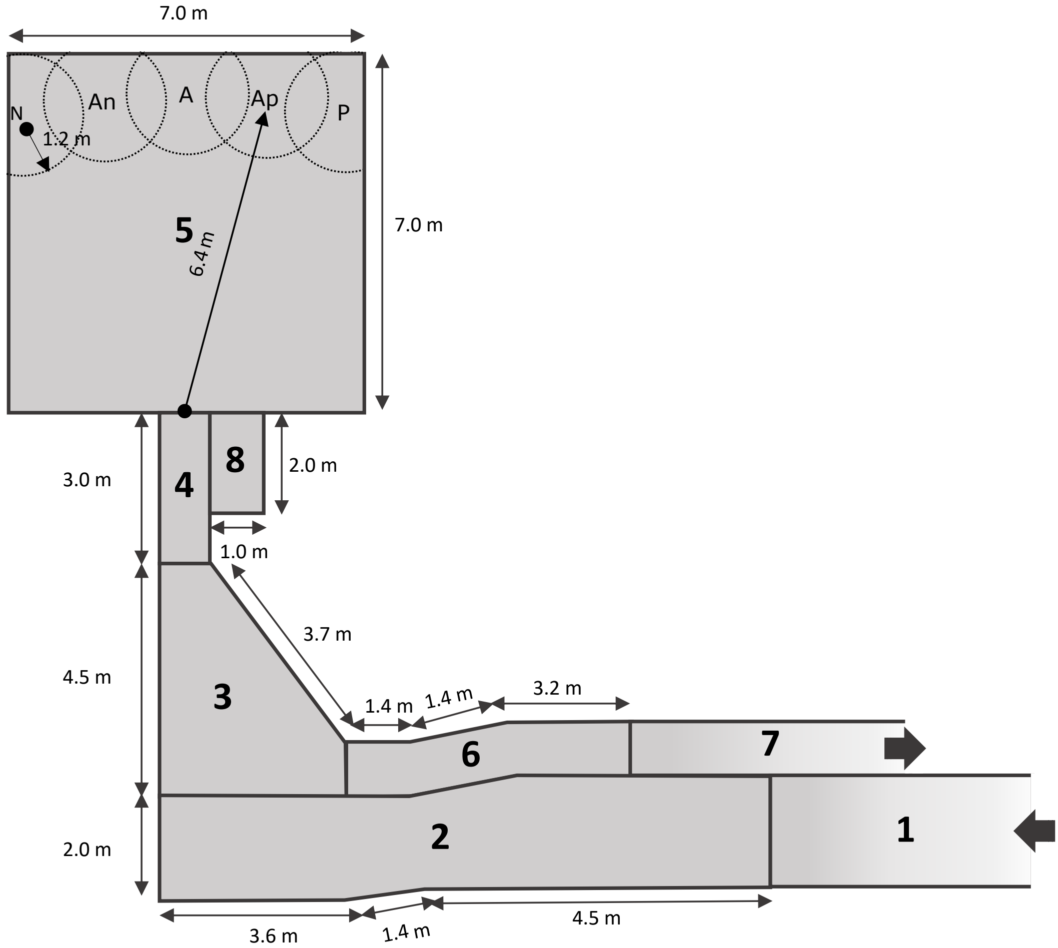


Figure 4. Layout of the experimental facilities used for the judgement bias procedures.

1: Entry corridor. It belongs to the milking corridor and is used to bring the groups of cows from the home pen to the waiting area.
2: Waiting area. Cows remained here while a cow from their group was habituated, trained or tested in the arena.
3: Turning area. Cows were used to turn on themselves inside the turning area to re-enter the arena in-between two consecutive trials. In the meanwhile, experimenter 1 positioned the feeder at the correct location for the next trial, and out of sight from the cow.
4: Starting box. Cows remained in the starting box for 15 s before each trial. A metal bar was positioned behind the cow to prevent her from going backward. The metal bar was lifted up shortly before the saloon door was opened.
5: Testing arena. The feeder was always presented at one of the 5 indicated locations (N/An/A/Ap/P) during a training or a testing trial. P and N locations were balanced across punisher.
6: Resting area. Cows were released onto the resting area at the end of a session while a cow from their group was trained or tested inside the arena.
7: Exit corridor. All cows from a group were thereafter brought back together to their home pen via the exit corridor.
8: Experimenters' office. Experimenters could observe the cows inside the testing arena via a screen connected to the cameras located inside the testing arena. The experimenters also had access to an automatic console, that they used to release concentrate and air puff by distance. The remote-control to deliver the electric shock was also placed in the experimenters' office.

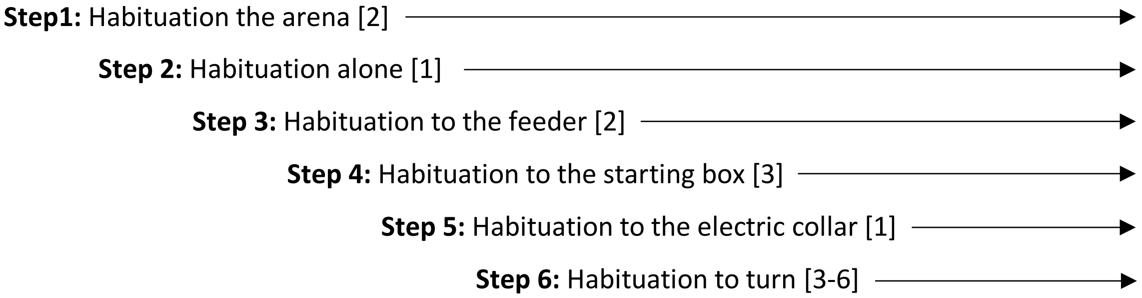


Figure 5. The six steps of habituation to the judgement bias task. The numbers in square brackets refer to the number of habituation sessions given per step of habituation.

Training

Once habituated, cows were subjected to eight training sessions. Cows were brought in groups of three to the waiting area, before being individually trained. During the training sessions, cows had to learn to discriminate between two feeder locations situated either on the far right or far left corner of the arena (N and P in **Figure 4**). One location signalled a positive outcome (P) while the other signalled a negative outcome (N). P and N cues were randomly assigned to the far-left or the far-right corner of the arena per cow, and the rewarding corner was balanced across punishers to avoid side-bias. For practical reasons, the rewarding corner remained the same within a test group of 3 cows. Training sessions consisted of 7 consecutive trials, of maximum 90 s each. Regardless of the trial type (N or P), the receptacle of the feeder was always filled with 150 g of concentrates to prevent cows from relying on olfactory cues to discriminate between the spatial cues. Before each trial, the focal cow remained in the starting box for 15 s, after which a saloon door leading onto the arena was opened. Experimenter 2 then firmly slapped three times with both hands on the cow's hips to encourage her to enter the arena. If the cow did not enter the arena, three additional slaps were given. If unsuccessful, the experimenter eventually pushed the cow inside the arena. The trial would only start once the cow crossed the virtual door line with one hoof. The order of exposure to P and N was pseudo-randomly determined, so that each training session would start with a positive followed by a negative trial and would always end with a positive trial (e.g. **P**-**N**-N-P-N-P-**P,** where letters in bold indicate trials that remained fixed across all sessions). The cows could not be exposed more than two consecutive times to the same trial type, mostly to minimise the negative experiences following the negative trials and ensuring the cow's willingness to participate in the training. Cows were trained to display Go-responses to P and to reach the feeder to get 150 g of concentrates. Experimenter 1 scored a Go-response if the cow reached P within 20 s. If the cow did not reach P within 90 s, concentrates were still released. If the cow still did not reach the feeder on her own 30 s after the release of concentrates, experimenter 1 entered the arena and gently orientated the cow toward the feeder while encouraging her vocally and petting her hips. Cows were also trained to display NoGo-responses to N to avoid their assigned punisher. Experimenter 1 scored a NoGo if the cow did not reach N within 90 s. When the cow reached N, AIR or ELEC punisher was immediately delivered; and thereafter the cow was immediately released from the arena. NOTH-cows remained ten additional seconds in the arena before being released, to prevent them from associating the N cue with the immediate end of the trial. The trial ended either after 90 s, or once the cow reached N - i.e. when the cow crossed the 1.2 m – radius circle around the feeder (**Figure 4**). Cows were considered trained once they made at least 13 correct responses out of 14 trials during two consecutive sessions. This criterion was selected based on existing literature in farm animals, where it ranges between 80 % of correct responses over 20 trials (Hintze et al. 2018) to 100 % over 10 trials (Lecorps et al. 2018). All cows had exactly 8 training sessions (i.e. 56 trials), regardless of when they met the training criteria. Latencies to reach the conditioned cues during training were video recorded.

Testing

After eight training sessions of seven consecutive trials each, the testing phase started. One testing phase consisted of two testing sessions of seven trials conducted on two separate days. The testing phase always started with a positive and a negative trial – as a reminder of the task for the cows – and ended with a negative and a positive trial, in this order. The same procedures as those used during the training were applied for P and N cues. Cows were exposed to the ambiguous cues during three consecutive trials (3rd, 4th and 5th). The feeder, always filled with 150 g of concentrates, was positioned either in the middle of the arena (A), in-between A and P (Ap), or in-between A and N (An) at 6.40 m from the starting line (**Figure 4**). The order of exposure to the ambiguous cues was based on studies conducted in sheep and calves (Destrez et al. 2012; Lecorps et al. 2018). Ambiguous cues were presented in the order Ap/An/A during the first testing session, and in the order A/Ap/An during the second testing session. The ambiguous locations were neither rewarded nor punished with an air puff or an electric shock. The ambiguous trials ended *as soon as* the cow entered the 1.2 m-zone around the feeder with one front hoof. Therefore, cows did not experience a 10 s time-out when they reached the ambiguous cues – unlike NOTH-cows when they reached N. Latencies to approach the cues were video-recorded and scored as done during the training sessions. The same procedures as the ones used during the first testing session were applied in a second testing session, and cows were tested in the same order as they were during the first testing session. The second testing phase occurred after a wash-out period of 10 days. The wash-out period consisted of three sessions of regular training, aiming at reducing the risk of cows remembering the outcomes of the ambiguous cues. Furthermore, maintaining training until the second testing period minimised the risk of altering cow affective states between the two testing sessions. The JBT training may indeed provide a form of cognitive enrichment (Roelofs et al. 2016), which could improve animal affective states (Pomerantz and Terkel 2009; Zebunke et al. 2013). Stopping animal training could, therefore, negatively influence affective states. In total, each cow was therefore tested 4 times, two times before and two times after the wash-out period.

Data analyses

All statistical analyses were conducted using R version 4.0.5 (R Core Team 2020). The significance level was set at $\alpha < 0.05$. The tendency level was set at $\alpha < 0.10$. Data and scripts are available in a public repository DOI.4121/15125193.

Feasibility

The feasibility of each JBT was assessed based on cows' responses to the conditioned cues during training. One AIR-cow was removed from the study due to aggressiveness towards the experimenters, and two cows were excluded from the analyses since their punishers (ELEC and NOTH) were mistakenly switched during the third training session. Thus, in total, thirty-six cows were included in the dataset (NOTH-cows: 12, AIR-cows: 12, ELEC-cows: 12).

Learning success was assessed based on the proportion of trained cows after 8 training sessions and cows' learning speed for each JBT. The effect of the punisher on the proportion of trained cows was investigated using a Fischer exact test. Learning speed was the number of sessions required for each cow to reach the training criterion. Learning speed was scored as a 9 when cows did not reach the training criterion after 8 training sessions. Differences in learning speed according to the punisher were investigated using a Friedman test and specifying Group as block.

Cow discrimination between P and N was assessed by calculating latencies to reach P and N during training. Response variables were expressed as remaining latencies, i.e. $remaining\ latency\ (Cue_i) = 1 - \frac{Latency(Cue_i)}{90}$, in such a way that a NoGo response corresponds to a remaining latency of zero. Remaining latencies were analysed using a generalized linear mixed model (GLMM, McCulloch and Neuhaus, 2014). Analysis was by approximate maximum likelihood estimation using Laplacian integration, employing routine glmmTMB (Brooks et al. 2017). The used GLMM comprises a logit link for fixed and random effects and a beta distribution for the proportions of remaining latencies. It allowed to model NoGo responses with a probability *p* and Go responses with a probability *1-p*. For a Go response, where a non-zero proportion of remaining latencies was observed, fixed effects on the logit scale included main effects for batch, punisher (NOTH, AIR, or ELEC), cue type (P or N) and the interaction term between punisher and cue type. Random effects of the intercepts were included for sessions nested within cows nested within groups. In the NoGo part, the logit of *p* was modelled with fixed effects for punisher and cue type, and random effect of the intercept for cows. The interaction term between punisher and cue type in the NoGo part was dropped from the final model because it had no significant effect on cows' probability to display NoGo responses. Wald tests were performed to assess the fixed effects, both for the Go and NoGo parts of the model. For subsequent pairwise comparisons, based on estimated marginal means (on the logit scale) a Bonferroni correction was applied for multiple testing. As is customary, adjusted p-values higher than 1 were rounded to 1.

Each response to P and N was also scored as either correct (1) or incorrect (0), based on the conditioned cue and the latency to reach the cue. In response to P, latencies smaller or equal to 20 s were scored as 1 (based on Henry et al. 2017), while latencies above 20 s were scored 0. In response to N, a NoGo response was scored as 1, and a Go response was scored as 0. These binary data were also analysed using a GLMM employing routine glmmTMB. This specific GLMM comprised a logit link and a binomial (Bernouilli) distribution. Fixed effects on the logit scale included main effects for batch, punisher (NOTH, AIR, or ELEC), cue type (P or N) and the interaction term between punisher and cue type. Random effects included sessions nested within cows nested within groups, following the recommendation from Gygax (2014). Wald tests were performed to assess the fixed effects. Again, subsequent pairwise comparisons included a Bonferroni correction.

Internal validity: discrimination among the cues

The validity of our JBT was assessed based on cows' responses to the cues during testing sessions of the first period. Analyses were conducted on cows who met the training criterion. One NOTH-cow was removed from the study due to miscarriage (n = 25; NOTH-cows: 6, AIR-cows: 9, ELEC-cows: 10).

Internal validity was evaluated based on cow discrimination of the cues. For each cow, latencies to reach the cues were averaged over the two testing sessions (hence not training) of the first testing period. Adjusted latencies were thereafter calculated using the following expression (Mendl et al. 2010a):

$$Adjusted\ latency\ (Cue) = \frac{Latency(Cue) - Latency(mean(P))}{Latency(mean(N)) - Latency(mean(P))}$$

Adjusted latencies were used to account for differences in walking speed between cows. First, the overall validity of our spatial Go/NoGo task (i.e. for all punishers combined) was assessed by investigating differences in adjusted latencies between two adjacent cues (e.g. P and Ap; An and N) using Wilcoxon signed rank tests. Second, the internal validity of each JBT was assessed separately by investing differences in adjusted latencies between the truly ambiguous cue A and the conditioned cues P and N using Wilcoxon signed rank tests.

Sensitivity

The sensitivity of the test was assessed by calculating the divergence of cow experimental response from the theoretical unbiased baseline. During the first testing period, the positive area $\mathrm{A}+$, and negative area $\mathrm{A}-$ were calculated for each cow between the curves obtained for the experimental adjusted latencies and the theoretical line of adjusted latencies (i.e. respectively $\mathrm{A}+$ above, and $\mathrm{A}-$ below the theoretical unbiased baseline). To assess the divergence from the expected theoretical line according to the punisher, the response variable signed area *SA* was determined as follow: $SA = \mathrm{sign}(\max(\mathrm{A}+, \mathrm{A}-)) \times \max(\mathrm{A}+, \mathrm{A}-)$. Negative *SA* thus indicates a punisher-driven positive judgement bias, while a positive *SA* indicates a punisher-driven negative judgement bias. *SA* differences according to the punisher were analysed by using Wilcoxon rank-sum tests. At cue level, differences between the adjusted latencies to reach each ambiguous cue according to the punisher were also assessed using Wilcoxon rank-sum tests.

Repeatability

The repeatability of each JBT was assessed based on cows' responses to the cues during the first and the second testing periods. Analyses were conducted on cows who met the criterion established during the initial training period (n = 25, NOTH-cows: 6, AIR-cows: 9, ELEC-cows: 10). For each cow, latencies to reach each cue were averaged over the two testing sessions of one testing period, before calculating the respective adjusted latencies. Differences between adjusted latencies during the first and the second period were analysed using Wilcoxon signed rank test for each type of ambiguous cues separately. Spearman's rank correlation coefficients between adjusted latencies to reach the ambiguous cues during the first and second testing periods were also calculated as a measure of JBT repeatability.

Results

Odds ratios (OR) and associated 95% confidence intervals (CI) are specified for binary data (including the NoGo part associated with a probability *p*); means of the raw data ± standard error and inter-quartile ranges (IQR) are given otherwise.

Feasibility

Learning success

The Fisher exact test did not reveal statistical evidence for significant differences in proportion of trained cows between punishers (NOTH: 7/12, AIR: 9/12, ELEC: 10/12, p=0.526). Similarly, there was no statistical evidence for significant differences in learning speed between punishers (NOTH: 7.2 ± 0.63, IQR=3.3; AIR: 6.6 ± 0.64, IQR=3.5; ELEC: 5.9 ± 0.68, IQR=4.3, p=0.249, χ^2=2.78, df=2).

Go responses to the conditioned cues

There was a significant interaction effect between the punisher and the cue type on cows' latencies to reach the cues (p=0.005, χ^2=10.6, df=2). Pairwise comparisons showed no statistical evidence that NOTH-cows were significantly faster to reach P than AIR-cows (t=1.38, df=1954 for all pairwise comparisons) or that AIR-cows were significantly faster to reach P than ELEC-cows (t=1.68). However, NOTH-cows were significantly faster to reach P than ELEC-cows (t=3.02). Furthermore, there was no evidence for significant differences in latencies to reach N between the punishers (NOTH vs AIR: t=-0.29, AIR vs ELEC: t=0.05, NOTH vs ELEC: t= -0.20). Finally, latencies to P were significantly smaller than latencies to N, regardless of the punisher. Results are detailed at the punisher and the cue levels in **Figure 6a**.

NoGo responses to the conditioned cues

There was a significant effect of the punisher on cows' probability to display NoGo responses to the cues ($p<0.001$, $\chi^2=23.0$, df=2). Pairwise comparisons revealed that the probability to display NoGo responses was significantly lower for NOTH-cows compared with AIR-cows (OR=0.42, 95% CI [0.23-0.77], p=0.014, t=-2.82, df=1954 for all pairwise comparisons), and for AIR-cows compared with ELEC-cows (OR=0.24, 95% CI [0.08-0.71], p=0.029, t=-2.60). Additionally, the probability to display NoGo responses was significantly lower for NOTH-cows compared with ELEC-cows (OR=0.10, 95% CI [0.05-0.19], p<0.001, t=-7.23). There was also a significant effect of the cue type on cows' probability for NoGo responses: cows were significantly less likely to display NoGo responses to P than to N (OR: 67, CI [23-196], p<0.001, χ^2 =467.4, df=1, t=7.67).

Proportion of correct responses

There was a significant interaction between the punisher and the cue type on cows' proportion of correct responses to the cues ($p<0.001$, $\chi^2=146.3$, df=2). Pairwise comparisons showed no statistical evidence that NOTH-cows displayed significantly more correct responses to P than AIR-cows (OR=2.67, 95% CI [1.17-6.07], t=2.34, df=1974 for all pairwise comparisons), but they revealed that AIR-cows displayed significantly more correct responses to P than ELEC-cows (OR=3.12, 95% CI [1.53-6.33], t=3.14). NOTH-cows also displayed significantly more correct responses to P than ELEC-cows (OR=8.31, 95% CI [3.77-18.32], t=5.25). Furthermore, NOTH-cows displayed significantly less correct responses to N than AIR-cows (OR=0.37, 95% CI [0.19-0.73], t=-2.86), and AIR-cows displayed significantly less correct responses to N than ELEC-cows (OR=0.28, 95% CI [0.14-0.58], t=-3.42). Likewise, NOTH-cows displayed significantly less correct responses to N than ELEC-cows (OR=0.11, 95% CI [0.05- 0.22], t=-6.12). Results at the cue level are detailed in **Figure 6b**.

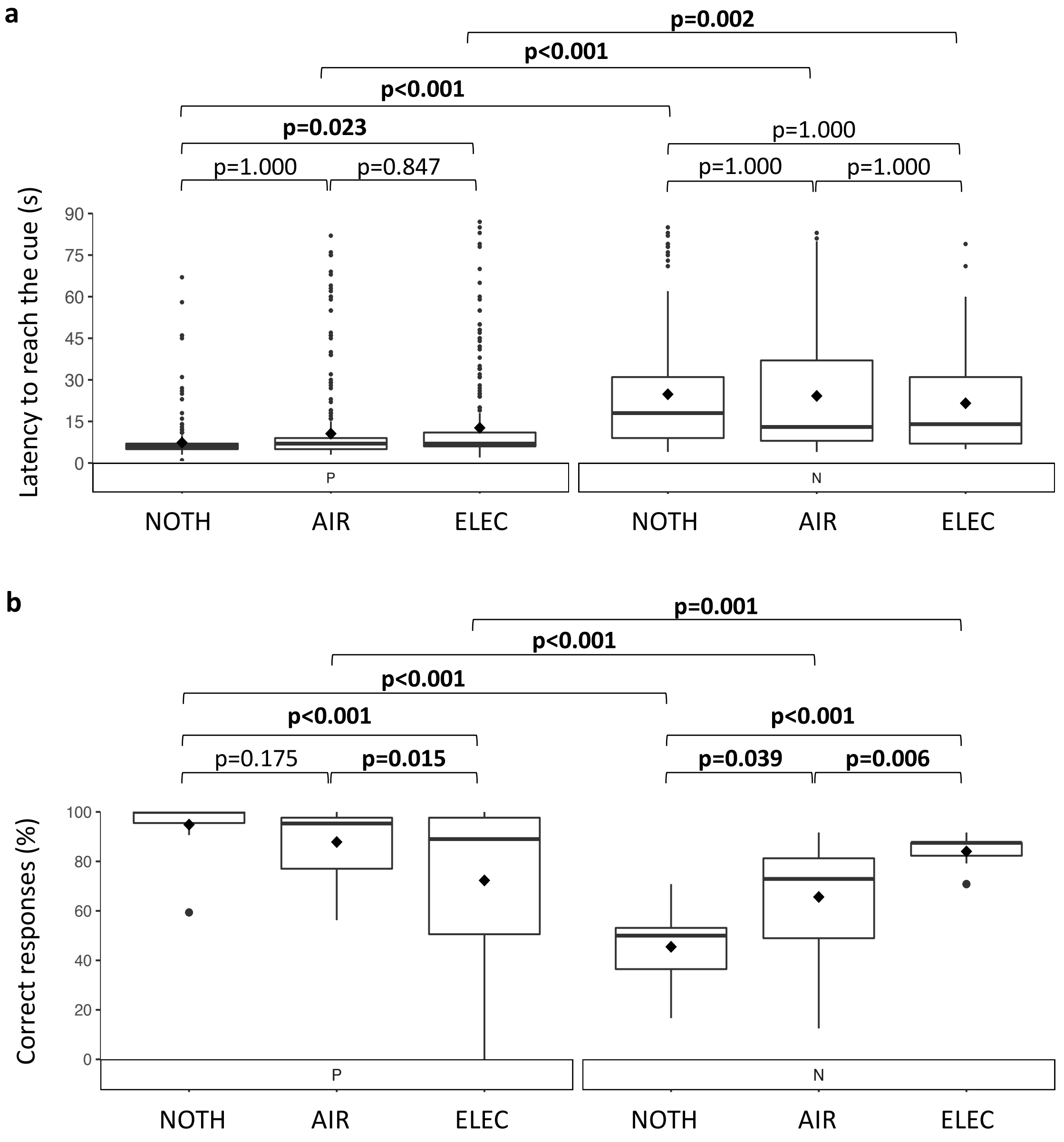


Figure 6. The graph depicts the box plots of (**a**) the latencies to reach the conditioned cues for Go-responses and (**b**) the proportion of correct responses to the conditioned cues according to the punisher across the eight training sessions. For each box, ◆ represents the mean value. NOTH: punisher is the absence of a reward; AIR: punisher is an air puff; ELEC: punisher is an electric shock. P: positive cue; N: negative cue. Significant p-values are written in bold.

Internal validity

Regardless of the punisher, cows that reached the training criterion went significantly faster to P than to Ap (V=44), to Ap than to A (V=29) and to A than to An (V=1). However, there was no statistical evidence that cows reached An significantly faster than N (V=21) Results are detailed in **Figure 7**.

At the punisher level, there was no statistical evidence that NOTH-cows reached P significantly faster than A (adjusted latency to A: 0.32 ± 0.253, IQR=0.39, p=0.281, V=12) but NOTH-cows tended to reach A faster than N (p=0.094, V=2). AIR-cows reached A significantly slower than P (0.29 ± 0.072, IQR=0.42, p=0.004, V=0) and faster than N (p=0.004, V=0). ELEC-cows reached A significantly slower than P (0.91 ± 0.060, IQR=0.0, p=0.004, V=0), but there was no statistical evidence that they reached A significantly faster than N (p=0.371, V=0).

Sensitivity

Within cows who met the training criteria, there was no statistical evidence for significant differences in the signed area *SA* between NOTH-cows and AIR-cows (NOTH: 0.33 ± 0.618, IQR=2.30; AIR: -0.20 ± 0.181, IQR=0.92 ± 0.140; p=0.776, W=30), but *SA* was significantly smaller for AIR-cows compared to ELEC-cows (ELEC: *SA*=0.96 ± 0.140, IQR=0.68; p<0.001, W=87). Additionally, there was no statistical evidence for significant differences in *SA* between NOTH- and ELEC-cows (p=0.355, W=21).

At the ambiguous cue level, there was no statistical evidence that the adjusted latencies to Ap were significantly smaller for NOTH-cows compared with AIR-cows (W=24), but the adjusted latencies to Ap tended to be smaller for AIR-cows compared with ELEC-cows (W=21.5). Additionally, the adjusted latencies to Ap tended to be smaller for NOTH-cows compared with ELEC-cows (W=12.5). Similarly, there were no statistical evidence that the adjusted latencies to A were significantly smaller for NOTH-cows compared with AIR-cows (W=19), but the adjusted latencies to A were significantly smaller for AIR-cows compared with ELEC-cows (W=2). Similarly, the adjusted latencies to A were smaller for NOTH-cows compared to ELEC-cows (W=11). Finally, there was no statistical evidence for significant differences in adjusted latencies to An between the punishers (NOTH vs AIR: W=36, AIR vs ELEC: W=32, NOTH vs ELEC: W= 32). Results are detailed at the punisher and the ambiguous cue levels in **Figure 8**.

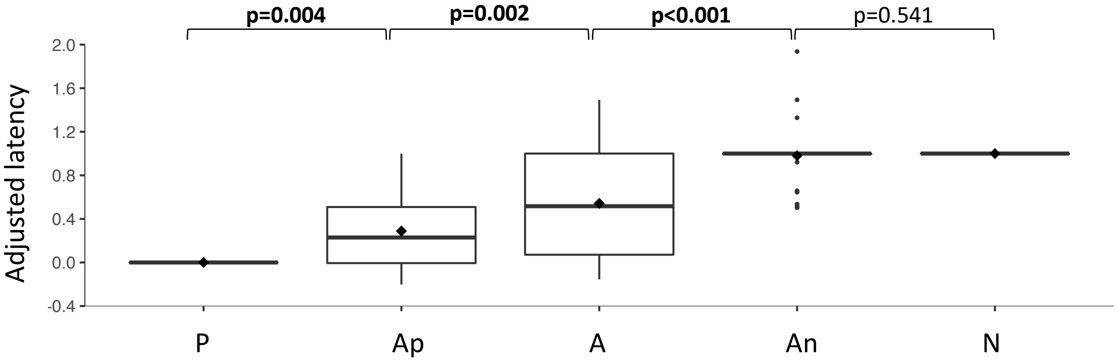


Figure 7. Adjusted latency according to the cue type during the first testing session. The graph represents the overall adjusted latencies obtained for all cows regardless of the punisher. For each box, ◆ represents the mean value. P: positive cue; Ap: ambiguous positive cue; A: ambiguous cue; An: ambiguous negative cue; N: negative cue. NOTH: punisher is the absence of a reward; AIR: punisher is an air puff; ELEC: punisher is an electric shock. Significant p-values are written in bold.

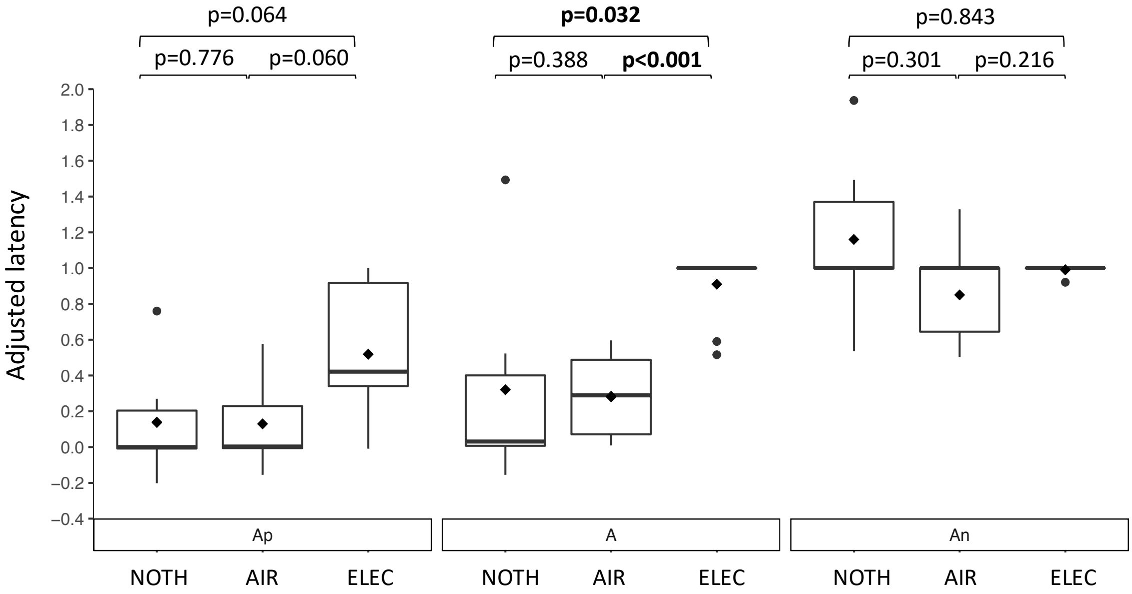


Figure 8. The graph depicts the box plots of the adjusted latency to the ambiguous cues according to the punisher during the first testing period. For each box, ◆ represents the mean value. NOTH: punisher is the absence of a reward; AIR: punisher is an air puff; ELEC: punisher is an electric shock. Ap: positive ambiguous cue; A: ambiguous cue; An: ambiguous negative cue. Significant p-values are written in bold.

Repeatability

Regardless of the punisher, there was no statistical evidence for significant differences in adjusted latencies to reach Ap (NOTH: V=8, AIR: V=19, ELEC: V=27) or An (NOTH: V=14, AIR: V=1, ELEC: V=0) between the first and the second testing periods.

However, adjusted latencies to A were significantly higher for AIR-cows in the second testing period compared with the first (V=0), while there was no evidence for significant differences in adjusted latencies to reach A between the two testing periods for NOTH-cows (V=5) and ELEC-cows (V=7). **Figure 9** provides an overview of these results. Furthermore, there was no statistical evidence for significant correlations between adjusted latencies to the ambiguous cues in the first testing period on the one hand and adjusted latencies to the ambiguous cues in the second testing period on the other hand – when cows reached the ambiguous cues (**Table 1**).

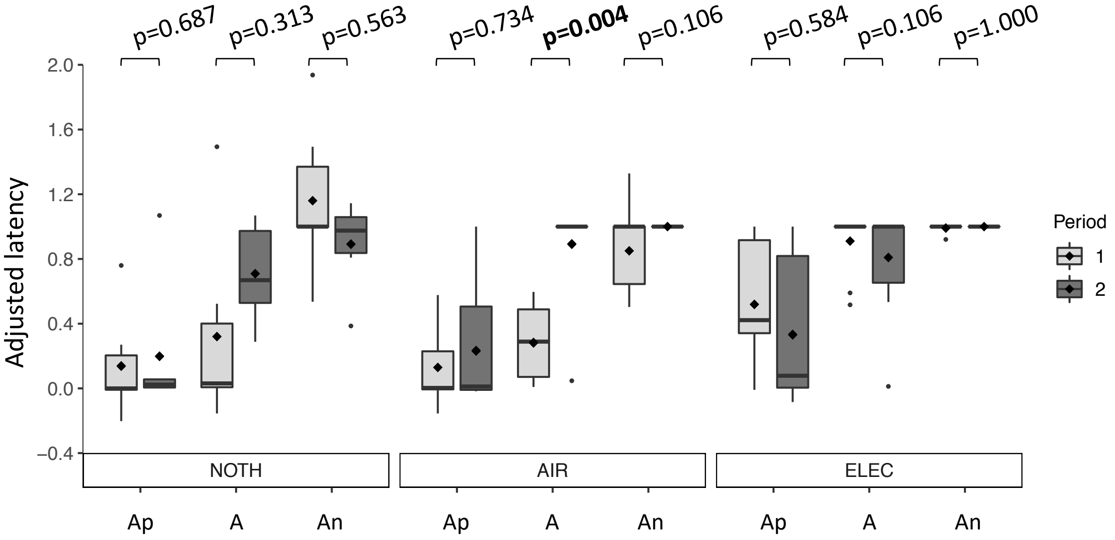


Figure 9. Adjusted latency to the ambiguous cues according to the punisher and the testing period. The higher the adjusted latency, the slower the cow reached the cue. For each box, ◆ represents the mean value. NOTH: punisher is the absence of a reward; AIR: punisher is an air puff; ELEC: punisher is an electric shock. Ap: ambiguous positive cue; A: truly ambiguous cue; An: ambiguous negative cue. Significant p-values are written in bold.

Table 1. Spearman's rank correlation coefficients (and respective p-values and S values) between adjusted latencies during the first testing session and adjusted latencies during the second testing session for each ambiguous cue according to the punisher. Ap: ambiguous positive cue; A: truly ambiguous cue; An: ambiguous negative cue. NOTH: punisher is the absence of a reward; AIR: air puff; ELEC: electric shock.

PUNISHER	Ap	A	An
NOTH	-0.37 (p=0.497, S=48)	0.49 (p=0.356, S=18)	0.03 (p=0.957, S=34)
AIR	-0.12 (p=0.776, S=134)	0.55 (p=0.127, S=54)	NA*
ELEC	0.17 (p=0.643, S=137)	0.15 (p=0.681, S=140)	NA*

* Coefficients with p-values and S-values are unavailable due to null variance in adjusted latencies to reach An for AIR-cows and ELEC-cows during the second testing period.

Discussion

The study aimed to develop for dairy cows a feasible, valid and sensitive JBT - the repetition of which does not lead to ambiguity loss. To this end, we investigated the influence of different punisher/reward combinations on the responses of cows trained and tested repeatedly in a JBT paradigm.

Feasibility of the Judgement Bias Tasks

The feasibility of the JBTs, here defined as the cow's aptitude to learn the discrimination task, was assessed based on both learning success and contingency learning. The effect of the reward/punisher combination on the JBT's feasibility yielded mixed results. Hypothesis 1 is, therefore, only partially supported by the present study.

We did not find statistical evidence for differences in learning success between cows exposed to different combinations of reinforcers. This result suggests that all three combinations of reinforcers can be successfully used to train the majority of dairy cows on spatial discrimination relatively fast (i.e. within 56 trials in 8 sessions). This result is consistent with previous findings showing that dairy cows could successfully learn to discriminate between two conditioned spatial cues associated with either concentrates or 'no-reward' (Crump et al. 2021). Former studies also demonstrated that calves were able to discriminate between two spatial cues paired either with a milk reward or an air puff (e.g. Lecorps et al. 2018).

Although learning success appeared not affected, the combination of reinforcers significantly influenced cow contingency learning (i.e. Go-to-P and NoGo-to-N). The greatest difference in responses to the conditioned cues was observed between NOTH-cows and ELEC-cows. NOTH-cows were the most likely to reach both conditioned cues and they displayed the lowest number of correct responses to N; while ELEC-cows were the least likely to reach both conditioned cues and they displayed the lowest number of correct responses to P. NOTH-cows were also faster to reach P than ELEC-cows. Additionally, AIR-cows made more correct responses to P and less correct responses to N than ELEC-cows. The 'no-reward' punisher was thus associated with the worse NoGo-to-N contingency learning, while the electric shock was associated with the worse Go-to P contingency learning. In comparison to the air puff, the 'no-reward' punisher hence encourages active responses to N while the electric shock inhibits active responses to P. We, therefore, question the suitability of 'no-reward' and electric shock as appropriate punishers in Go/NoGo JBTs for adult cows. Instead, we recommend using an air puff to design relatively feasible JBTs for dairy cows. There are several possible and compatible explanations for these findings.

The differences in probabilities of reaching the conditioned cues based on the punisher could arise from differences in affective responses to the punishers themselves. These punisher-induced affective states may be associated with distinct behaviours. We hypothesise that the air puff and electric shock elicited fear – an emotional state experienced in anticipation of threatening or dangerous stimuli (Papini et al. 2019). AIR- and ELEC-cows may hence display NoGo responses to N to avoid subsequent negative outcomes. Avoidance behaviours are expected in animals experiencing fear (Gray and McNaughton 2000). We hypothesise that 'no-reward' instead elicited frustration – *'a temporary state that results when a response is nonreinforced […] in the presence of a reward expectancy'* (Amsel 1992). A vast behavioural repertoire has been linked to frustration, including goal-oriented behaviours such as aggressiveness (Dantzer et al. 1980) and response invigoration (Papini et al. 2019). Response invigoration is characterised by an increased motivation to engage in the dominant behaviour – here, reaching the feeder. Punishers inducing avoidance rather than goal-oriented behaviours are likely to ensure more efficient learning of the NoGo response. Response suppression (i.e. NoGo) to a punisher inducing frustration is likely to be less natural than response suppression to a punisher inducing fear or pain. Therefore, the congruence between the expected behavioural response to the conditioned cue and the punisher-driven affective state is likely to reduce the required training period, thereby leading to a more feasible Go/NoGo JBT. In this respect, species-specific differences should be considered, as the adaptive responses to fear may differ from one species to the next. For instance, mice more readily learnt the Go-to-N contingency than the NoGo-to-N contingency, while rats more readily learnt the NoGo-to-N contingency (Jones et al. 2017).

Additionally, differences in probabilities to reach the conditioned cues created by the punisher may reflect differences in speed-accuracy trade-offs made by the cow during decision-making. In sensory discrimination tasks, accuracy and speed of decision are two key conflicting factors that contribute to decision quality – decisions taken faster more likely leading to errors (Chittka et al. 2009). In our experiment, ELEC-cows may have perceived the cost of making an error as higher than NOTH-cows – which may have led them not to respond to certain P trials. As a result, decision-making in ELEC-cows may have predominantly relied on accuracy gain over speed gain, while the opposite may be true for NOTH-cows.

Discrepancies in contingency learning may also have arisen from differences in punishers' aversiveness. We hypothesise that cows experienced the electric shock as more aversive than the air puff or 'no-reward'. Manipulation of the punisher aversiveness has been found to influence the acquisition of the NoGo-to-N contingency. For instance, rats subjected to electric shocks of high intensity learn to avoid the negative cue faster than rats subjected to electric shocks of lower intensity (Feigley and Spear 1970). The validity of this theory could be investigated by assessing cows' affective arousal during their exposure to the punishers, via analyses of heart rate or thermography data for instance (Sinha et al. 1992; Clay-Warner and Robinson 2015).

Two main elements may explain the fact that we detected an effect of the reward/punisher combination on cow learning contingency (number of correct responses) but not on learning success. First, our measure of learning speed was by definition dependent upon our training criterion. Each cow was allowed to make one incorrect response to either P or N over two consecutive sessions to be considered trained. NOTH- and ELEC-cows may thus have reached the training criterion in a similar timespan while predominantly displaying incorrect responses to N and P, respectively. Although our training criterion is in range with other criteria found in literature (Hintze et al. 2018; Lecorps et al. 2018), opting for a different criterion (e.g. one incorrect response to N only) may have led to more contrasted results in terms of learning speed according to the allocated punisher. Second, the limited number of cows used in our experiment may have impacted the statistical power of our test – thereby reducing our ability to spot a potential effect of the punisher on cow learning success. Replication studies at a larger scale are, therefore, required to draw more reliable conclusions on the effect of the punisher/reward combination on JBT feasibility.

Careful consideration of factors other than the reward/punisher combination may help optimise discrimination training in JBT paradigms. The congruence between the selected punisher and the cue modality could facilitate the acquisition of the discrimination task. In rats, for instance, pairing the ingestion of a toxin with tasty water (i.e. gustatory cue) led to aversive reactions to water consumption but not pairing the ingestion of a toxin with noisy and bright water (i.e. auditory-visual cue) (Garcia and Koelling 1966). Evolutionary mechanisms may explain that internal discomfort is more readily associated with gustatory over auditory cues (Garcia and Koelling 1966). We advise that the rationale behind the choice of a punisher integrates the nature of the expected behavioural response as well as the cue modality (**Figure 10**). Additionally, allowing the animal to self-initiate the trials by displaying a natural behaviour (Jones et al. 2018) may also reduce the training duration. In their study, Jones et al. (2018) developed an automated JBT during which they trained rats to self-initiate a trial by nose-poking into a feed trough. Depending on the tone of a sound cue, rats learnt to leave their head in the feed trough for 2 s to get a feed reward, or to remove their head from the trough to avoid an air puff. Compared with results obtained in other studies where rats could not initiate the trials on their own (e.g. Parker et al. 2014), Jones and colleagues (2018) reported that their rats needed fewer training sessions to be considered trained. Translation of this task to cows seems promising since dairy cows in modern commercial farms are generally used to receive concentrates in their home pen from an automatic feeder and based on voluntary approach. Voluntary testing would also alleviate some of the feasibility issues encountered in the present study. First, habituation time would be significantly reduced since training would occur in a familiar environment with a familiar device. Second, automation of the delivery of concentrates/air puff would considerably reduce the time allocated by researchers to training. Third, self-initiation of trials would give cows control over the task which is expected to guarantee motivation to participate (Hintze et al. 2018).

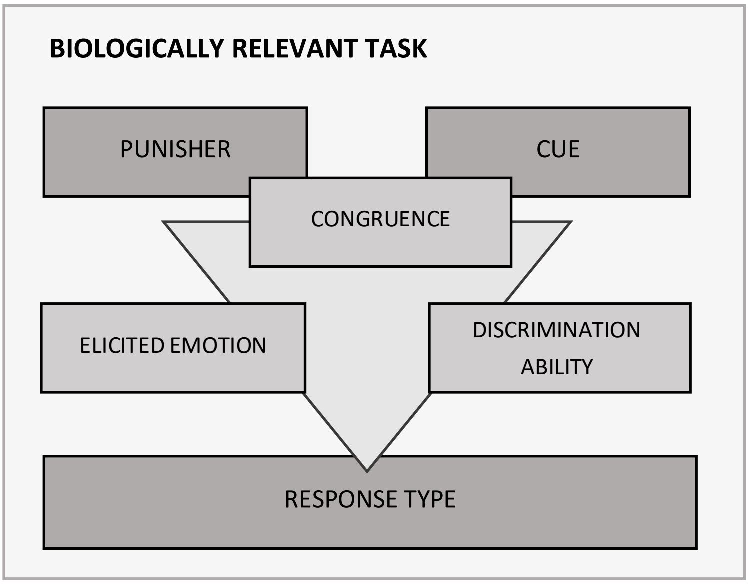


Figure 10. Punisher-Cue-Response triad. The rationale behind the selection of a punisher within the context of a JBT should integrate the existing congruence between the punisher and the cue modality. The chosen cue modality should be evolutionary relevant and match the species-specific discrimination ability. The punisher should elicit an emotion in line with the expected response

Internal validity of the Judgement Bias Tasks

The internal validity of a spatial task was assessed by investigating the pattern of cow responses (expressed in adjusted latencies) from the positive to the negative cues for all punishers. Overall, cows reached the adjacent spatial cues at significantly different speeds regardless of the punisher used, which is in line with our hypothesis 2. This finding corroborates the idea that cows are able to discriminate between two cues separated from 1.6 m, and is in agreement with another study assessing the horizontal visual acuity of Friesian-Holstein dairy bulls at 1.6 c/deg (Rehkämper et al. 2000). Furthermore, when combining all punishers, the monotonic graded pattern of responses from the positive to the negative cues was effectively observed, which indicates that cows exhibited more optimistic responses to ambiguous cues positioned closer to the positive conditioned cue, and vice versa. This result supports the idea that cows perceived the positions of the feeder between the conditioned cues as ambiguous rather than novel (Hintze et al. 2018) and further validates the use of a spatial Go/NoGo JBT in dairy cows.

Nonetheless, differences in adjusted latencies to reach the truly ambiguous cue A relative to P and N were noted. While NOTH-cows reached A and P at a similar speed (statistically speaking) and ELEC-cows reached A and N at a similar speed, AIR-cows reached A slower than P and faster than N. We suggest that, within each JBT task, individual responses to the ambiguous cue are partially determined by the asymmetry in the affective ladder delineated by the reinforcers (from rewarding/pleasant to aversive/unpleasant). Affective asymmetry may lead to a phenomenon known as 'peak shift' (Roelofs et al. 2016) which results in biased responses to the ambiguous cues toward the most salient cue – i.e. the cue associated with the reinforcer of the highest value, be it negative or positive (**Figure 11**). For NOTH-cows, the perceived positive value of the concentrates may have largely exceeded the perceived aversive value of the absence of reward, resulting in responses to the ambiguous cue biased toward P. In contrast, the aversive value of the air puff appears to balance the positive value of the concentrates, leading AIR-cows to reach A at an intermediary speed. Finally, the aversive value of the electric shock may have outweighed the rewarding value of the concentrates, leading to responses to ambiguous cues biased to N in the ELEC-cows. Above a certain intensity, cow motivation to avoid an electric shock becomes greater than feed motivation – as demonstrated by Lee et al. (2007) who successfully trained cows to stop reaching a feeding trough filled with hay by using an electric shock. Further research is required to refine our understanding of how the punisher/reward balance may influence the JBT's internal validity, as well as its sensitivity.

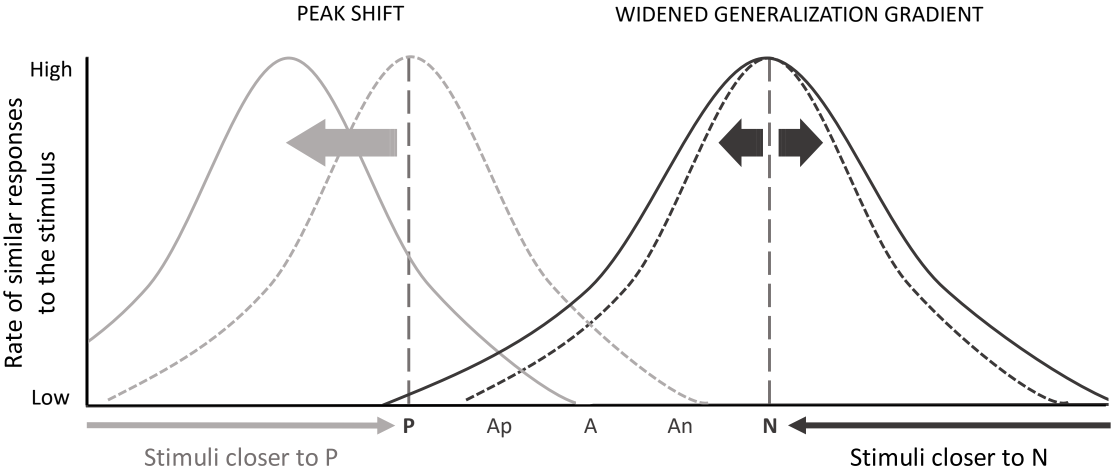


Figure 11. Peak shift phenomenon and generalization gradient expansion (adapted from Roelofs et al. 2016). When the negative value of N outweighs the rewarding value of P, the generalization gradients around P may shift away from N. When the threat linked to N is severe, the discrimination thresholds between perceptually similar cues (for instance A, An and N) may decrease.

Sensitivity of the Judgement Bias Tasks

The sensitivity of each JBT was assessed by calculating the divergence (expressed using signed area) between cow adjusted latencies to reach the cues and the expected unbiased baseline (linear and monotonic graded pattern). Divergence from this line before the application of any affective treatment reflects a construct punisher-driven bias that can mask treatment-induced judgement bias – therefore altering the sensitivity of the task. In line with our hypothesis 3, the combination of reinforcers influenced the sensitivity of the task. ELEC-cows responses to the ambiguous cues were negatively biased, while AIR-cows responses to the ambiguous cues were the closest to the expected theoretical line. In other words, ELEC-cows exhibited a negative baseline judgement bias. In our experimental design, the electric shock is therefore not associated with a sensitive JBT, and may lead to a failure to detect a treatment-induced negative affective shift. By contrast, a 5 bar air puff is a punisher suitable for a valid and sensitive JBT in cows, since the punisher-induced judgement bias was minimal for AIR-cows. Unexpectedly, NOTH-cows also exhibited an overall negatively biased baseline. The positive area between the experimental curve of NOTH-cows responses to the ambiguous cues and the theoretical unbiased baseline was thus larger than the negative area (as noticeable in **Figure 12**). This finding can partially be explained by the fact that 3/6 trained NOTH-cows kept responding to N during testing, which resulted in adjusted latencies to An above 1. This result could also be attributed to the fact that trained NOTH-cows associated the outcomes of the unrewarded An to the outcome of N – and thus stopped responding to An in the same way they learnt to suppress their behavioural response to N. This result is similar to previous findings demonstrating that JBTs involving 'no-reward' punishers are less sensitive than JBTs involving more salient punishers (Lagisz et al. 2020) since the 'no-reward' punisher-induced judgement bias around An may hamper the detection of treatment-induced negative shifts in animal affective states. Therefore, we hypothesise that the sensitivity of such tasks could be increased by positioning An further away from N, in an attempt to reduce individual expectations of a negative outcome associated with An.

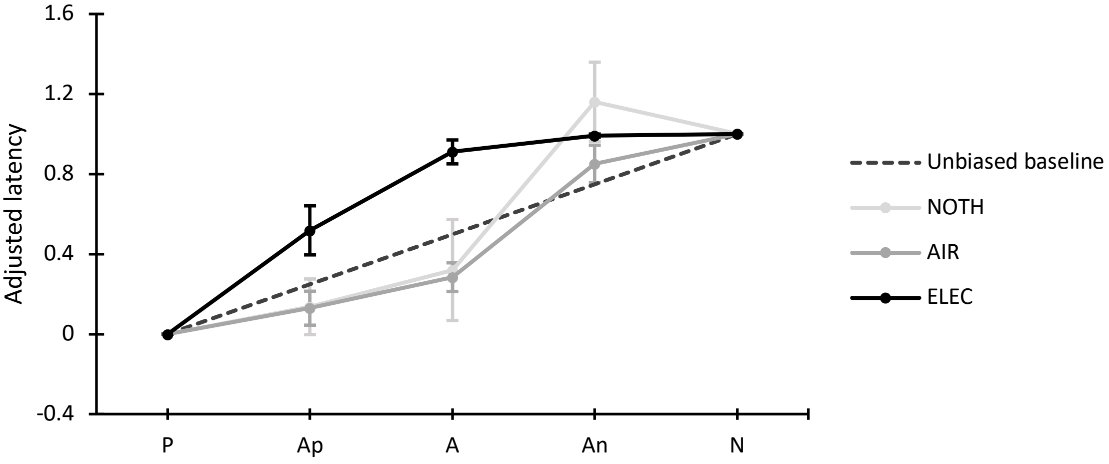


Figure 12. Adjusted latencies according to the cue type and the punisher for the first testing period. NOTH: punisher is the absence of a reward; AIR: punisher is an air puff; ELEC: punisher is an electric shock.

Differences in response to the ambiguous cues may have emerged from variations in cow affective states. The type of punisher may have impacted 1) cow perception of the JBT as a whole and 2) cow affective states within the JBT. First, the JBT has often been considered as a potential cognitive enrichment that could enhance animal welfare (Roelofs et al. 2016). However, although habituated, cows may still dislike being isolated and regularly handled by the experimenters. Their repeated exposure to aversive experiences when making incorrect responses to N during training may also negatively influence cow welfare – and consequently alter their motivation to engage in the judgement bias task. Thus, since cows were exposed to the training procedures up to three times per week in our experiment, we cannot rule out the possibility that the JBT training itself induced an affective shift in cows. Based on the assumption that the electric shock was more aversive than the air puff or 'no-reward', ELEC-cows may hence have been in a worse welfare state than AIR-cows and NOTH-cows, which led them to make more pessimistic decisions. Future studies are required to elucidate the influence of JBT procedure on animal welfare. This could be assessed, for instance, by investigating how motivated animals are to participate in the JBT. Behavioural and physiological differences between groups of animals exposed or not to the JBT could also be scrutinised to assess the long-term effect of JBT on animal affective states. Of note, the animal perception of the JBT is also likely to evolve – depending on how fast the individuals are able to cope with the exposure to a threat (i.e. the negative cue signalling the punisher). Over time, one might expect that the negative impact of the negative conditioned cue will decrease as the animals learn to avoid the associated punisher. The animal perception of the JBT could thus also improve over time, as the animals gain control over the task.

Second, even trained cows may experience negative affect inside the testing arena when faced with the negative cue that they may perceive as a threat. The (supposedly) relatively high aversiveness of the electric shock may have induced anxious-like affective states in cows, resulting in pessimistic responses to the ambiguous cues. Anxiety has been linked with negative interpretation bias, judgement bias and decision-making bias in humans (Blanchette and Richards 2010) and animals (e.g. rats: Burman et al. 2009; dogs: Karagiannis et al. 2015). Thus, compared with NOTH- and AIR-cows, ELEC-cows may have interpreted the ambiguous cues as more negative ('interpretation bias') and overestimated the likelihood of the ambiguous cues to be associated with a negative outcome ('judgement bias'). In mechanistic terms, the supposedly higher aversiveness of the electric shock compared with the air puff or the absence of reward may have widened the generalisation gradient around N, resulting in more risk-averse decisions (decision-making bias) to the ambiguous cues. From an evolutionary perspective, it is more advantageous to react similarly to a wide range of stimuli with characteristics common to these of a stimulus associated with a severe threat (e.g. predator attack), while sharper discrimination may be more advantageous when the threat is less severe (e.g. insect bite). The influence of N on animal short-term affective states could be investigated by analysing individual behavioural indicators of affect (e.g. ear postures and vocalisations) when the latter are faced with N within the testing arena.

Factors other than the combination of reinforcers must be taken into account to design a sensitive JBT. Following the example of previous JBTs designed for herbivores, we opted for a spatial Go/NoGo discrimination task. However, a recent meta-analysis from Lagisz et al. (2020) revealed that the most sensitive JBTs rely on active choice tasks and involve either tactile or auditory cues. Considering the relatively wide hearing range of cows (Heffner and Heffner 1983), researchers may thus consider developing an auditory active choice task suitable for dairy cows rather than a spatial Go/NoGo paradigm. In active choice tasks, cows could be taught to press either a right or left panel in response to the conditioned cues – as performed by calves on a double demand operant conditioning setup (Webb et al. 2015). Moreover, Lagisz et al. (2020) demonstrated that JBTs using reward/smaller reward as combination of reinforcers are more sensitive than JBTs using reward/punisher as reinforcers. Future studies should therefore investigate cows' discrimination ability among different reward quantities, and subsequently determine whether using a smaller reward instead of an air puff leads to more sensitive (and potentially feasible) JBTs in dairy cows than those presented in this study.

Repeatability of the Judgement Bias Tasks

The repeatability of each JBT was assessed based on 1) cows' differences in adjusted latencies to reach the ambiguous cues between the first and the second periods of testing and 2) individual consistency in response to the ambiguous cues between both testing periods. In contradiction with hypothesis 4, not all three JBTs appeared repeatable – despite the inclusion of a washout period between the two testing periods.

Our assessment of repeatability for the JBTs associated with 'no-reward' and the electric shock was inconclusive. While there was no statistical evidence that NOTH- or ELEC-cows reached the ambiguous cues slower during the second testing period compared to the first testing period (although means did differ, hence suggesting that our sample size was too small to pick up differences), there was also no statistical support for consistency in individual responses between both periods. Non-significant results being no proof of an absence of effect, we are, here, unable to reject or confirm the hypothesis that these JBTs are repeatable. For a JBT to be characterised as repeatable, two requirements must be met – none of which is self-sufficient. First, the population study must, on average, reach the ambiguous cues in a similar timespan for every testing session. Second, each individual must be consistent in their responses to the ambiguous cues (i.e. a relatively optimistic individual should remain relatively more optimistic than the other individuals of the study population in every testing session).

Nowadays, however, JBTs' repeatability is often investigated based on the sole assessment of the first requirement. While such a strategy can be validly used to demonstrate that a task is not repeatable (Doyle et al. 2010b), it is not sufficient to demonstrate that the task *is* repeatable. In the latter case, a JBT could be falsely advertised as repeatable despite little correlation in individual responses to the ambiguous cues across testing sessions. As recommended elsewhere (Carreras et al. 2015), we, therefore, encourage researchers to assess JBT repeatability by investigating individual consistency in response to the ambiguous cues – in addition to exploring differences in response means over testing sessions.

Finally, we were able to demonstrate that the JBT associated with the air puff was not repeatable. AIR-cows reached the truly ambiguous cue slower during the second testing period compared with the first testing session. This finding could either indicate a negative shift in cow perception of the JBT between the two testing sessions or demonstrate that cows progressively learnt that the ambiguous cues were not rewarded. The second assumption seems the most plausible since our cows were subjected four times to the test, and previous studies already reported a loss of ambiguity after repeated testing (Doyle et al. 2010b; Scollo et al. 2014; Karagiannis et al. 2015). The loss of ambiguity in our study could also be explained by the low reference:ambiguous trial ratio compared with other studies (e.g. 4:3 versus 50:3 for rodents in Hintze et al. 2018). Therefore, while the air puff is associated with a relatively feasible and sensitive judgement bias task for dairy cows, we recommend reducing the number of testing sessions to minimise ambiguity loss. Our experimental design could also be combined with one of the following strategies to counteract ambiguity loss.

Partial reinforcement of the conditioned cues has been proposed to minimise ambiguity loss (Roelofs et al. 2016). In calves, Neave et al. (2013) applied a partial reinforcement ratio schedule during training – by progressively reducing the positive reinforcement by 50 %. As a result, the outcome of unrewarded ambiguous cues remained unclear to calves since the uncertainty of the reinforcement value (i.e. positive or less positive/neutral) of Go-response to P was already introduced at the end of the training period. Training animals to associate ambiguous cues with pre-determined reinforcement ratio would also eliminate the risk of ambiguity loss (Lecorps et al. 2019). Lecorps et al. (2019) recently developed an innovative spatial judgement paradigm during which calves were directly trained to discriminate among the usual reference cues and three ambiguous cues. Responses to the ambiguous cues were reinforced based on the expected probability of reinforcement according to the cue position. For instance, calves were trained to associate the ambiguous cue positioned exactly in between the positive and the negative conditioned cues with a probability of getting a reward or a punisher of 50 %. Affect-driven judgement bias due to a specific treatment was therefore assessed by comparing latencies to reach the ambiguous cues before and after the treatment induction.

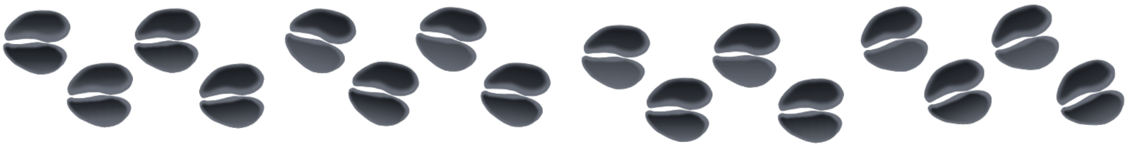
As much as possible, repeatable JBT procedures should be developed to assess treatment-induced affective states within longitudinal studies. Longitudinal assessment of animal judgement bias allows to control for the effect of endogenous factors – otherwise known to influence individual responses to ambiguous cues in cross-sectional studies. In future studies aiming at developing a JBT suitable for dairy cows, we thus suggest the use of an auditory Go/NoGo discrimination task in which cows learn the outcome of three ambiguous sound cues at the concentrate station. An air puff could be used as a punisher to ensure the acquisition of the NoGo-to-N contingency.

Conclusion

The aim of our study was to design a relatively feasible, valid and sensitive JBT for adult dairy cows that could be used in longitudinal studies. Here, we demonstrated the validity of using spatial JBTs for dairy cows and confirmed the effect of the combination of reinforcers on JBT feasibility. Despite having no detectable effect on learning success, the combination of reinforcers influenced cow contingency learning during training. Cows displayed more Go responses to both conditioned cues when using ‘no-reward’ as a punisher, while cows displayed more NoGo responses to both cues when using an electric shock. We also demonstrated the impact of the combination of reinforcers on JBT sensitivity, and we identified the JBT associated with the air puff as the most sensitive JBT within our study. Although unfit for longitudinal studies, spatial discrimination tasks using concentrates and air puff as reinforcers constitute valid, sensitive and relatively feasible JBTs for dairy cows. Other methodological aspects, like the type of task, should be considered in the future to refine this already promising JBT for dairy cows.

Chapter 4

Housing and personality effects on judgement and attention biases in dairy cows



Scientific Reports (2021) 11(1), 1-18.
https://doi.org/10.1038/s41598-021-01843-w

L. Kremer[1,2*], J.D. Bus[1], L.E. Webb[1], E.A.M. Bokkers[1], B. Engel[3], J.T.N. van der Werf[2],
S.K. Schnabel[3], C.G. van Reenen[1,2]

[1] Animal Production Systems Group, Wageningen University & Research, Wageningen, the Netherlands

[2] Livestock Research, Wageningen University & Research, Wageningen, the Netherlands

[3] Biometris, Wageningen University & Research, Wageningen, the Netherlands

Abstract

Affective states can be inferred from responses to ambiguous and threatening stimuli, using Judgement Bias Tasks (JBTs) and Attention Bias Tasks (ABTs). We investigated the separate and interactive effects of personality and housing conditions on dairy cattle affective states. We assessed personality in 48 heifers using Open-Field, Novel-Object and Runway tests. Personality effects on responses to JBT and ABT were examined when heifers were housed under reference conditions. Heifers were subsequently housed under positive or negative conditions, and housing effects on animal responses in both tasks were investigated while controlling for personality. A Principal Component Analysis revealed three personality traits labelled Activity, Fearfulness and Sociability. Under reference conditions, personality influenced heifers' responses to JBT and ABT, therefore questioning the tasks' generalizability across individuals. Against expectations, housing did not influence responses to JBT and heifers in the negative conditions looked at the threat later than heifers in the positive or reference conditions. More research is warranted to confirm the validity and the repeatability of JBT and ABT as measures of affective states in dairy cows.

Keywords

Affective states – Judgement bias – Attention bias – Housing – Personality

Introduction

The welfare of dairy cows is a major societal concern (Weary and Von Keyserlingk 2017), insomuch that consumers are willing to pay more for products obtained from animals whose welfare has not been compromised (Spain et al. 2018). The concept of animal welfare initially revolved around major threats to animal survival (e.g. disease, thirst…), but it progressively evolved as research progressed and public values changed (Ohl and van der Staay 2012). Nowadays, the definition of animal welfare includes the notion of affective states (Duncan 2004; Weary and Robbins 2019), which reflects the animal subjective experience of events. Animal welfare is now considered optimal when the balance between positive and negative affective states is overall positive (Green and Mellor 2011). Adequate evaluation of cow welfare, therefore, requires valid measures of cow affective states (Boissy et al. 2007; Watanabe 2007; Yeates and Main 2008). Several methodologies have been developed to this end, based on research in human cognitive psychology (Dolan 2002).

Animal affective states can be inferred from biases in cognition (Mendl et al. 2009). Two cognitive biases, the judgement and attention biases, have been assessed in farm animals using standardised paradigms (Doyle et al. 2010a; Baciadonna and McElligott 2015; Lee et al. 2016, 2017). Judgement biases, which reflect affect-driven shifts in the interpretation of ambiguous stimuli (Paul et al. 2005), are assessed using Judgement Bias Tasks (JBTs) (Harding et al. 2004). The JBT principle relies on the idea that animals in positive affective states judge ambiguous situations more positively (i.e. more optimistically) than animals in negative affective states – and vice versa (Harding et al. 2004). Attention biases, which reflect affect-driven shifts in the allocation of attention to salient stimuli (Paul et al. 2005), are assessed using Attention Bias Tasks (ABTs) (Brilot and Bateson 2012). The ABT principle relies on the idea that attention to threat is influenced by one’s affective states (MacLeod et al. 1986). Cows and sheep, for instance, have heightened attention to threatening stimuli when in heightened anxious states (Paul et al. 2005; Lee et al. 2016). With JBTs and ABTs, researchers can investigate the effects of various treatments on animal affective states, by assessing changes in pessimism (Lagisz et al. 2020; Neville et al. 2020) and in attention to threat (Luo et al. 2019; Monk et al. 2019a; Raoult and Gygax 2019). Two main sources of variation in judgement and attention biases have commonly been explored, namely the living environment and animal personality.

The living environment influences animal affective states, and subsequent affect-driven cognitive biases, by conditioning individuals' propensity to experience positive and negative events. For captive animals, the housing conditions constitute the main aspect of their living environment. Previous studies have investigated the effects of housing conditions on cognitive biases in different farm species (Douglas et al. 2012; Stephenson and Haskell 2020), with a predominant focus on judgement bias. For instance, pigs housed in enriched conditions seemed more optimistic than those housed in barren conditions (Douglas et al. 2012). Similarly, paired-housed calves seemed more optimistic than individually-housed calves (Bučková et al. 2019). Recently, one study also revealed that housing conditions influenced attention biases in pigs (Luo et al. 2019). Negative and positive housing contrasts may, therefore, constitute promising models of animal affective states.

Beside housing conditions, personality may also modulate animal judgement and attention biases. Personality traits – defined as a correlated set of individual behavioural and physiological traits that are consistent over time and across contexts (Koolhaas et al. 1999; Finkemeier et al. 2018) – may influence animal affective states by mediating subjective experiences of events (Roelofs et al. 2016). For instance, calves characterised as 'fearful', i.e. relatively slow at reaching novel-objects and unfamiliar humans, were more pessimistic than non-fearful calves – while housed under the same conditions (Lecorps et al. 2018). Similarly, parrots characterised as neurotic, i.e. relatively excitable, fearful and non-social, had greater attention bias to an unfamiliar human than non-neurotic parrots, while housed under the same conditions (Cussen and Mench 2014). Accounting for individual variations, and in particular for personality differences, appears hence necessary to validate the use of housing contrasts as model of affective states.

Furthermore, housing and personality may exert an interactive effect on animal judgement and attention biases. Several studies revealed that housing-induced judgement biases (Asher et al. 2016; Ross et al. 2019) and attention biases (Luo et al. 2019) are dependent upon animal personality traits. For instance, Ross and colleagues (Ross et al. 2019) reported that the level of enrichment in the housing conditions of hens influenced the judgement bias responses of individuals characterised as exploratory – i.e. hens approaching a novel object relatively fast – but not the responses of individuals characterised as non-exploratory. Depending on their personality, certain sub-populations of animals may hence be more sensitive than others to specific housing conditions.

Therefore, in our study, we aimed at investigating:

1) the effect of personality traits on judgement and attention biases in dairy heifers kept under reference housing conditions (i.e. at baseline) to minimise variations in individual affective background
2) the effect of (supposedly) affectively-contrasted housing conditions on heifer judgement and attention biases by using a longitudinal approach to control for individual variation
3) individual variation consistency in heifer responses to the cognitive bias tests across (supposedly) contrasted housing conditions – either by focusing on untargeted sources of individual variation or by tentatively exploring the effects of targeted sources of individual variation, i.e. identified personality traits, on dairy heifers' cognitive bias responses

Results

Identification of dairy heifers' personality traits

Behavioural data from the personality tests of 47 heifers were suitable for PCA analysis, as reflected by the overall Kaiser-Meyer-Olkin criterion (=0.71) and each variable Kaiser-Meyer-Olkin criterion (from 0.48 to 0.77). The hypothesis of all zero correlation was rejected (Bartless's sphericity test, $p<0.001$). The first three Rotated Components (RCs) explained 80 % of the total variance. Loadings on the first three RCs are presented in **Table 1.** Heifers scoring high on RC1 explored and walked the most in the arena. Those scoring high on RC2 spent the most time in contact with the Novel Object (NO) and were the fastest to reach the NO for the first time. Finally, heifers scoring high on RC3 spent more time within the 2 m zone, i.e. close to other cows, in the runway (RW) test. For ease of reading, RCs are hereafter referred to as 'personality traits'. RC1 is labelled the Activity trait, RC2 the Fearfulness trait and RC3 the Sociability trait. The repartition of heifers per different personality trait and within each type of housing conditions is detailed in **Table 2**.

Table 1. Loadings of the behavioural measures on the 3 main rotated components (RCs). Loadings rated 'excellent' (i.e. (|values| > 0.71) are written in bold. Novel-Object (NO), Open-Field (OF), Runway (RW).

Behavioural measures	**RC1**	**RC2**	**RC3**
Number of locomotion bouts in OF and NO	**0.817**	0.314	-0.079
Time spent in locomotion in OF and NO	**0.854**	0.153	0.114
Time spent in contact with walls/floor in OF and NO	**0.843**	0.071	-0.068
Time spent in contact with NO	0.120	**0.886**	-0.128
Latency to touch NO	-0.241	**-0.841**	-0.012
Time spent within 2 m from the group in RW	-0.011	-0.085	**0.990**
Eigenvalue	2.72	1.19	0.91

Table 2. Number of heifers in each housing conditions per personality trait. Heifers were divided in two classes per personality trait based on their behavioural scores on the related personality trait in comparison to the median score. Superior (Sup.), Inferior (Inf.), RC rotated component, n number of heifers per personality class in the reference conditions, n_1 number of heifers per personality class in the positive conditions, n_2 number of heifers per personality class in the negative conditions.

Personality			**Housing conditions (number)**		
Trait (median score)	**Class**	**Definition**	**Reference ($n = n_1+n_2$)**	**Positive (n_1)**	**Negative (n_2)**
RC1: Activity (-0.08)	Active	Sup. to -0.08	23	11	12
	Inactive	Inf. or equal to -0.08	24	13	11
RC2: Fearfulness (0.23)	Fearful	Inf. or equal to 0.23	24	13	11
	Non-fearful	Sup. to 0.23	23	11	12
RC3: Sociability (0.26)	Social	Sup. to 0.26	24	10	14
	Non-social	Inf. or equal to 0.26	23	14	9

Personality effect on heifers' responses to judgement and attention bias tests under the reference conditions.

The main effect of personality on individual responses to the cognitive bias tasks was assessed in the reference conditions only, when variations in individual affective background were expected to be minimal. Regarding the JBT results, there was a significant interaction of Activity and Fearfulness on heifers' *Average latency to reach the ambiguous cues* (p=0.001). In particular, inactive fearful heifers were slower to reach the ambiguous cues (i.e. more pessimistic) than inactive non-fearful heifers (inactive fearful: 73 % ± 8.7, inactive non-fearful: 33 % ± 8.8, p=0.032). Sociability had no significant effect on latency to reach the cues (p=0.150).

Regarding the ABT results, there was a significant interaction between Fearfulness and Sociability on *Time spent eating* (p=0.018). Non-fearful non-social heifers spent more time eating (46 % ± 13.8) than fearful non-social heifers (9 % ± 5.1, p<0.001), fearful social heifers (9 % ± 5.0, p<0.001) and non-fearful social heifers (26 % ± 9.0, p=0.027). The effect of personality on the behaviours observed under the reference conditions during the ABT are detailed in **Table 3.**

Housing effects on heifers' responses to the judgement and attention bias tests.

The longitudinal analyses of cognitive biases *both* in the reference and in the experimental conditions were used to assess the main effect of housing on heifers' cognitive biases while controlling for individual variations - including personality differences. Regarding the JBT results, housing did not influence heifers' *Average latency to reach the ambiguous cues* (p=0.700, reference: 52 % of total trial duration ± 5.0, positive: 54 % ± 6.4, negative 59 % ± 6.9).

Regarding the ABT results, heifers in the negative conditions looked at the threat later and walked less than heifers in the positive conditions (p<0.001 and p=0.011, respectively). Furthermore, heifers in the positive, but not negative, conditions spent less time looking at the threat than they did during the reference conditions (p=0.005).

Table 3. Behavioural measures obtained in the Attention Bias Task under the reference conditions. Results are presented according to the personality traits and classes of personality trait. Results are expressed in proportion of trial duration (120 s), except for *Relative positive attention*, which is expressed in proportion of heifer's total time spent at looking at the stimuli. Significant results are written in bold. [a]NA (not applicable) instead of an exact p-value is indicated in case of interaction effects between personality traits.

Personality		**Latency to look at the threat**		**Latency to eat from the bucket**	
Trait	**Class**	Mean (%) ± s.e.m	p-value	Mean (%) ± s.e.m	p-value
Activity	Active	2 ± 0.6	0.500	57 ± 8.4	0.410
	Inactive	2 ± 0.5		58 ± 9.5	
Fearfulness	Fearful	2 ± 0.4	0.990	64± 8.2	0.260
	Non-fearful	2 ± 0.7		51 ± 9.4	
Sociability	Social	**2 ± 0.4**	**0.043**	63 ± 8.3	0.130
	Non-social	**3 ± 0.7**		51 ± 9.4	

Personality		**Time spent looking at the threat**		**Time spent eating**		**Relative positive attention**	
Trait	**Class**	Mean (%) ± s.e.m.	p-value	Mean (%) ± s.e.m.	p-value	Mean (%) ± s.e.m.	p-value
Activity	Active	13 ± 2.6	0.330	**20 ± 5.8**	**0.013**	53 ± 6.9	0.360
	Inactive	14 ± 3.2		**23 ± 7.4**		52 ± 8.8	
Fearfulness	Fearful	**17 ± 3.3**	**0.039**	9 ± 3.5	NA[a]	**41 ± 6.4**	**0.023**
	Non-fearful	**10 ± 1.9**		34 ± 7.8		**64 ± 8.3**	
Sociability	Social	13 ± 1.9	0.970	19 ± 5.7	NA[a]	52 ± 7.5	0.560
	Non-social	14 ± 3.7		24 ± 7.4		53 ± 8.1	

Personality		**Head up**		**In locomotion**		**In contact with walls/floor**	
Trait	**Class**	Mean (%) ± s.e.m.	p-value	Mean (%) ± s.e.m.	p-value	Mean (%) ± s.e.m.	p-value
Activity	Active	24 ± 4.9	0.670	**27 ± 3.3**	**0.026**	15 ± 2.0	0.052
	Inactive	28 ± 5.3		**17 ± 2.0**		11 ± 2.7	
Fearfulness	Fearful	32 ± 4.8	0.098	**27 ± 3.2**	**0.013**	12 ± 2.0	0.810
	Non-fearful	19 ± 4.9		**18 ± 2.5**		14 ± 2.7	
Sociability	Social	19 ± 3.5	0.089	22 ± 3.0	0.600	**16 ± 2.5**	**0.013**
	Non-social	33 ± 6.0		23 ± 3.1		**10 ± 2.0**	

Housing effects on heifers' behaviours during ABT are presented in **Table 4**. Behavioural responses obtained for each housing conditions are detailed according to personality traits, personality classes and housing conditions in **Supplementary Tables 2-4** online.

Table 4. Average ± s.e.m. of each behavioural response observed during the Attention Bias Tasks according to the housing conditions. Different letters indicate statistical differences between the housing conditions and were extracted from post-hoc testing ($p < 0.05$ after Bonferroni correction).

Response variables	**Reference**	**Positive**	**Negative**
Latency to look at the threat	2 ± 0.4^{a}	1 ± 0.3^{a}	7 ± 2.4^{b}
Latency to eat	54 ± 6.5^{a}	36 ± 8.6^{b}	31 ± 9.1^{b}
Time spent looking at the threat	13 ± 2.2^{a}	6 ± 1.5^{b}	9 ± 3.2^{ab}
Time spent eating	22 ± 4.6^{a}	36 ± 7.2^{b}	49 ± 8.3^{b}
Relative positive attention	55 ± 5.7^{a}	78 ± 5.6^{b}	78 ± 8.1^{b}
Time spent in locomotion	22 ± 2.1^{a}	18 ± 2.3^{a}	13 ± 2.0^{b}
Time spent in contact with walls	13 ± 1.8^{a}	13 ± 3.3^{a}	10 ± 2.3^{a}
Time spent with head-up	25 ± 3.7^{a}	20 ± 5.2^{a}	11 ± 3.3^{a}

Relationships between cognitive bias responses in the reference and in the experimental conditions.

The analyses of covariance allowed for the assessment of consistency in heifers' judgement and attention biases across the different housing conditions, which is used as a measure of unspecific personality influences on heifers' responses to the cognitive bias tests. Regarding the JBT results**,** there was a positive linear relationship between *Average latency to reach the ambiguous cues* in the reference and in the experimental conditions (ß=0.168, p=0.034). There was no evidence of a housing effect on *Average latency to reach the ambiguous cues* in the experimental conditions when controlling for individual response in the reference conditions (p=0.660).

Regarding the ABT results, there was an interaction effect between the covariate *Time spent with head up* in the reference conditions and housing on *Time spent with head up in the experimental conditions* (p=0.007). We found a negative linear relationship between *Time spent with head up* in the reference conditions and in the positive housing conditions, but not between *Time spent with the head up* in the reference conditions and in the negative housing conditions. Results from the ANCOVAs are presented in **Table 5**.

Table 5. Regression coefficients (ß), standard error (in brackets) and p-values of behavioural responses measured during the Attention Bias Tasks in the experimental conditions in relation with their respective measures (covariates) in the reference conditions. (**a**) presents the parameters of the equation lines for both levels of housing, when no significant interaction between the covariate and housing were found. (**b**) presents the parameters of the equation line for each level of housing, when an interaction between the covariate and housing was found. Significant values are in bold.

a

Response variables (experimental conditions)	**Explanatory variables**	
In the Attention Bias Task	Covariate (reference)	Housing (positive – negative)
Latency to look at the threat	ß = -0.063 (0.299), p = 0.820	**p < 0.001**
Latency to eat	ß = 0.317 (0.116), **p = 0.003**	p = 0.950
Time spent looking at the threat	ß = 0.149 (0.202), p = 0.430	p = 0.150
Time spent eating	ß = 0.216 (0.114), **p = 0.043**	p = 0.390
Relative positive attention	ß = 0.109 (0.104), p = 0.260	p = 0.560
Time spent in locomotion	ß = 0.417 (0.173), **p = 0.010**	**p = 0.009**
Time spent in contact with walls/floors	ß = -0.026 (0.170), p = 0.870	p = 0.320

b

Response variables	**Explanatory variables**
In the Attention Bias Task	Covariate (reference)
Time spent with head-up in *positive housing*	ß = -0.396 (0.172), **p = 0.011**
Time spent with head up in *negative housing*	ß = 0.224 (0.195), p = 0.210

Exploratory analyses: Personality and housing interactions on heifers' responses to the judgement and attention bias tests in the experimental conditions

Transverse analyses were used to tentatively explore whether the identified personality traits, *specifically*, may influence individual responses to the cognitive bias tests when heifers were housed under the experimental housing conditions. Regarding the JBT results, there was an interaction between housing (positive versus negative) and Fearfulness (p=0.007) on *Average latency to reach the ambiguous cues*. Non-fearful heifers were faster to reach the ambiguous cues than fearful heifers in the positive conditions only (Non-fearful: 34 % ± 8.9, Fearful: 66 % ± 7.0, p=0.014). There was also an interaction between Activity and Sociability (p=0.004). Inactive social heifers were slower to reach the ambiguous cues than inactive non-social heifers in the experimental conditions (Inactive social: 72 % ± 8.0, Inactive non-social: 52 % ± 8.8, p=0.038). Other relations were not significant (for more detail, see the **Supplementary Table 1**).

Regarding the ABT results, neither housing nor personality significantly influenced heifers' *Latency to eat* or *Time spent eating*. There was no evidence that housing significantly influenced *Relative positive attention* either (positive: 78 % ± 5.6, negative: 78 % ± 8.1, p=0.620), but *Relative positive attention* was lower for fearful heifers than non-fearful heifers (fearful: 68 % ± 7.6, non-fearful: 86 % ± 5.0, p=0.002) and for social heifers than non-social heifers (social: 75 % ± 7.6, non-social: 80 % ± 5.2, p=0.033). There was a significant interaction between housing and Activity on *Latency to look at the threat* (p<0.001). Inactive heifers under the positive conditions looked at the threat sooner than inactive heifers under the negative conditions (positive: 1 % ± 0.4, negative: 13 % ± 5.5, p=0.003). Furthermore, under the negative housing, active heifers looked at the threat sooner than inactive heifers (active: 6 % ± 2.4, p=0.014). There was also a significant interaction of housing and Sociability on *Latency to look at the threat* (p=0.013). Non-social heifers under the positive conditions looked at the threat sooner than non-social heifers under the negative conditions (positive: 2 % ± 0.4, negative: 10 % ± 4.6, p=0.003). Furthermore, under the negative housing, social heifers looked at the threat sooner than non-social heifers (social: 6 % ± 2.7, p=0.012). Similarly, there was a significant interaction between housing and Activity, as well as Fearfulness and Sociability, on *Time spent looking at the threat* (p=0.035 and p=0.011, respectively), but subsequent pairwise comparisons did not reveal any significant differences in responses after Bonferroni correction. Housing and personality effects on *Time spent with head up* also appeared. Heifers in the positive conditions spent more time with the head-up than heifers in the negative conditions (positive: 20 % ± 5.2, negative: 11 % ± 3.3, p=0.005). Regardless of the housing conditions, fearful and social heifers spent more time with the head up than non-fearful and non-social heifers, respectively (fearful: 22 % ± 5.9, non-fearful: 11 % ± 2.6, p=0.003; social: 18 % ± 5.5, non-social: 13 % ± 3.3, p=0.003). Other relations were not significant (for more detail, see the **Supplementary Table 1** online).

Relationships between heifers’ responses to the judgement and attention bias tests

Under the reference conditions, there was no evidence of significant correlations between any of the behavioural responses obtained in the ABT and the average latency to reach the ambiguous cues in the JBT (p-values between 0.186-0.789). Similarly, there was no evidence of significant correlations between judgement and attention bias responses, after correction for a housing effect (p-values between 0.409-0.906).

Discussion

The objective of this study was three-fold. First, we investigated the effects of cattle personality on judgement and attention processes while heifers were kept in similar housing conditions to investigate personality-dependent cognitive biases. Heifers were initially housed under reference conditions in an attempt to standardise their background affective states. Second, we investigated the effects of contrasted housing conditions on cattle responses to the JBT and to the ABT by using a longitudinal approach to control for individual differences. Modifications of the housing conditions were used as a procedure to elicit changes in heifers’ affective valence. Third, we examined whether individual variation in responses to the JBT and to the ABT were consistent across putative affectively-contrasted housing conditions using two complementary approaches (i.e. an untargeted approach using heifers’ response to the JBT/ABT in the reference conditions as a covariate, and a targeted approach focusing on the identified personality traits).

This study supports the idea that cattle personality is multi-dimensional (van Reenen et al. 2005; Graunke et al. 2013; Webb et al. 2015; Koolhaas and van Reenen 2016; Lecorps et al. 2018). We have identified at least three personality traits. RC1 may reflect *Activity/Exploration*. This result is in line with similar studies conducted in cattle (Graunke et al. 2013; Foris et al. 2018; Neave et al. 2020), although one study with a very low number of calves suggested two separate constructs for ‘Activity’ and ‘Exploration’ (Webb et al. 2015). RC2 may reflect *Fearfulness* since behaviours reflecting interactions with the NO were strongly correlated on this axis (van Reenen et al. 2005, 2009; Webb et al. 2015). RC3 may reflect *Sociability* (Lecorps et al. 2018) since it loaded high on latency to reach pen mates and heifers are considered social when they look for conspecifics’ proximity (Gibbons et al. 2010). For ease of reading, the three RCs are hereafter simply referred to as Activity, Fearfulness and Sociability with a capital.

This study investigated the effect of personality on perception of ambiguity within the context of a JBT. In our study, inactive fearful heifers were more pessimistic than inactive non-fearful heifers in the reference conditions. Other studies already noted the influence of Activity and Fearfulness on animal judgement biases. Pigs classified as active personality-wise, for instance, were consistently less pessimistic than inactive ones regardless of their housing conditions (Asher et al. 2016), and fearful calves consistently showed more pessimistic responses over time than non-fearful individuals (Lecorps et al. 2018). These results may be due to affective differences between animals' perception of the task and its settings, depending on their personality. Because fearful individuals are prone to neophobia (Boissy and Bouissou 1995), they may have perceived the exposure to the ambiguous – and intrinsically novel – cues more negatively than non-fearful individuals. Alternatively, the set-up of the JBT itself may have triggered personality variations in pessimism, in an affect-independent manner. Regardless of their affective states, active individuals may be more likely to engage in any kind of locomotor response, e.g. reach the ambiguous cues, than inactive ones. This hypothesis is in agreement with previous research showing that personality, in particular coping-style, predicts decision style – which reflects individual predispositions for decisions involving risk/reward trade-offs (Mazza et al. 2019). Moreover, our study demonstrates that personality profile predicts animal responses to the ambiguous cues better than a unique personality trait does. In our conditions, predispositions to Inactivity and to Fearfulness exerted a synergistic effect on heifers' likelihood to reach the ambiguous cues. To better understand the role of individual differences in judgement processes, we therefore encourage researchers to characterise animal individuality based on personality profile rather than a single personality trait.

This study also investigated the influence of personality on the perception of threat in the context of a newly developed ABT. First, Activity did not significantly influence cattle threat-directed nor food-directed behaviours, therefore suggesting that Activity does not alter heifers' affective perception of threat. This theory is consistent with the idea that Activity is independent from an emotionality dimension (Koolhaas et al. 2007; Koolhaas and van Reenen 2016) and in line with Luo and colleagues (2019) who found no significant effect of coping-style on *Time spent looking at the threat* or *Time spent eating* in pigs after the threat exposure – although they reported that proactive pigs looked at the threat more frequently than reactive pigs in enriched conditions. As expected, Activity influenced heifers' locomotor behaviours – active heifers walked more than inactive ones during the ABT. This finding further supports the validity of Activity as a personality trait in cattle, because heifers displayed consistent locomotor behaviours across contexts. Unlike Activity, both Fearfulness and Sociability influenced attention bias in the reference conditions. In humans and farm animals, certain underlying traits, like trait anxiety, have also been associated with sustained attention to threat (Veerapa et al. 2020, Lee et al. 2016, 2017) and with longer latencies to engage with positive stimuli (Campbell et al. 2019). In our conditions, fearful heifers were more biased towards the threat than the bucket (i.e. the positive cue) compared with non-fearful heifers. We suggest that fearful heifers may have experienced the exposure to the dog model more negatively than non-fearful heifers. Considering that heifers in our study had no previous experience with the dog model during the first ABT, we speculate that fearful heifers were more scared of the dog model or more anxious about the threat after the dog model was covered. Moreover, fearful heifers walked more during the ABT compared with non-fearful heifers. This finding is in accordance with studies demonstrating that drug-induced anxiety increases locomotion in hens (Campbell et al. 2019) and beef cattle (Lee et al. 2017) during ABTs. Our results must, nonetheless, be interpreted with caution, since we did not validate beforehand that our dog model was truly perceived as threatening for heifers. Our ABT was adapted from Lee and colleagues (2017), who validated their task as a reliable tool to assess beef cattle anxiety. However, Lee and colleagues (2017) were authorised to use a live dog, a procedure against the safety hazard policy of our experimental farm. We also found that social heifers looked faster at the threat than non-social heifers, which may suggest that social heifers were in worse affective states than non-social heifers. This presupposed difference in affective states could be explained by the fact that social heifers may have suffered more than non-social heifers from being separated from their companions during the ABT. In addition, we found that social heifers spent more time in contact with the floor/walls of the arena than non-social heifers. We speculate that this behaviour may reflect social heifers' heightened motivation to find an exit from the arena in order to reunite with their pen mates, since heifers classified as social in our study were – by definition – more willing to stay in proximity to their conspecifics than non-social heifers. From an evolutionary perspective, and in line with this idea, we hypothesise that social heifers may be more susceptible than non-social heifers to anti-predation grouping, an adaptive strategy used by ungulates to dilute predator risks (Creel et al. 2014). This presumed susceptibility to grouping might, therefore, have mediated social heifers' motivation to escape the arena in response to our predator-like dog model. Finally, similar to JBT findings, we found an interaction effect of personality traits on responses to the ABT, with non-fearful and non-social heifers spending more time eating than heifers of other personality profiles. This finding, once more, highlights the need to characterise individuality among heifers using personality profile rather than personality trait.

Overall, this study confirms that stable traits in cattle are associated with differences in behaviours during the ABT, as demonstrated by Lee and colleagues (2017), who showed that beef cattle temperament index (measured from flight speed and crush score) was positively associated with the number of zones crossed and the attention toward the threat. The exact nature of the personality traits underlying this temperament index remains, however, unclear. Lee and colleagues (2017) suggested that their index reflected individual general agitation, a theory partially supported by our findings showing that Activity influences cattle locomotor behaviour in the ABT. However, since differences in Activity do not explain variation in threat-related behaviours, we suggest that other personality traits, like Sociability, may underlie this temperament index.

Altogether, these results further demonstrate that JBT and ABT paradigms are not purely state-sensitive but are also trait-sensitive. These results may reflect variations in heifers' background affective states, which could be either due to personality-based differences in individual perception of the reference conditions or due to a failure to standardize background affective states within 9 weeks. Alternatively, our findings may reflect personality-based differences in heifers' perception of the tasks' set-up itself (e.g. with regard to the type of response, type of cue, level of isolation, etc.). This idea questions the generalizability of our cognitive bias tasks across individual of various personality and highlights the need to control for individual variations when assessing cognitive biases in our study.

This study also investigated the sole effect of housing on judgement processes, while controlling for individual differences. Surprisingly, changes in housing did not impact heifers' pessimism. One explanation could be that housing did not elicit the expected shifts in heifers' affective states. Background affective states are thought to result from the accumulation of positive and negative experiences (Mendl et al. 2010b), but our housing changes may have been too infrequent (i.e. once a week for an entire week) and predictable (i.e. every Friday afternoon) to truly impact heifers' opportunity to experience positive or negative events. Alternatively, background affective states may not emerge from a general accumulation of positive and negative experiences, as initially hypothesised, but they may arise – more specifically – from an accumulation of mismatches (Eldar et al. 2016; Raoult et al. 2017). More research is required to understand the aetiology of background affective states (Raoult et al. 2017). Another explanation could be that heifers were affected by the changes, but in the short-term only. In our conditions, heifers might have had the ability to habituate to the housing changes within days, while our experiment was designed to detect the long-lasting consequences of housing changes by exposing heifers to the JBT a week after the last housing modifications. Cows may, therefore, be more resilient to subsequent changes than initially anticipated. Another explanation is that the JBT itself failed to detect the affective difference between heifers housed in contrasted housing conditions. This lack of treatment detection may be due to an impaired sensitivity of our own JBT set-up. We used a Go/NoGo task based on a spatial discrimination among a female population – while a recent systematic review revealed that JBTs yield larger treatment-induced judgement biases when using Go/Go tasks based on auditory or tactile cues in males (Lagisz et al. 2020). Therefore, we cannot exclude the possibility that our housing conditions effectively influenced heifers' affective states but that our JBT set-up was not sensitive enough to detect shifts in judgement bias. Our results are in line with Crump and colleagues (Crump et al. 2021) who also failed to detect a shift in cows' pessimism given access to pasture, while using a similar Go/NoGo spatial JBT. Therefore, we encourage researchers to develop alternatives of our JBT set-up (e.g. auditory Go/Go tasks as previously suggested: Kremer et al. 2021b) that would be more sensitive to affective shifts when investigating the effects of common farm practices on cow affective states. Once more, this study highlights the necessity for researchers to combine indicators of various nature (cognitive, behavioural and physiological) to assess animal affective states reliably. During our experiment, samples from various biofluids were collected on a weekly basis during both reference and experimental conditions. Samples were also collected while the heifers were exposed to an acute-stress test at the end of both reference and experimental conditions. Results from these physiological markers (in prep.) will allow us to draw more solid conclusions with regard to potential treatment-induced affective states in heifers. In particular, heifers' physiological responses to the acute-stress tests will help us identify whether housing effectively influenced individuals' ability to cope with stressors, since long-term negative affective states are often associated with physiological dysregulations (McEwen 2004).

As for JBT, this study investigated the sole effects of housing on behavioural responses to the ABT, while controlling for individual differences. Contrary to expectations, heifers in the negative conditions looked at the threat later and walked less than heifers in the positive and in the reference conditions. Furthermore, although non-significant, heifers in the negative conditions spent on average less time with the head up than heifers in the reference conditions. Although unexpected, these results are in line with another study conducted in sheep where chronically stressed individuals exhibited reduced vigilance towards a live predator threat (Verbeek et al. 2019). As hypothesised by the authors of the aforementioned study, these findings are consistent with a phenomenon known as attentional avoidance, where attention is allocated away from the threat location. Similarly, Bethell and colleagues (2012) reported that rhesus macaques avoided looking at threatening faces of conspecifics following an acute-stress procedure. Interestingly, attentional avoidance effects – as opposed to facilitated attention to threat – have repeatedly been reported in anxious humans when threatening stimuli are presented for long (superior to 20 s) but not short durations (Cisler and Koster 2010). Therefore, considering the duration of our trials, our ABT was potentially more likely to detect anxiety-driven differences in attentional avoidance strategies rather than differences in threat detection. Overall, our results could, therefore, suggest that heifers in the negative housing conditions became chronically stressed. This theory, yet, remains to be verified using validated indicators of chronic stress such as heart rate variability indices (Kovács et al. 2015). Alternatively, we cannot rule out the idea that these results could also indicate that heifers in the negative conditions learnt to cope better with challenges due to their repeated exposure to stressors during the experimental periods, or that heifers housed in the negative conditions were in relatively better affective states than heifers housed in the positive conditions. Heifers in the negative conditions may have been temporarily relieved to exit their home pens and became momentary less scared/anxious during the ABT. This assumption is strengthened by the fact that heifers in the negative conditions spent on average less *Time with the head-up,* which is a measure of vigilance (Welp et al. 2004), than heifers in the reference and positive conditions – although statistical differences between housing conditions were not significant. Nonetheless, this idea is speculative and remains to be further validated – by comparing, for instance, heifers' home pen behaviours in the different housing conditions. Besides, the use of attention bias – unlike judgement bias – as a valid indicator of positive affective shift remains to be proven (Monk et al. 2020). Finally, heifers in both negative and positive housing conditions spent more time eating and shifted their attention more towards the positive cue during the second ABT compared with the first. Overall, these results may indicate that heifers habituated to the task and became either less scared of the dog model or remembered that the familiar bucket was also filled with concentrates during this task. This theory, however, contrasts with a previous study conducted in rhesus macaques, where a week interval between testing seemed sufficient to supress the effect of repeated testing on animal responses to ABT (Howarth et al. 2021).

Overall, there is little evidence that our housing conditions influenced dairy heifers' cognitive biases when controlling for inter-individual variation. While this finding may suggest that our housing conditions did not substantially influence heifers' affective states, this lack of statistical support may also reflect methodological limitations of our study. In particular, we question the repeatability of our cognitive bias tasks: we cannot exclude the idea that, in the reference conditions, heifers' responses to the tests influenced individual affective experience of the tests themselves – which might have, in turn, influenced heifers' responses to the cognitive bias tests in the experimental conditions. More research is, therefore, required, to assess whether animal prior experience of the cognitive bias tests influences individual responses to subsequent tests.

This study also aimed to investigate whether variations in responses to JBT were mediated by underlying traits. Assuming that influences of individual traits on judgement bias are constant over time and across contexts, we explored the predictive value of heifers' pessimism in the reference conditions on subsequent pessimism when heifers were under supposedly affectively-contrasted housing conditions. Of note, in this study, we restricted ourselves to the sole investigation of linear relationships between pessimism in the reference conditions and pessimism assessed in the experimental conditions. The regression ANCOVA analysis revealed that pessimistic heifers in the reference conditions remained pessimistic in the experimental conditions. This finding is in line with that of Lecorps and colleagues (2018) who found that pessimism is constant in calves. This result supports the idea that cattle decision-making under ambiguity is influenced by stable individual characteristics, i.e. personality traits independent from environmental context. This hypothesis is furthermore supported by the fact that inactive social heifers responded in a more pessimistic manner than inactive non-social heifers, irrespective of the experimental housing conditions. Such findings may, once more, reflect personality-based differences in heifers' affective background or personality-based differences in heifers' perception of the JBT set-up. A word of caution with regard to these results is, however, due here. Considering the relatively small sample size of our population study, we cannot rule out the possibility that our model failed to detect a significant interaction effect between pessimism in the reference conditions and experimental housing (p=0.082). In agreement with this idea, pessimism in the reference conditions appeared to more reliably predict pessimism in the negative conditions than in the positive housing conditions. This could suggest that personality-based differences in response to JBT may be exacerbated during challenging conditions, as found in humans and non-human animals (Goldsmith et al. 1987; van Reenen et al. 2004). Considering that pessimism may be both affective state- and trait-dependent (Faustino et al. 2015), the relative lack of consistency between pessimism in the reference conditions and pessimism in the positive housing conditions may indicate a certain variability in the affective states experienced by heifers in the positive conditions compared to heifers in the reference conditions. This theory is strengthened by the interaction effect between Fearfulness and housing on heifers' pessimism – non-fearful heifers being less pessimistic than fearful heifers in the positive, but not in the negative conditions. We suggest that fearful heifers may have experienced the repeated positive changes in their environment less pleasantly than fearful heifers, which resulted in greater affective differences among the two sub-populations. Fearful heifers may have suffered from the weekly changes occurring in their home pen, particularly from the repeated introduction of new enrichment. Similarly, frequent rotation of enrichment objects in parrots was shown to successfully reduce individual fear behaviours – except for the most fearful parrots who displayed even more fear behaviours (Fox and Millam 2007). Therefore, our findings corroborate the idea that individual differences must be carefully considered when designing animals' enclosure to improve their welfare.

Finally, we investigated whether variations in responses to ABT were mediated by underlying traits. Interestingly, positive linear relationships were found between feeding-directed behaviours in the reference and in the experimental conditions. This result suggests the existence of one (or more) underlying stable traits accounting for inter-individual variations in feed-directed behaviours – a result consistent with Melin and colleagues (2005), who found that individual differences in dairy cattle explain 84 % to 98 % of the variation in feeding patterns. Furthermore, our exploratory analyses seem to indicate that Fearfulness may mediate cow feeding-directed behaviours, as suggested elsewhere (Neave et al. 2018). More research is, nonetheless, warranted to validate this preliminary finding, since we used a relatively low number of individuals per personality trait and housing conditions in our study. In contrast, there were no linear relationships between threat-directed behaviours observed during the reference and during the experimental conditions. We could hypothesise that threat-directed behaviours are relatively insensitive to trait-differences in cattle, but this theory seems unlikely considering recent findings demonstrating a high degree of repeatability (R=0.63) in attention to threat over several years in rhesus macaques (Howarth et al. 2021), and the evidence of stable differences in attention biases among humans according to their trait-anxiety scores (Veerapa et al. 2020). Alternatively, the lack of consistency in threat-directed behaviours may, once more, suggest that heifers experienced the housing conditions differently depending on their personality. This theory is partially supported by the fact that inactive heifers looked at the threat later than active heifers in the negative conditions. Likewise, non-social heifers looked at the threat later than social heifers in the negative conditions. There are several potential explanations to these preliminary findings. Inactive and non-social heifers may have been more relieved to exit their home pens, and therefore temporarily in better affective states during ABT, than active and social heifers. Conversely, and in congruence with the attentional avoidance theory, inactive and non-social heifers may have been in worse affective states than active and social heifers during ABT. Although mutually exclusive, both theories seem to indicate that inactive and non-social heifers may have experienced the negative conditions more aversively than active and social heifers. However, we cannot rule out the possibility that these personality-based differences in response to the ABT may also be independent from any affective processes. For instance, active heifers may have been less flexible in their responses to the ABT than inactive heifers, as a result of personality-dependant behavioural inflexibility (Koolhaas et al. 2010). In line with this idea, proactive pigs were shown to be more optimistic in the JBT independent of their housing conditions (Asher et al. 2016). Considering that behavioural flexibility depends on individual personality, we therefore question the generalizability of both cognitive bias tasks across individuals of different personality – since both tasks solely rely on behavioural outcomes. Lastly, no clear relationships between heifers' responses to ABT and JBT were identified, which is in line with previous finding obtained in sheep (Monk et al. 2019a). This result could potentially indicate that separate mechanisms underlie the aetiology of attention and judgement biases.

In conclusion, we did not find substantial evidence that housing conditions influenced heifers' affective states since housing had relatively little effect on heifers' cognitive biases. Nonetheless, when housing effects on cognitive biases were identified, they appeared to be mediated by heifers' personality. On the one hand, this result could indicate that heifers' affective experience of their housing conditions differs according to individual personality. On the other hand, this finding questions the validity of both cognitive bias tasks as repeatable tools for affective states assessments, since personality-based differences in response to the JBT and to the ABT may also be affect-unrelated.

Materials and methods

Animals and husbandry system

The experiment took place between February 2019 and January 2020. The study was divided into three batches of fifteen weeks each. Each batch was composed of four groups of eleven Friesian Holstein dairy cows. Among the eleven individuals, four animals were focal individuals, while seven animals were companion animals. Focal animals (N=48) were first parity heifers between the third and seventh lactation month when the batch started. The term 'focal group' is used here to refer to a subset of four heifers housed in the same pen. Heifers were pseudo-randomly allocated to their group based on their days in milk (165 d ± 5.5), milk production (25.16 kg ± 0.609) and body weight (606 kg ± 6.0). Companion animals were cows between the second and sixth parity. They were pseudo-randomly allocated to their group based on their parity (3 ± 0.1), milk production (30.39 kg ± 0.663) and body weight (707 kg ± 7.4). All heifers and cows were healthy at the beginning of the experiment (i.e. somatic cell count within normal-range, absence of fever and absence of lameness) and confirmed pregnant. However, one companion animal was removed on week fourteen from the first batch because she was found contagious for para-tuberculosis. Furthermore, one heifer was replaced in the second week of the third batch due to miscarriage.

All groups were housed in the same barn, but visual and tactile contacts between the groups were prevented via 2 m-high solid partitions. Milking occurred twice a day at around 05:00 h and 15:00 h. Cows received a total mixed ration of maize silage (35 % of dry matter), grass silage (30 %), concentrates (20 %), grinded whole soy (10 %), grinded whole wheat (3 %) and minerals (2 %) at around 7:00 h, which was pushed up again at around 17:00 h. Within a pen, cows had free access to one automatic concentrate dispenser delivering a pre-set daily amount of concentrates based on individual milk production, and they had *ad libitum* access to one water trough. For each batch, the study was divided into two phases, hereafter referred to as 'reference conditions' and 'experimental conditions'. At the beginning of our study, heifers were housed in a stable environment in the reference conditions, while they were subsequently housed either under environmental conditions that were assumed to progressively (i.e. every week) worsen, or under environmental conditions that were assumed to progressively improve in the experimental conditions. During both the reference and experimental conditions, heifers were exposed to a series of behavioural tests, including the JBT and the ABT testing (**Figure 1**). Heifers' responses to cognitive bias tests in the reference conditions were used as baseline, while their responses to the tests in the experimental conditions were used as measures of housing-induced affective states.

All experimental procedures were approved by the Central Committee on Animal Experiments (the Hague, the Netherlands), Approval Number AVD4010020174306. All methods involving animals during the study were carried out in accordance with the 'Wet op de dierproeven' (law on animal experiments) and ARRIVE guidelines. Methods requiring plant materials were also carried out in accordance with the institutional guidelines and regulations.

Housing conditions

Reference conditions

For nine weeks, heifers were housed under reference conditions (**Figure 2a**). In each pen, cows had access to eleven flexible cubicles with gel mattresses (AgriProm) covered with sawdust, eleven feeding gates, and one simple, fixed brush. Mixing within the groups was prohibited.

Experimental conditions

Following the reference conditions, two pens per batch (i.e. six in total) were allocated to weekly-improved housing conditions (hereafter 'positive housing'), while the other two pens were subjected to weekly-worsened housing conditions (hereafter 'negative housing'). For a detailed description of the treatments, see **Table 6**. These weekly changes aimed at inducing positive/negative shifts in heifers' background affective states, under the assumption that the latter emerge from the accumulation of respectively positive or negative events (Mendl et al. 2010b). To induce a relative positive affective shift, we manipulated the housing conditions based on existing literature in cattle that links specific housing elements with 1) preferential/motivational behaviours, 2) an increase in comfort indicators, or 3) a decrease in behavioural/physiological indicators of stress. Three actions were performed to design the positive housing conditions: under-crowding conditions (extra space: Schütz et al. 2015, under-crowding: Wang et al. 2016), environmental enrichment (provision of brushes: McConnachie et al. 2018, installation of feed partition: Huzzey et al. 2006) and social stability (familiarity among cows: Gutmann et al. 2015).

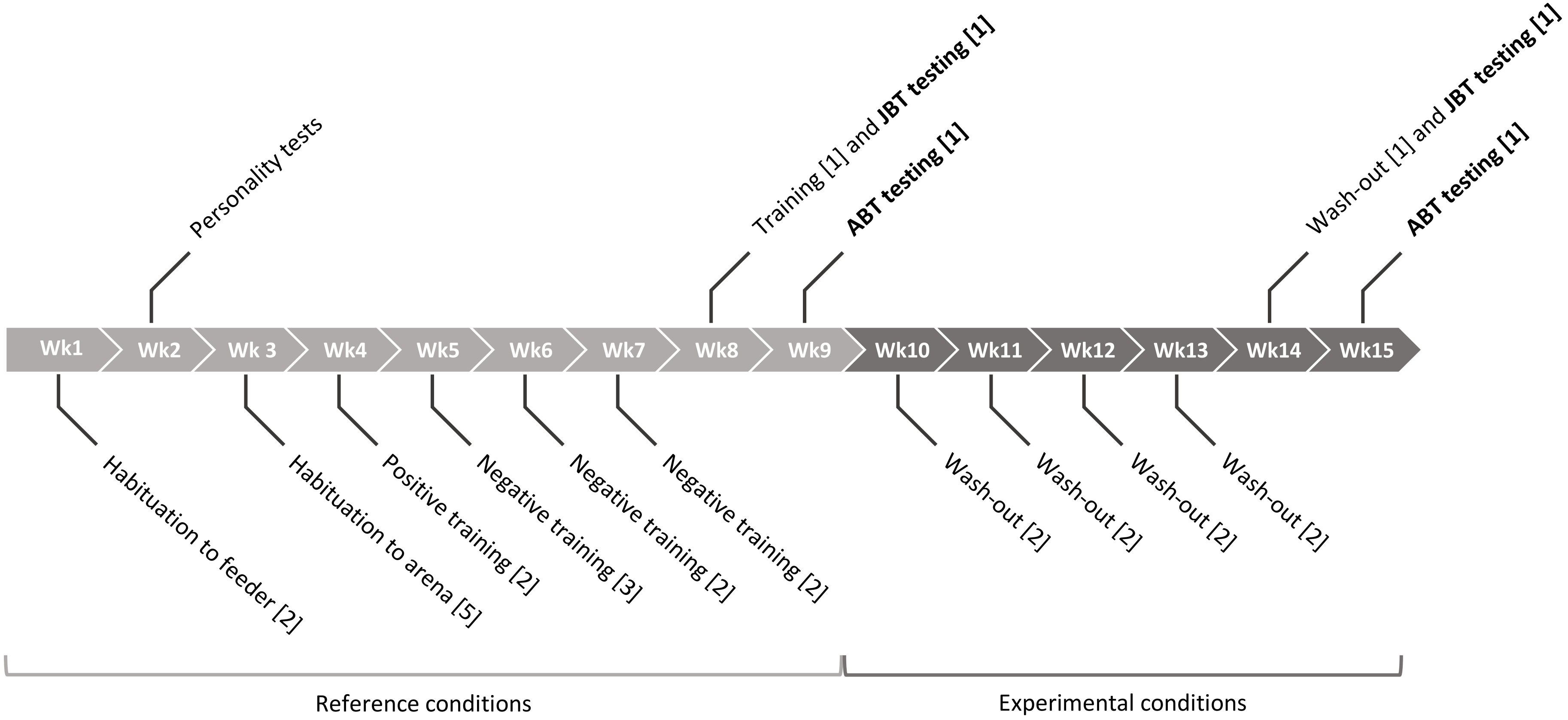


Figure 1. Timeline of the experimental procedures during each batch. Habituation, training and wash-out sessions were part of the Judgement Bias Task (JBT). Numbers in square brackets indicate the number of sessions conducted per heifer per week on weekdays. If necessary, additional sessions were added during the weekends or during the evenings. Testing sessions of the JBT and the Attention Bias Task (ABT) are indicated in bold. Weeks in light grey (wk1-wk9) depict the reference conditions, while weeks in dark grey depict the experimental conditions. Heifers were housed under stable housing conditions during the reference conditions; while they were housed under supposedly weekly-improved or weekly-worsened conditions during the experimental conditions.

To induce a relative negative affective shift, we manipulated the housing conditions according to the existing literature demonstrating a link between specific housing elements and an increase in cattle physiological or behavioural stress markers. Three actions were performed to design the negative housing conditions: over-crowding (Fregonesi et al. 2007; Winckler et al. 2015), barren housing conditions and social instability (Schirmann et al. 2011; Wilcox et al. 2013). As a result, the negative housing conditions did not meet the European minimum recommendations – especially in terms of stocking density (Standing Committee of Council of Europe 1988). In both negative and positive conditions, the weekly housing changes was always performed on Friday afternoons for a period of 1 week. Each experimental week started from Saturdays, i.e. once the weekly treatment had been applied. Successive positive/negative contrasts were thus created, and the risk of heifers habituating to the treatment was therefore minimised. **Figure 2b**. provides an overview of the housing conditions at the end of the experimental conditions, for both treatments.

Table 6. Detailed treatment applied every Friday during the experimental conditions to design the positive and negative housing conditions.

Week	Levers of actions	Positive housing	Negative housing
Wk10	Crowding conditions	Increase space allowance Open 1 cubicle and 1 feeding gate	Decrease space allowance Close 2 cubicles and 2 feeding gates
	Social stability	Add feeding partitions	
Wk11	Enrichment	Add a fixed brush or replace a fixed brush by a rotating one	Remove the fixed brush
	Social stability	Keep stable groups	Mix two companion animals
Wk12	Crowding conditions	Open 1 cubicle and 1 feeding gate	Close 2 cubicles and 2 feeding gates
Wk13	Enrichment	Add a fixed brush or replace a fixed brush by a rotating one	Switch to another home pen
	Social stability	Keep stable groups	Mix two companion animals
Wk14	Crowding conditions	Open 1 cubicle and 1 feeding gate	Close 1 cubicle and 1 feeding gate
	Social stability	Keep stable groups	Mix two companion animals

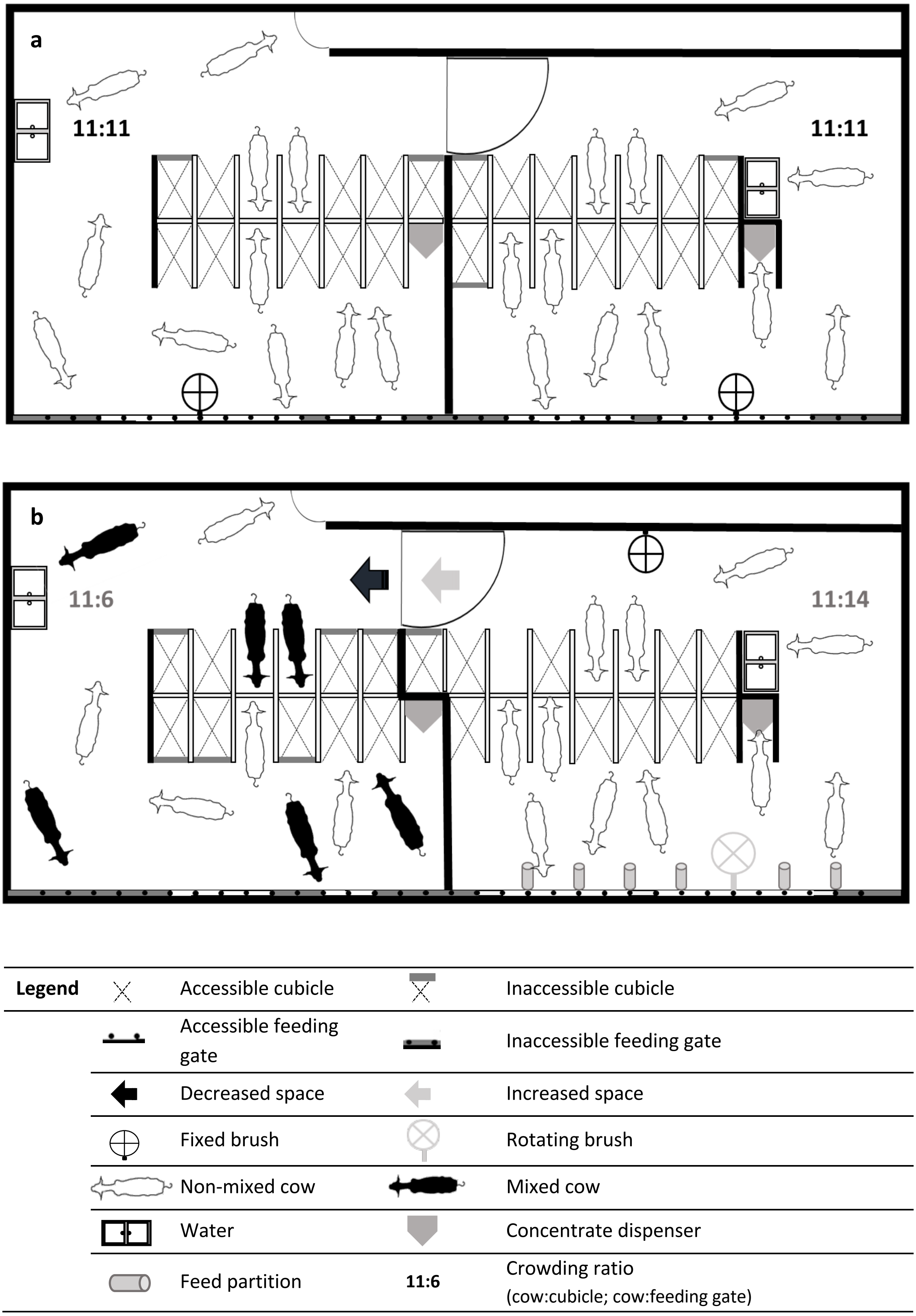


Figure 2. Schematic layout of two adjacent pens. (**a**) represents a schematic layout of two pens in the reference conditions. (**b**) represents a schematic layout of two pens at the end of the experimental negative housing conditions (left pen) and at the end of the positive housing conditions (right pen).

Personality tests

On week two, heifers were subjected to three standard personality tests: the Open-Field (OF), Novel-Object (NO) and Runway (RW) tests, in this order. The OF and NO were video recorded (CAMCOLBUL2, Velleman, Belgium), while the RW was live scored. Behaviours were scored using The Observer XT 10 (Noldus Information Technology BV, Wageningen, the Netherlands). For each personality test, the testing order of the experimental groups was pseudo-randomly determined based on pen allocation. Each heifer was consecutively subjected to the OF and NO on the same day. Two days were needed to test all sixteen heifers to the OF and the NO tests. The OF and NO protocols were adapted from those developed for calves by van Reenen and colleagues (2004). All heifers were subjected to the RW test on the same day. The RW protocol was based on Gibbons and colleagues (2010). For each test, two experimenters were in charge of handling the cows and scoring the heifers' behaviours. All behavioural measures are detailed in **Table 7**.

Table 7. Definitions of the behavioural measures recorded or live scored across the three personality tests.

Variable	Definition
Open-Field and Novel-Object tests	
In locomotion (% of time)	Movement of front legs or all four legs once one of the two front hooves is off the floor (adapted from van Reenen et al. 2004). The locomotion bout stops when both front hooves touch the floor for more than 1s
In contact with floor and walls (% of time)	Muzzle below heifer's carpal joint, or head oriented towards the wall with the muzzle in proximity/in contact with the wall
Novel Object test	
Latency to touch the object (s)	Time until the first contact with the object (van Reenen et al. 2004)
In contact with the object (% of time)	Touching the object with the muzzle, the head or the shoulder
Runway test	
Time spent in the 2 m zone (s)	Time spent with both front hooves within 2 m from the gate separating the runway and the waiting area

Open-Field test

Groups of three individuals from the same pen were brought into the waiting area. Each group consisted of two heifers and a third parity companion cow. The companion cow was included in the group to prevent each heifer from being isolated in the waiting area while the other heifer was being tested. Heifers were then individually brought into a 7 m x 7 m testing arena, unfamiliar to the animals. Before entering the arena, the focal heifer was positioned inside a 2 m x 1 m starting box, where she remained for 3 min. The door of the testing arena was then manually opened, and the experimenter tapped three times on the heifer's hips to make her enter. This procedure was applied to ensure that all heifers entered the arena in a standardised manner. The test started once the heifer crossed the virtual line of the entrance door with two front hooves and lasted ten minutes.

Novel-Object test

Immediately following the OF test, a novel object attached to a rope was quickly lowered in the middle of the arena until it touched the floor. The novel object was then lifted up at 1 m above the floor for 10 min, i.e. for the entire test duration. The novel object was new to the heifers and consisted of two orange cones filled with stones (for weight) and attached together.

Runway test

A runway test was conducted in the corridor leading to the milking parlour (**Figure 3**). From each pen, six cows were brought into a waiting area – the focal group of four heifers and two companion cows of second and third parity. The cows remained in the waiting area for 10 min before the test. Each focal heifer was then tested individually in a random order. The focal heifer was brought by an experimenter into the starting area located 18 m away from the group. A removable bar prevented the heifer to reach the group for 1 min, before being gently removed by the experimenter. The test lasted 5 min and started once the heifer voluntarily crossed the starting line with both front hooves. If the heifer did not cross the starting line within 5 min, the experimenter would encourage the heifer to walk by doing circular forearm movements in the air, without physical contact. At the start of the test, the experimenter slowly withdrew from the runway. During the RW, behaviours were live scored by using a portable computer equipped with The Observer XT 10.

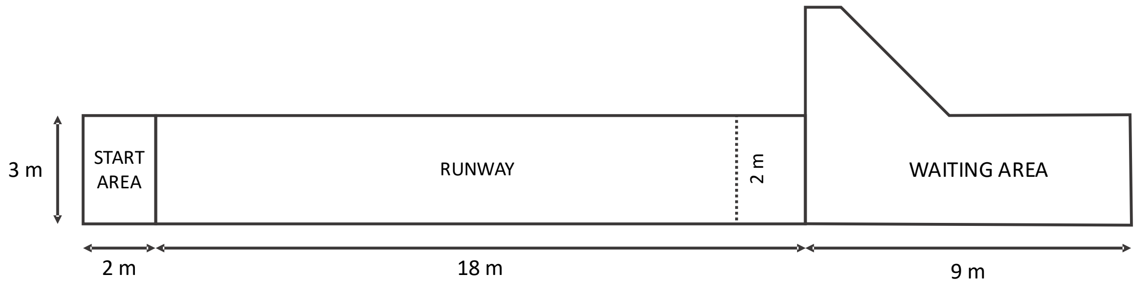


Figure 3. Layout of the runway. During the test, five cows were herded in the waiting area, and the focal heifer was brought by one experimenter to the start area. After 1 min, the heifer was released onto the runway for 5 min. A second experimenter scored the heifer's time spent within 2 m (dashed line) from the gate separating the runway and the waiting area.

Judgement Bias Task

Judgement bias apparatus

JBT was carried out in the same arena as the one used for the OF test. It was adapted from previous studies conducted in ruminants (Kremer et al. 2021b; Destrez et al. 2012; Lecorps et al. 2018). It consisted of a Go/No-Go task, based on a spatial cue characterised by an automatic feeder always filled with 150 g of concentrates that could be remotely released. For a detailed description of the facilities used for the JBT, the reader is referred to Kremer et al. (2021b).

Judgement bias habituation

The habituation took place in 6 steps. First, a focal group (i.e. four heifers from the same pen) was brought to the waiting area, and pairs of heifers were then habituated to eat from the feeder in a subsection of the waiting area, hereafter called the turning area. Concentrates were released when one heifer was about 50 cm away from the feeder. Once one heifer had eaten from the feeder, she was brought to the exit corridor, to limit competition for the feeder and to allow the second heifer to eat from the feeder. Once the second heifer also ate, the next pair of heifers was brought to the turning area – and the same procedure was applied. Two days later, heifers were once again brought in focal groups of four to the waiting area, but they were individually introduced to the turning area. The focal heifer was released into the exit corridor once she ate three consecutive times from the feeder without startling in response to the concentrate release. Once habituated to the feeder, heifers were familiarised to the testing arena itself two weeks later. Heifers from the same pen were first brought in groups of four inside the testing arena, where three buckets filled with concentrates and one feeder were present in each corner of the arena. Concentrates were remotely released from the feeder when one heifer was 50 cm away from the feeder. The corner attribution was randomly selected. The door of the arena remained opened, and the habituation trial stopped once all heifers exited the arena by themselves or after approximatively 5 min. The experimenter then re-filled all buckets and the feeder; and the trial was repeated four to five times in total. On the following day, heifers were brought in pairs inside the testing arena for two consecutive trials of 10 min each. Three buckets and one feeder were located in each corner of the arena. The next day, heifers were introduced alone to the arena for 2 consecutive trials of 5 min each. Heifers still had access to one bucket and one feeder filled with 150 g of concentrates – the positions of which were once again pre-randomly selected. Eventually, the feeder was solely positioned in one of the two far corners of the arena (hereafter called ‘positive’ location and abbreviated ‘P’), and the heifer was considered as habituated once she reached the feeder within 3 min for two consecutive habituation trials. Extra habituation sessions were provided until all heifers reached the habituation criterion.

Judgement bias positive training

Once habituated, each heifer was subjected to at least two positive training sessions. Each session consisted of three trials of 90 s where the feeder remained in P. The corner attribution was balanced across groups and pens - and remained the same within a pair of heifers for practicality. Before each trial, the heifer was kept for 30 s inside the starting box adjacent to the testing arena. The entrance door was subsequently opened, and the heifer was tapped three consecutive times on her hips to encourage her to enter the arena. If the heifer did not enter, the taps were repeated, and the heifer was eventually physically encouraged if needed. When the heifer’s muzzle reached the 50 cm-radius circle around the feeder, 150 g of concentrates were delivered. If the heifer did not reach the feeder within 90 s, the trial was extended for an additional 30 s. If the heifer still did not reach the feeder, one experimenter entered the arena and gently orientated the heifer towards the feeder while talking to her and petting her on the hips – until she reached the feeder and ate from it. Heifers were considered trained once they reached the feeder within 30 s for three consecutive times. Additional positive training sessions were provided where necessary.

Judgement bias negative training

Following the positive training, heifers were trained to discriminate between two feeder’s locations, either on P or in the opposite corner for at least eight sessions. The opposite corner will hereafter be referred to as ‘N’, which stands for negative location. Heifers were trained to display Go-responses to the feeder to obtain 150 g concentrates when the feeder was located on P. The response was considered correct and deemed a Go when the heifer reached the feeder within 20 s. If she did not reach the feeder within trial duration (90 s), the same procedure as the one used for the positive training session was applied. Alternatively, heifers were trained to display NoGo responses in order to avoid a 6 bar air puff when the feeder was located on N. This combination of reward/punisher was selected to maximise the sensitivity of our JBT to shifts in heifers’ affective states (Kremer et al. 2021b). The response to N was considered correct and deemed a NoGo if the heifer did not reach the feeder during the whole trial duration, i.e. 90 s. If the heifer reached the feeder during the negative trial, an air puff was released from the bottom of the feeder’s bowl, and the trial ended 10 s later. Each training session was composed of 10 consecutive trials: 6 positive trials, and 4 negative ones. The order of trials was pseudo-randomly determined: the negative training session always started with a positive and a negative trial, and always ended with a positive trial. This was done to ensure the heifer’s motivation to participate in the task. From this phase of JBT, heifers remained 20 s inside the starting box before each trial. Heifer were considered trained if they displayed 10 correct responses during one training session.

Judgement bias testing

On weeks 8 and 14 of the experiment, all heifers were subjected to the judgement bias testing. The testing session was composed of 10 consecutive trials, among which 4 positive trials, 3 negative trials and 3 ambiguous trials. The order of trials was pre-determined, and the ambiguous trials were interspersed by one positive trial and one negative trial in this order. Furthermore, the session always started with a positive and a negative trial, and ended with a positive one. All heifers were first exposed to a truly ambiguous cue (A), positioned between P and N. On the sixth and ninth trials, heifers were either exposed to a positive ambiguous cue (Ap) positioned in between A and P; or to a negative ambiguous cue (An) positioned in-between A and N. The order of Ap and An trials was balanced across pairs, groups and treatment. Latencies to reach the cues were video recorded. Animals were exposed to the same sequence of trials on weeks 8 and 14, and were tested exactly in the same order in the two sessions.

Judgement bias wash-out

During the experimental conditions, training sessions were maintained. This was made to minimise the risk of heifers remembering their exposure to the ambiguous cues; and to maintain heifers' routine since JBT training may provide cognitive enrichment (Roelofs et al. 2016). The same procedures as those used during the training sessions were followed. In total, heifers were exposed to nine wash-out sessions.

Attention Bias Task

Attention bias apparatus

The ABT was adapted from previous studies conducted in ruminants (Lee et al. 2016, 2017). The arena consisted of a subsection of the milking parlour's waiting area delimited by a 1.5 to 2.0 m high tarp (**Figure 4**). As previously recommended (Monk et al. 2018a), the positive stimulus and the threat were positioned in a such a way that the heifers could not look at both simultaneously. The positive stimulus consisted of a familiar bucket filled with 500 g of concentrates and was located in the right corner of the arena. The threat consisted of a dog model positioned on the left of the arena, behind a hole (1.0 m x 1.3 m) in the tarp. The dog model was built from a combination of visual, olfactory and auditory cues. In batch 1 and 2, the visual cue consisted of the statue of a blond standing Labrador (73 cm high on a 35 cm elevation). In batch 3, the dog statue was replaced by a sitting brown and white Bulldog (37 cm high on a 61 cm elevation) because the former statue broke prior to the test. The olfactory cue consisted of 2 urine-saturated cotton pads obtained from American Bulldogs (Dierenopvang de Wissel, Leeuwarden, The Netherlands) and positioned underneath the dog statue. Samples were stored at minimum -18 °C and thawed 24 h before use. The auditory cue was a 5 s recording of a growling dog, played with a Bluetooth speaker located underneath the dog statue.

Attention bias test

On weeks 9 and 15 of the experiment, all heifers were subjected to the attention bias test. The order of testing was pseudo-randomised based on the experimental housing conditions and kept identical between the two tests. For practical considerations, heifers from the same focal group were brought together to a waiting arena located approximatively 17 m from the testing arena. Within a group, heifers' testing order was randomised. At heifer's entrance in the arena, the dog statue was visible, and the urine sample's lid was open. Once the heifer had crossed the start line and made visual contact with the statue, the 5 s auditory cue was played. Ten seconds after the visual contact, the dog model was removed by covering the hole in the tarp and closing the urine sample's lid. If the heifer did not see the dog statue within 30 s, her attention was drawn to it by playing the auditory cue, and the dog model was hidden 10 s later. The test started once the dog model was removed and heifer's behaviour was scored during 120 s, from video (**Table 8**). Faeces were removed between trials.

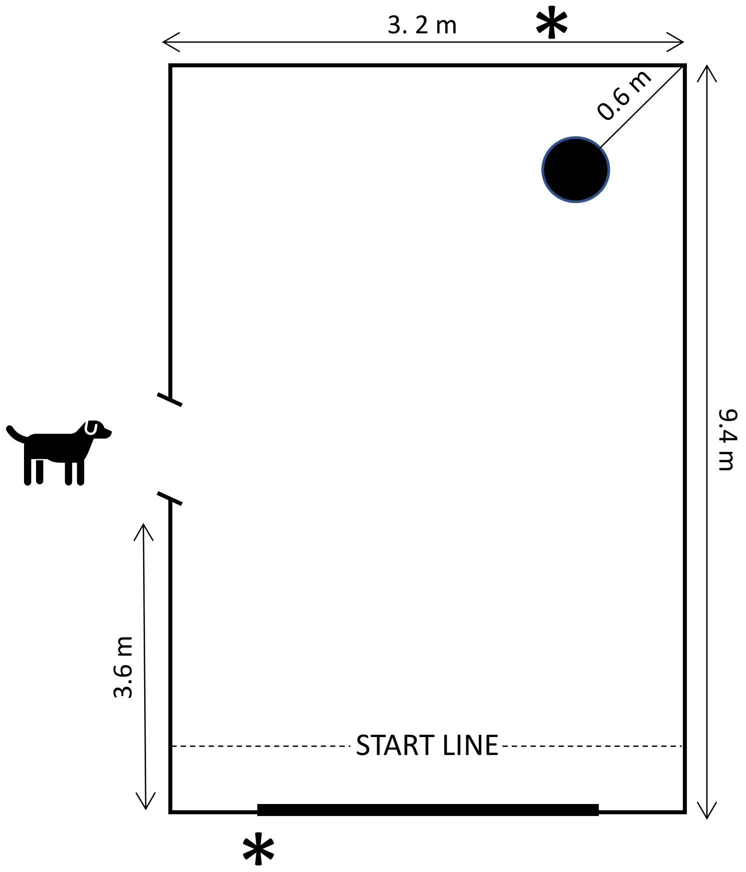


Figure 4. Schematic layout of the attention bias arena. The black circle is the positive cue (i.e. a bucket with 500 g concentrates) and the dog is the threat (dog model). The stars show the locations of the cameras.

Table 8. Definitions of the behavioural measures recorded during the Attention Bias Tests.

Behaviour	Definition
Attention to the threat	Looks at the closed tarp (next 120 s) with binocular vision – i.e. the head is directed towards the threat
Attention to the bucket	Looks at bucket with binocular vision, in a direct line – i.e. head is directed toward the bucket
Feeding	Places the muzzle within cm from or inside the bucket
Relative positive attention	Time spent looking at the bucket and feeding relative to the total time spent looking at the bucket, feeding and looking at the threat (adapted from Raoult and Gygax 2019)
In locomotion	At least one leg moves
In proximity with walls/floors	Sniffs, touches, licks or chews on the floor or the walls (tarp) of the arena
Head up	Head raised above the withers, when the heifer is not *in proximity with walls/floors*

Statistical analyses

Datasets and scripts are available online (**Supplementary Dataset 1** and **Supplementary Script 1**, respectively). All statistical analyses were performed using R 4.0.5.

Identification of dairy cattle personality traits

A Principal Component Analysis was used to identify the dimensions of personality among heifers (N=47). The substitute heifer was excluded from the analyses. PCA analysis followed researchers' recommendation (Budaev 2010). In total, six measures were included within the PCA: the proportion of time spent in contact with the object, the latency to touch the object, the proportion of time spent in contact with the walls during the OF and the NO tests, the proportion of time spent in locomotion during the OF and the NO tests, the number of locomotion bouts during the OF and the NO tests, and the proportion of time spent less than 2 m away from the group during the RW test. As recommended, the number of behavioural measures included within the PCA was minimised to ensure an [animals : parameters] ratio superior to 5 (Budaev 2010). The selected behaviours were reduced to behaviours that were not highly correlated with each other (r>0.7), and behaviours with small in-between animal variability were also disregarded. Latencies were expressed as proportions of total test duration. Latencies and count data were log-transformed (log(y+1)) or square transformed; and proportion of times were logit transformed (log(y/(1-y)) using y=0.1*minimum(proportion of time) when y=0 (van Reenen et al. 2004) to achieve approximate normality (Webb et al. 2015). PCA was performed on the correlation matrix, and the first three factors were included and subjected to varimax rotation. The number of rotated components (RC) was selected based on the number of components explaining more than 75 % of the total variance. Loadings rated 'excellent' (|value|> 0.71) were considered for further interpretation (Comrey and Lee 1992). All communalities were higher than 0.7. Heifers' scores on the three main components were extracted from the PCA. For each component, heifers were thereafter classified into two classes, based on the component's median score (RC1: -0.08, RC2: 0.23, RC3: 0.26).

Personality effect on heifers' responses to the judgement and attention bias tests under the reference conditions

The effect of personality traits on heifers' responses to the cognitive bias tests was assessed in the reference conditions only, i.e. when differences in housing-induced affective states between heifers were minimal. Response variables were expressed as proportions of trial duration (i.e. 90 s for the JBT and 120 s for the ABT) or as proportions of total time spent paying attention to one cue or the other (i.e. *Relative positive attention*). Analyses were conducted using generalized linear mixed models (GLMMs). Analyses were performed by penalized quasi-likelihood (Breslow and Clayton 1993) employing routine glmmPQL from the MASS library. The GLMMs comprised of a logit link and a binomial variance function with an extra multiplicative overdispersion parameter. Fixed effects on the logit scale included batch, each personality trait (expressed as two-levels categorical variables), as well as two-way and three-way interactions between personality traits. Random effects included group (Gygax 2014). Wald tests were performed to assess the main fixed effects in all GLMM analyses. Pairwise comparisons were based on a Fisher's LSD procedure with Bonferroni correction. Analyses of personality effects on heifers' responses to the judgement and attention bias tests were conducted on 42 (i.e. trained heifers) and 43 heifers, respectively. Two heifers were excluded from the ABT analyses because they saw the experimenters behind the curtains, and three additional heifers were excluded due to a technical failure of the Bluetooth speaker.

Housing effect on heifers' responses to judgement and attention bias tests

The effect of housing on heifers' responses to the cognitive bias tests in the reference and in the experimental conditions was investigated in a longitudinal fashion to control for inter-individual variation. Response variables were thus longitudinal data defined as heifers' behavioural responses to the cognitive bias tests *both* in the reference and in the experimental conditions. Again, analyses were conducted using GLMMs which comprised a logit link and a binomial variance function. Fixed effects included batch and housing (reference, positive, negative) while random effects included heifer nested in group (Gygax 2014) to account for any source of individual variation, including personality. Analyses of housing effects on heifers' responses to JBT and ABT were conducted on 41 and 38 heifers, respectively. One trained heifer was removed from the JBT analyses because she suddenly stopped reaching the feeder a week preceding the second testing. Five heifers were removed from ABT analyses because of technical issues.

Relationships between cognitive bias responses in the reference and in the experimental conditions

Relationships were investigated using analyses of covariance. Consistency was assessed by investigating the effects of behavioural response *i* to the JBT and to the ABT in the reference conditions on behavioural response *i* to the JBT and to the ABT in the experimental conditions. All responses to the cognitive bias tests were expressed in proportions. For each GLMM, fixed effects included the logit-transformed behavioural response under the reference conditions, housing (positive, negative) and their interaction. The random effects included group. Analyses of responses consistency to JBT and ABT were conducted on 41 and 38 heifers, respectively.

Exploratory analyses: personality and housing interactions on heifers' responses to the judgement and attention bias tests in the experimental conditions

The examination of potential interaction effects between personality and housing on heifers' responses to the judgement and attention bias tests in the experimental conditions was purely explorative, and not part of the original experimental design. Hence, heifers were not allocated to the positive or to the negative housing conditions based on their personality traits. Consequently, the analyses described are preliminary and the results derived from these analyses should be considered as such. Models were built following a step-by-step approach. The analyses initially included the fixed effects for each personality trait, housing (positive, negative) and all possible two-way and three-way interactions – except for Activity:Fearfulness:Sociability and Treatment:Activity:Sociability due to singularity and convergence issues. From this model skeleton, the selection process of the final models was carried out as follow: 1) removal of three-way interactions with p-values higher than 0.10, 2) removal of two-way interactions between two personality traits with p-values higher than 0.10 and 3) removal of two-way interactions between personality traits and housing with p-values higher than 0.10. Analyses of personality and housing interactions on heifers' responses to the JBT and to the ABT were conducted on 41 and 38 heifers, respectively.

Relationships between heifers' responses to the judgement and the attention bias tests

Relationships between heifers' *Average latency to reach the ambiguous cues* and each behavioural response obtained in the ABT were examined in the reference and in the experimental conditions separately, using Spearman's rank correlations. In the reference conditions, tests were performed on the raw data expressed as proportions. In the experimental conditions, tests were performed on the residuals extracted from the GLMM analyses modelling the sole effect of housing (positive, negative) on heifers' responses to the judgement and attention bias tests. In total, 38 and 32 heifers were included for the analyses in the reference and in the experimental conditions, respectively.

Supplementary materials

Supplementary Table 1. Effects of Personality and Housing on heifers' responses to the Judgement and Attention Bias Tasks. Housing (Hous), Activity (Act), Fearfulness (Fear), Sociability (Soc). Significant values ($p<0.05$) are written in bold. Dashes for two-way interactions indicate that the interaction was dropped from the model ($p>0.10$). Dashes for main effect indicate that the respective factor is engaged in an interaction.

Response variables	**Explanatory variables**								
In the Judgement Bias Task	Hous	Act	Fear	Soc	Hous:Act	Hous:Fear	Hous:Soc	Act:Soc	Fear:Soc
Average latencies to reach the ambiguous cues	-	-	-	-	-	**0.007**	-	**0.004**	-
In the Attention bias task									
Latency to look at the threat	-	-	0.140	-	**<0.001**	-	**0.013**	-	-
Latency to eat	0.790	0.430	0.084	0.530	-	-	-	-	-
Time spent looking at the threat	-	-	-	-	**0.035**	-	-	**-**	**0.011**
Time spent eating	0.084	0.170	0.075	0.100	-	-	-	-	-
Relative positive attention	0.620	0.066	**0.002**	**0.033**	-	-	-	-	-
Time spent in locomotion	**0.024**	0.380	0.058	0.075	-	-	-	-	-
Time spent with head up	**0.005**	0.062	**0.003**	**0.003**	-	-	-	-	-
Time spent in contact with the walls/floor	0.490	0.420	0.880	0.980	-	-	-	-	-

Supplementary Table 2. Behavioural measures obtained in the attention bias test under the reference and the positive conditions *only* for heifers eventually housed under the positive conditions. Results are presented according to personality traits and classes of personality trait. Results are expressed in proportion of trial duration (120 s), except for *Relative positive attention* that is expressed in proportion of heifer's total time spent at looking at the stimuli.

Personality		**Latency to look at the threat**		**Latency to eat from the bucket**	
Trait	**Class**	**Reference**	**Positive**	**Reference**	**Positive**
Activity	Active	2 ± 1.0	1 ± 0.4	58 ± 12.3	31 ± 14.0
	Inactive	3 ± 0.6	1 ± 0.4	67 ± 11.3	40 ± 11.1
Fearfulness	Fearful	2 ± 0.6	1 ± 0.5	69 ± 11.3	50 ± 12.0
	Non-fearful	3 ± 1.0	1 ± 0.4	56 ± 12.1	24 ± 11.5
Sociability	Social	1 ± 0.4	1 ± 0.4	69 ± 11.6	38 ± 15.6
	Non-social	3 ± 0.9	2 ± 0.4	58 ± 11.6	35 ± 10.0

Personality		**Time spent looking at the threat**		**Time spent eating from the bucket**		**Relative positive attention**	
Trait	**Class**	**Reference**	**Positive**	**Reference**	**Positive**	**Reference**	**Positive**
Activity	Active	16 ± 4.8	5 ± 1.4	18 ± 7.6	34 ± 11.3	50 ±11.0	80 ± 6.7
	Inactive	17 ± 4.7	7 ± 2.4	12 ± 5.8	38 ± 9.6	46 ± 10.3	76 ± 8.5
Fearfulness	Fearful	21 ± 5.4	8 ± 2.8	8 ± 5.1	25 ± 7.7	33 ± 8.8	66 ± 9.9
	Non-fearful	10 ± 2.8	5 ± 1.5	23 ±7.5	46 ± 11.3	64 ± 10.4	88 ± 4.0
Sociability	Social	13 ± 3.0	5 ± 1.7	14 ± 6.7	30 ± 11.1	55 ± 9.9	74 ± 10.2
	Non-social	19 ± 5.3	7 ± 2.4	16 ± 6.6	41 ± 9.6	42 ± 10.6	80 ± 6.3

Personality		**Head up**		**In locomotion**		**In contact with walls/floor**	
Trait	**Class**	**Reference**	**Positive**	**Reference**	**Positive**	**Reference**	**Positive**
Activity	Active	30 ± 7.6	22 ± 5.7	23 ± 4.1	21 ± 3.8	13 ± 2.3	11 ± 4.1
	Inactive	35 ± 7.0	19 ± 7.7	19 ± 2.2	16 ± 2.8	13 ± 3.9	14 ± 4.8
Fearfulness	Fearful	38 ± 6.4	28 ± 9.5	23 ± 3.7	22 ± 3.7	11 ± 2.4	13 ± 3.3
	Non-fearful	27 ± 7.8	13 ± 4.2	19 ± 2.3	15 ± 2.4	16 ± 4.0	13 ± 5.8
Sociability	Social	23 ± 6.5	27 ± 12.3	22 ± 4.2	16 ± 3.5	19 ± 3.9	14 ± 7.8
	Non-social	40 ± 6.8	16 ± 4.2	20 ± 2.5	20 ± 3.1	8 ± 1.9	13 ± 3.1

Supplementary Table 3. Behavioural measures obtained in the attention bias test under the reference and the negative conditions *only* for heifers eventually housed under the negative conditions. Results are presented according to personality traits and classes of personality trait. Results are expressed in proportion of trial duration (120 s), except for *Relative positive attention* that is expressed in proportion of heifer's total time spent at looking at the stimuli.

Personality		**Latency to look at the threat**		**Latency to eat from the bucket**	
Trait	**Class**	**Reference**	**Negative**	**Reference**	**Negative**
Activity	Active	1 ± 0.6	6 ± 2.4	56 ± 11.9	38 ± 11.9
	Inactive	2 ± 1.1	13 ± 5.5	45 ± 16.5	15 ± 9.5
Fearfulness	Fearful	2 ± 0.6	7 ± 3.0	58 ± 12.3	37 ± 14.8
	Non-fearful	2 ± 1.0	8 ± 3.7	45 ± 14.9	26 ± 11.6
Sociability	Social	2 ± 0.7	6 ± 2.7	58 ± 11.9	28 ± 11.5
	Non-social	2 ± 0.9	10 ± 4.6	40 ± 16.3	36 ± 15.8

Personality		**Time spent looking at the threat**		**Time spent eating from the bucket**		**Relative positive attention**	
Trait	**Class**	**Reference**	**Negative**	**Reference**	**Negative**	**Reference**	**Negative**
Activity	Active	11 ± 2.2	11 ± 4.3	21 ± 9.0	44 ± 10.6	55 ± 8.9	70 ± 10.7
	Inactive	9 ± 3.7	3 ± 0.4	40 ± 14.9	63 ± 11.0	62 ± 16.0	96 ± 0.5
Fearfulness	Fearful	11 ± 2.7	9 ± 2.7	10 ± 4.9	45 ± 13.8	51 ± 8.8	70 ± 12.7
	Non-fearful	9 ± 2.8	9 ± 5.7	46 ± 13.4	53 ± 10.2	64 ± 13.8	85 ± 10.3
Sociability	Social	13 ± 2.6	10 ± 4.8	23 ± 8.7	47 ± 10.0	50 ± 11.1	75 ± 11.5
	Non-social	6 ± 1.8	6 ± 2.1	38 ± 16.9	53 ± 15.7	73 ± 8.9	81 ± 10.1

Personality		**Head up**		**In locomotion**		**In contact with walls/floor**	
Trait	**Class**	**Reference**	**Negative**	**Reference**	**Negative**	**Reference**	**Negative**
Activity	Active	18 ± 6.0	13 ± 4.3	31 ± 5.0	13 ± 2.7	17 ± 3.2	10 ± 2.6
	Inactive	7 ± 4.3	8 ± 3.7	13 ± 2.2	12 ± 2.0	6 ± 1.9	11 ± 5.2
Fearfulness	Fearful	25 ± 6.9	15 ± 6.2	30 ± 5.5	13 ± 3.7	14 ± 3.3	8 ± 3.1
	Non-fearful	9 ± 4.3	8 ± 2.9	18 ± 4.7	12 ± 2.0	12 ± 3.8	11 ± 3.4
Sociability	Social	16 ± 3.6	13 ± 4.5	22 ± 4.4	13 ± 2.3	13 ± 3.0	11 ± 3.1
	Non-social	18 ± 10.3	8 ± 4.6	28 ± 7.6	11 ± 4.0	12 ± 4.7	8 ± 3.6

Supplementary Table 4. Behavioural measures obtained in the attention bias test under the positive and the negative conditions. Results are presented according to personality traits and classes of personality trait. Results are expressed in proportion of trial duration (120 s), except for *Relative positive attention* that is expressed in proportion of heifer's total time spent at looking at the stimuli.

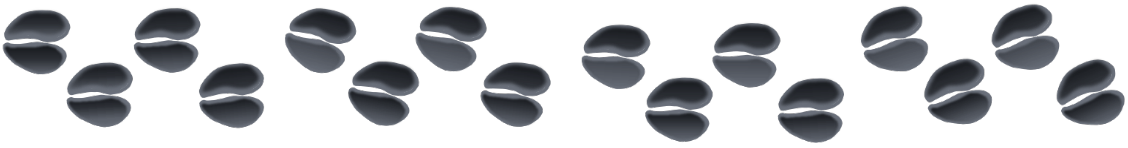
Personality		**Latency to look at the threat**		**Latency to eat from the bucket**	
Trait	**Class**	**Positive**	**Negative**	**Positive**	**Negative**
Activity	Active	1 ± 0.4	6 ± 2.4	31 ± 14.0	38 ± 11.9
	Inactive	1 ± 0.4	13 ± 5.5	40 ± 11.1	15 ± 9.5
Fearfulness	Fearful	1 ± 0.5	7 ± 3.0	50 ± 12.0	37 ± 14.8
	Non-fearful	1 ± 0.4	8 ± 3.7	24 ± 11.5	26 ± 11.6
Sociability	Social	1 ± 0.4	6 ± 2.7	38 ± 15.6	28 ± 11.5
	Non-social	2 ± 0.4	10 ± 4.6	35 ± 10.0	36 ± 15.8

Personality		**Time spent looking at the threat**		**Time spent eating from the bucket**		**Relative positive attention**	
Trait	**Class**	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
Activity	Active	5 ± 1.4	11 ± 4.3	34 ± 11.3	44 ± 10.6	80 ± 6.7	70 ± 10.7
	Inactive	7 ± 2.4	3 ± 0.4	38 ± 9.6	63 ± 11.0	76 ± 8.5	96 ± 0.5
Fearfulness	Fearful	8 ± 2.8	9 ± 2.7	25 ± 7.7	45 ± 13.8	66 ± 9.9	70 ± 12.7
	Non-fearful	5 ± 1.5	9 ± 5.7	46 ± 11.3	53 ± 10.2	88 ± 4.0	85 ± 10.3
Sociability	Social	5 ± 1.7	10 ± 4.8	30 ± 11.1	47 ± 10.0	74 ± 10.2	75 ± 11.5
	Non-social	7 ± 2.4	6 ± 2.1	41 ± 9.6	53 ± 15.7	80 ± 6.3	81 ± 10.1

Personality		**Head up**		**In locomotion**		**In contact with walls/floor**	
Trait	**Class**	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
Activity	Active	22 ± 5.7	13 ± 4.3	21 ± 3.8	13 ± 2.7	11 ± 4.1	10 ± 2.6
	Inactive	19 ± 7.7	8 ± 3.7	16 ± 2.8	12 ± 2.0	14 ± 4.8	11 ± 5.2
Fearfulness	Fearful	28 ± 9.5	15 ± 6.2	22 ± 3.7	13 ± 3.7	13 ± 3.3	8 ± 3.1
	Non-fearful	13 ± 4.2	8 ± 2.9	15 ± 2.4	12 ± 2.0	13 ± 5.8	11 ± 3.4
Sociability	Social	27 ± 12.3	13 ± 4.5	16 ± 3.5	13 ± 2.3	14 ± 7.8	11 ± 3.1
	Non-social	16 ± 4.2	8 ± 4.6	20 ± 3.1	11 ± 4.0	13 ± 3.1	8 ± 3.6

Chapter 5

Relationships between non-invasive physiological measures and indicators of affective states in dairy heifers



L. Kremer[1,2]*, C.G. van Reenen[1,2], A. Kok[3], E.A.M. Bokkers[1], G. Gort[4], J. Engel[4], J.T.N. van der Werf[2], L.E. Webb[1]

[1] Animal Production Systems group, Wageningen University & Research, Wageningen, the Netherlands

[2] Livestock Research, Wageningen University & Research, Wageningen, the Netherlands

[3] Adaptation Physiology Group, Wageningen University & Research, Wageningen, the Netherlands

[4] Biometris, Wageningen University & Research, Wageningen, the Netherlands

Abstract

Practical indicators of affective states are necessary to ensure that dairy cattle welfare can be adequately monitored on-farm. The gold-standard methodology to assess affective states, the judgement bias task (JBT), is however not suitable for on-farm assessment as it requires animal training. This study investigated the validity of three categories of non-invasive and relatively easy-to-collect physiological measures as indicators of affective states in dairy heifers: hair cortisol, heart rate variability (HRV) measures and milk-derived measures. These physiological measures – alongside individual responses to the JBT – were assessed when heifers (n=48) were first kept under reference housing conditions and once brought under housing conditions of contrasted hedonic quality, either positive or negative. The validity of the physiological measures as indicators of affective states was investigated following two lines of reasoning. First, we examined the influence of the different housing conditions on the physiological measures. At this point, heifers' personality was also accounted for – by measures of activity, fearfulness and sociability traits – as it influences the perception of the environment. We expected valid physiological measures of affective states to vary according to the valence continuum delineated by the housing contrasts (i.e. from positive to negative). Second, we examined the correlations between the physiological measures and heifers' responses to the JBT. We hypothesised that valid indicators of affective states would co-vary with measures of judgement bias and share common underlying traits. None of the physiological measures varied according to the valence continuum or were consistently correlated with measures of judgement bias across housing contexts. Yet, we found that the housing conditions influenced the relative power of the high-frequency band – a measure of HRV, and milk fluctuations in a personality-dependent manner. Inactive heifers (from 9.0 % ± 1.40 to 11.3 % ± 2.16), non-fearful heifers (from 9.4 % ± 1.76 to 16.5 % ± 3.52) and non-social heifers (from 8.4 % ± 0.91 to 10.0 % ± 1.64) moved to the positive conditions had higher relative power of the high-frequency band, and fearful heifers moved to the negative conditions had greater milk fluctuations (from 0.55 kg ± 0.136 to 1.33 kg ± 0.350). These results suggest that the relative power of the high frequency band and milk fluctuations may constitute promising indicators of positive and negative valence, respectively, but further research is warranted to confirm this idea.

Keywords

Dairy cattle – Hair cortisol – Heart rate variability – Milk yield – Cognitive bias – Personality

Introduction

Within the farm animal industry, the dairy sector would particularly benefit from research on animal welfare. Dairy cattle welfare is still considered the second greatest livestock welfare problem in the European Union (Broom 2017). Such a statement, therefore, suggests that there is ample room for dairy cattle welfare improvements and urges research efforts to go in this direction. Research progress on cow welfare could lead to substantial ethical progress as the quality of life of up to 22.6 million European dairy cows could be enhanced (European Commission 2019). Such progress would also help the dairy industry to meet the societal demand for better animal treatment, thereby mitigating the risk of economic loss due to bad publicity (Sinclair et al. 2019).

As the scientific understanding of non-human animal (hereafter animal) welfare evolves, methodologies to assess animal welfare must be refined. Nowadays, the definition of animal welfare hinges on the notion of affective states, insomuch that animal welfare is considered optimal when the balance between positive and negative affective states is overall positive (Green and Mellor 2011). Animal affective state reflects individual subjective experiences of events. It is commonly characterised in a two-dimensional model according to arousal (i.e. low to high activation) and valence (positive to negative) (Mendl et al. 2010b). In the context of animal welfare studies, the measure of valence, over arousal, is of primary importance since valence codes information on the intrinsic pleasantness or aversiveness of the animal experience. Therefore, adequate evaluation of animal welfare requires valid measures of affective states (Watanabe 2007), in particular of affective valence.

The gold-standard methodology to assess animal affective valence is not a suitable tool for on-farm assessment. Nowadays, researchers commonly infer affective valence based on measures of judgement biases (Harding et al. 2004; Bateson and Nettle 2015), which reflect affect-driven shifts in individual interpretation of ambiguous stimuli (Harding et al. 2004; Paul et al. 2005). Judgement biases are usually assessed using Judgement Bias Tasks (JBTs), the principle of which relies on the idea that animals in negatively-valenced affective states judge ambiguous situations more negatively (i.e. are more pessimistic) than animals in positively-valenced affective states – and vice versa (Harding et al. 2004). The JBT, however, does not constitute a practical tool for on-farm assessment of affective valence since it often requires extensive periods of animal training (Roelofs et al. 2016). Recently, a second cognitive bias tool called the Attention

Bias Task (ABT) has been developed, the principle of which relies on the idea that anxious individuals have heightened attention towards threatening stimuli. Unlike the JBT, the ABT does not require animal training, and therefore allows researchers to assess certain animal affective states (e.g. anxiety) more rapidly than JBT does. In practice, nonetheless, ABT is also not suitable for on-farm assessment of animal affective states since its implementation often requires specific facilities and its analysis involves extensive periods of behavioural observation. More practical measures of affective valence must, therefore, be identified to provide dairy farmers with a tool allowing them to detect shifts in their cows' affective valence. Physiology could provide such a tool.

Three non-invasive categories of physiological measures have been identified as promising indicators of affective valence in human and animal literature. First, hair cortisol (HC) may constitute a practical and valid metric of affective valence. HC can be collected via simple hair shaving and its concentration has been shown to increase in humans in response to negative affective states such as chronic stress (Stalder et al. 2017) and psychiatric disorders (Wosu et al. 2013). In cows, elevated concentrations of HC have been associated with impaired welfare (e.g. body injuries and dehydration: Sharma et al. 2019). However, no correlations were found between HC concentrations and traditional on-farm welfare assessments scores (van Eerdenburg et al. 2021). Therefore, the extent to which HC constitutes a valid indicator of cow affective state remains to be confirmed.

Second, measures of Heart Rate Variability (HRV) may constitute promising indicators of affective valence of practical relevance (Boissy et al. 2007) since they can be remotely collected on commercial farms using wearable belts. Heart rate variability reflects the response flexibility of the heart to environmental demands (Thayer et al. 2009) and its association with psychological processes has often been studied. In humans, negative affective states (e.g. depression and chronic stress) have repeatedly been associated with low HRV (Sgoifo et al. 2015; Da Estrela et al. 2021), whereas positive affective states (e.g. cheerfulness and calmness) have been associated with higher HRV (Geisler et al. 2010). In dairy cows, similar results have been reported. HRV was lower in cows subjected to chronic stress (Kovács et al. 2015), while it increased in response to a positive experience (e.g. human stroking combined with gentle live talking: Lange et al. 2020). Further studies are, nonetheless, required to investigate the generalizability of these promising findings to different affective contexts in dairy cows.

Finally, measures derived from milking data could also provide valuable, non-invasively collected and easy-to-access information regarding cow affective valence. Milk yield (MY), for instance, is automatically recorded in modern dairy farms and known to decrease when cow health is impaired (e.g. mastitis: Rajala-Schultz et al. 1999) and in response to stressors (e.g. mixing: Phillips and Rind 2001). Alternatively, however, research has yet to determine whether MY increases in response to affective improvements. Other measures such as daily milk fluctuation and milk composition may also constitute promising indicators of affective valence, which can be easily obtained from MY recordings and routine milk analyses. Recent studies in dairy cows showed that lower milk fluctuation and lower fat content were associated with increased animal ability to recover from challenges (Poppe et al. 2020; Poppe et al. 2021). While the underlying mechanisms behind individual resilience remain largely unknown (Southwick and Charney 2012; Southwick et al. 2014), empirical studies have demonstrated the contribution of psychological factors to one's ability to respond to adversity (de la Fuente et al. 2021). It is therefore plausible that indicators of resilience also convey information relative to affective valence, but this theory remains to be proven.

Based on these considerations, the objective of this study was to investigate the validity of the aforementioned physiological measures as indicators of affective valence. Two approaches were adopted to reach this goal. First, we investigated the effects of contrasted housing conditions on the physiological measures of interest since housing conditions of contrasted hedonic values may effectively induce positive and negative shifts in animals (Douglas et al. 2012; Bučková et al. 2019). We hypothesised that valid physiological indicators of affective valence would vary according to the valence continuum delineated by the housing conditions. Considering that personality may also influence how individuals perceive their living conditions (Kremer et al. 2021a), personality effects on the physiological measures were also investigated here – similarly to van Reenen et al. (2004). Second, we investigated the correlations between the physiological measures of interest and presumably valid indicators of affective valence – i.e. judgement and attention biases. We hypothesised that valid physiological indicators of valence would correlate (either positively or negatively) with measures of cognitive bias, considering that indicators of valence should have common underlying traits.

Materials and methods

Animals and management

The study was carried out at the Dutch experimental farm of Dairy Campus in Leeuwarden between February 2019 and January 2021. The experiment was approved by the ethical committee of Wageningen University & Research (the Netherlands).

The experiment lasted a year and was divided into three batches of four groups. Each group consisted of eleven lactating and pregnant Holstein Friesian cows, among which four primiparous cows (hereafter referred to as 'heifers') and seven multiparous companion cows. Heifers were pseudo-randomly allocated to their group to balance for their days in milk (165 d ± 5.5), milk production (25.2 kg ± 0.6) and body weight (606 kg ± 6.0). Companion animals were pseudo-randomly allocated to their group to balance for their parity (3 ± 0.1), milk production (30.4 kg ± 0.7) and body weight (707 kg ± 7.4). All individuals were healthy and confirmed pregnant at the beginning of the experiment. One heifer in batch 3 was replaced during the second week of the experiment due to miscarriage.

All groups were housed in adjacent pens within the same barn. Solid partitions were installed between the pens to prevent visual and tactile contacts between the groups. Every morning around 7:00 h, each group was delivered a total mixed ration of maize silage (35 % - percentage based on dry matter), grass silage (30 %), concentrates (20 %), grinded whole soy (10 %), grinded whole wheat (3 %) and minerals (2 %). Additionally, each group had free access to one water trough and to one automatic concentrate dispenser delivering a pre-determined daily amount of concentrates based on individual milk production. All cows were milked twice a day around 05:00 h and 15:00 h.

Housing conditions

Reference phase

Groups were first housed under reference conditions for nine weeks. Each group had access to exactly 11 cubicles, 11 feeding gates and one fixed brush. The social composition of the group was kept constant.

Experimental phase

Groups were subsequently housed under supposedly weekly-improved or weekly-worsened conditions (hereafter referred to as ‘positive’ and ‘negative’ housing, respectively) for six study weeks. Here, the term ‘study week’ refers to an experimental period of 7 d during which the housing conditions were constant, i.e. from Saturday to Friday since housing changes were always implemented on Fridays around the afternoon milking time (at 15:00 h). Housing changes are described in detail in the supplementary materials (**Supplementary Table 1**). Succinctly, they revolved around three housing facets which are known to influence cow welfare: the stocking density (Fregonesi et al. 2007; Schütz et al. 2015; Winckler et al. 2015), the social stability within the groups (Schirmann et al. 2011; Wilcox et al. 2013; Gutmann et al. 2015) and the level of enrichment (Devries and Keyserlingk 2006; Huzzey et al. 2006; McConnachie et al. 2018). The two groups in the positive conditions had more space and access to progressively more cubicles and feeding gates (maximum 14 per pen), while the other two groups in the negative conditions had less space and access to progressively fewer cubicles and feeding gates (minimum 6 per pen). Additionally, the group composition remained unchanged in the positive conditions, whereas groups in the negative conditions were subjected to frequent mixing. Finally, groups in the positive conditions were provided with additional brushes, while groups in the negative conditions were brush-deprived.

Personality tests

In week 2 of the reference conditions, heifers (n=47) were subjected to three standard personality tests, namely the open-field, the novel-object and the runway tests. The replacement heifer in batch 3 could not be subjected to the personality tests. Protocols were adapted from previous studies conducted in cattle (Gibbons et al. 2010; van Reenen et al. 2004) and are explained in detail elsewhere (Kremer et al. 2021a). Succinctly, during the open-field test, heifers were brought into an unfamiliar arena, where they remained alone for 10 min. At the end of the open-field test, the novel-object test immediately started, and a novel object was remotely lowered from the ceiling of the arena until it touched the floor. The novel object, which consisted in two orange cones attached together, was then lifted 1 m above the floor and remained in this position for the entire duration of the test, i.e. for 10 min. Heifers’ behaviours in the open-field and the novel-object tests were video recorded (CAMCOLBUL2, Velleman, Belgium) and analysed using The Observer XT 10 (Noldus Information Technology BV, Wageningen, the Netherlands). On the days following the open-field and the novel-object tests, heifers were subjected to the runway test, which was conducted in a corridor familiar to the heifers and leading to the milking parlour. Prior to the test, a group of four heifers and two companion cows from the same pen were brought to a waiting area for 10 min. The focal heifer was subsequently brought 18 m away from the waiting area and positioned into a starting area, where she could face the rest of group for 1 min. Thereafter, the metal bar preventing the heifer from reaching the rest of group was removed and the test started for 5 min. Heifers' behaviours in the runway test were live scored using a computer equipped with The Observer XT 10. The outcomes of these behavioural tests are described in a previous paper (Kremer et al. 2021a) and used in the current study. They revealed that heifers' personality could be characterised in a three-dimensional space according to their activity, fearfulness and sociability (hereafter referred to as personality trait). Heifers' scores on each personality trait were extracted from the PCA. For each personality trait, heifers were, thereafter, classified into two classes depending on the median score of the considered trait (activity: -0.08, fearfulness: 0.23, sociability: 0.26). The distribution of heifers in the different housing conditions according to their personality traits is detailed in the supplementary materials (**Supplementary Table 2**).

Cognitive tasks

Judgement Bias Task

For a detailed description of the Go/NoGo spatial JBT procedure, see Kremer et al. (2021a). The JBT protocol was based on Destrez et al. (2013) and Lecorps et al. (2018). In brief, heifers were trained to discriminate between two locations of a feeder, signalling either a reward (i.e. 'positive' location P) or a punisher (i.e. 'negative' location N). The reward consisted of 150 g concentrates and the punisher consisted of a 7 bar air puff. This combination of reward/punisher was selected to maximise the sensitivity of the JBT to shifts in heifers' affective states (Kremer et al. 2021b). A training session was composed of 10 trials, during which the feeder was located on P for 6 trials and on N for 4 trials. The order of the trials was pseudo-randomly selected to ensure the heifers' motivation to participate in the task. Heifers were considered trained if they reached the feeder on P within 20 s and if they did not reach the feeder on N within 90 s (i.e. the trial duration) during trials from the same training session. Once trained, heifers were exposed to three ambiguous locations of the feeder that were interspersed between P and N. One feeder location (A) was equidistant from P and N, and the two other locations were equidistant from A and P (Ap) and from A and N (An). Latencies to reach the cues were video recorded (CAMCOLBUL2, Velleman, Belgium). In total, heifers were exposed twice to the ambiguous cues: in week 8 of the reference conditions and in week 14 of the experimental conditions. As for the personality tests, the outcomes of the test are described elsewhere (Kremer et al. 2021a). Here, judgement bias is expressed as the average latency to reach all three ambiguous cues during one testing session, and heifers were assumed to become more optimistic (pessimistic) if they reached the ambiguous cues faster (slower) in week 14 compared with week 8.

Attention Bias Task

For a detailed description of the ABT procedure, see Kremer et al. (2021a). The ABT protocol was based on Lee et al. (2016, 2017). In brief, heifers were simultaneously exposed to a positive cue and a threat for 120 s in a familiar arena. The positive cue consisted of a bucket familiar to the heifers and filled with 500 g of concentrates. The bucket was positioned in the far-right corner from the arena entrance. The threat consisted of an out-of-sight dog model built from a dog statue associated with the growling sound and the urine smell of a dog. The dog model was positioned behind the left fence of the arena, at mid-distance from the entrance and the opposite wall of the arena. Prior to the test, the dog statue was made visible to the focal heifer, and the growling sound was played once the cow had made the first visual contact with the statue. After 10 s, the dog model was covered with tarpaulin and the attention bias test started. Heifer’s behaviour during the task was also video recorded. The outcomes of the ABT are also described elsewhere (Kremer et al. 2021a). Here, attention bias is expressed as heifers’ first latency to look at the threat, i.e. once the tarpaulin covered the dog model.

Physiological measures

Hair cortisol

Hair was shaved along the left side of heifers’ vertebral column using Aesculap Favorita CL hair clipper and was discarded on week 4. The clipper was dusted off between each heifer. Hair from the dominant colour of the focal heifer was then collected at the end of both reference and experimental phases in week 9 (N=48) and

15 (N=48) from the previously shaved region. Therefore, the hair collected in the reference phase was the hair grown during the last 6 weeks of the reference phase, and the hair collected in the experimental phase was the hair grown during the 6 weeks of the experimental phase. Hair samples were then stored in zip-lock plastic bags at -20°C and in darkness until assays. Hair samples were prepared and analysed at the Dutch Animal Health Service (Gezondheidsdienst voor Dieren, Deventer, the Netherlands) based on a protocol adapted from previous studies (Koren et al. 2002; Burnett et al. 2014). Hair samples were washed three times with decanted isopropanol and dried during approximatively 24 h. Samples were then cut with clean scissors and ground into powder using a Retch grinding cup with a 20 mm bead in a Retch beater for 5 min. After grinding, about 200 mg of hair powder were transferred into glass tubes wherein methanol was added and mixed with the powder. Tubes were thereafter sonicated for 30 min, placed in an oven for about 24 h and centrifuged. After centrifugation, a solution of 4 mL-supernatant was collected from each tube and its content in methanol was evaporated using nitrogen. Subsequently, 500 µl of physiological saline solution was added to each tube and mixed with the tube content. Tubes were then placed in an ultrasonic bath for 10 min and the solution was mixed. Cortisol concentration was then measured by chemiluminescence on an Immulite 1000 plus with the associated test kit for cortisol analysis (LKCO1 Siemens). Intra- and inter-assay coefficients of variability were of 6.5 % and 11.2 %, respectively.

Heart rate measures

Heifers were habituated to wearing Zephyr Bioharness^{TM} heart rate belts for minimum 1 h, 1 h 30 and 2 h for three days on study weeks 4 and 5. In study weeks 6 to 8 of the reference conditions and 10 to 14 of the experimental conditions, heart rate data were collected once a week for each heifer (N=16 per batch). On recording days (Tuesday to Friday), the heart rate data of four heifers were measured from 8:30 to 15:00. Heifers were pseudo-randomly attributed to one day of recording to ensure that 1) heart rate measurements were obtained from one heifer per group and 2) the focal heifers were not planned for judgement training or testing on that specific day. For practicality, the weekly order of heifers' heart rate recording was kept constant during the experiment.

Heart rate belts were adjusted to heifer physiognomy using rubber straps. Belts were attached around the heifers' thorax when individuals were either locked at the feeding gates or standing up in a cubicle. After securing the belts, the electrodes' sites were covered with ultrasound transmission gel for optimal electrode-skin contact. A pilot study was conducted to identify the optimal positions of the two electrodes on the heifers' thorax. Eventually, the electrodes were positioned above the focal heifer's left elbow joint and on the sternum. An accelerometer tag (IceQube, IceRobotics, South Queensferry, UK) was also attached to the individual left hind fetlock joint to synchronise heart rate data with lying bouts, since posture and activity are known to influence HRV measures (Hagen et al. 2005).

The RR data, which represent the time-intervals between the R peaks of two consecutive heartbeats, were exported and analysed using Kubios HRV Premium 3.5.0. The noise detection level was set at medium and the remaining artefacts were corrected using the automatic algorithm correction that was previously validated (Lipponen and Tarvainen 2019). Heart rate measures were calculated for every 5-min time window of each RR signal stream. The measures selected for statistical analyses are described in **Table 1**. Frequency-domain measures were estimated from the detrended RR series using Fast-Fourier Transformation. HRV frequency bands were set following the recommendations of von Borell and colleagues (2007). Nonlinear measures were estimated using Pointcaré plot. Finally, only time-windows during which heifers were lying down and which contained less than 5 % of corrected beats were included in the statistical analyses, as previously advised (von Borell et al. 2007). For analyses, all heart rate measures were averaged across both the reference and the experimental conditions for each heifer.

Table 1. Definitions and units of the heart rate and heart rate variability measures selected in this study (adapted from Shaffer and Ginsberg, 2017).

Heart rate measure	**Units**	**Definition**
Time-domain		
HR	beats.min^{-1}	Heart rate, i.e. the number of heart beats per minute
Frequency-domain		
LF	%	Relative power of the low frequency band (0.05-0.2 Hz)
HF	%	Relative power of the high frequency band (0.2-0.58 Hz)
Non-linear		
SD1	ms	Poincaré plot standard deviation perpendicular the line of identity. SD1 is equal to RMSSD, i.e. the root mean square of successive RR interval differences
SD2	ms	Poincaré plot standard deviation along the line ofidentity
SD1.SD2^{-1}	n.u	Ratio of SD1 to SD2
SampEn	n.u	Sample entropy, which measures the regularity and complexity of time series

Milk-derived measures

From study week 6 onwards, MY was automatically recorded at the individual level during each milking and subsequently averaged across days and study week (except for weeks 9 and 15). Additionally, 10 mL milk samples were collected four times a week in tubes containing Bronopol as a preservative to assess fat, protein and lactose composition (ISO, 2013; Qlip, Zutphen, Netherlands). Weekly fat- and protein- corrected milk (FPCM) was calculated using the MY and the milk composition of each study week based on the following formula (CVB, 2012):

$$FPCM\ (kg) = MY\ (kg) * [0.337 + 0.116 * fat\ (\%) + 0.06 * protein\ (\%)].$$

Furthermore, *one* value of milk persistency (kg.day^{-1}) was calculated for each heifer based on daily MY during both the reference (wk 6 to wk 8) and the experimental periods (wk 10 to wk 14); and it was defined as the slope of the lactation curve. Here, the lactation curve was modelled as a straight line since milk decline is said to be linear past the lactation peak (Wilmink 1987). The slope of the lactation line was determined for each heifer from the linear regression modelling MY according to the days of the aforementioned weeks.

Finally, fluctuations in daily MY were assessed by calculating the log-transformed variance (LnVar) of daily MY deviation from the fitted lactation curves, i.e. from the log-transformed variance of the residuals extracted from the fitted curve (Poppe et al. 2020). For each heifer, daily MY fluctuations in the reference and in the experimental periods were assessed separately – but from the same line used to obtain milk persistency.

Statistical analyses

All statistical analyses were performed using R 4.0.5. Scripts and data are available on request.

Regression analyses

Regression analyses were used to assess the influence of the housing conditions and of heifers' personality traits on a pre-selected set of non-invasively collected physiological measures. For response variables expressed as percentages, analyses were conducted using Generalized Linear Mixed Models (GLMMs) (Mcculloch and Neuhaus, 2014) based on approximate maximum likelihood estimation and using Laplacian integration. GLMMs employed routine glmmTMB (Brooks et al. 2017) and included a logit link for the fixed and the random effects while specifying a beta distribution for the percentages. For other (log-transformed) response variables, analyses were conducted using Linear Mixed Models (LMMs) and employing routine lmer from the lme4 package (Bates et al. 2015).

Unless specified otherwise, all models included batch, the three personality traits (each expressed as two-levels factors according to heifers' median score on the corresponding trait), the housing conditions (Reference, Positive, Negative) and the two-ways interactions between the personality traits and the housing conditions as fixed factors. Random factors included heifer nested in group. Hair colour (black or white) was also included as a fixed factor in the LMM assessing the effect of personality and housing conditions on HC, since hair colour influences HC concentrations (Ghassemi Nejad et al. 2017). Concerning models of HRV measures, heart rate was also included as a covariate to control for differences in arousal. The interaction effect between HR and housing was tested and dropped from all models, since it was not significant. Finally, for milk persistency, the explanatory variable housing was divided into two factors solely (either positive or negative) to model a single lactation curve per heifer between wk 6 and wk 14.

Potential heteroscedasticity was accounted for by assigning weights to the response variables based on 1) the number of valid time windows per heifer per housing conditions (which ranges from 6 to 184 per heifer) for the heart rate measures, and 2) the number of studied weeks per heifer per housing conditions (i.e. 3 and 5 for the reference and experimental conditions, respectively) for the milk-derived measures. Weights were scaled so that their sum were equal to the number of observations (i.e. 6680 and 8, respectively). Model residuals were inspected for normality and statistical outliers. If statistical outliers were identified, additional models from which the outliers were dropped were conducted to ensure that they did not influence the interpretation of the model outputs. Since no statistical outlier had a significant effect on any model outputs, results from the non-truncating models are presented below.

Finally, all pairwise comparisons were based on a Fisher's LSD procedure with Bonferroni correction. The Bonferroni correction was manually applied based on the number of pairwise comparisons of interest for the related effect.

Correlation analyses

Hierarchical clustering. Similarities among heifers' cognitive biases, physiological measures and personality scores were investigated in the reference and in the experimental conditions separately, using hierarchical clustering. The hierarchical clustering was performed on the correlation matrix of the residuals extracted from (1) the analyses modelling the effect of batch and group on the cognitive and physiological measures in the reference conditions and from (2) the analyses modelling the effect of batch, housing (positive, negative) and group on the cognitive and physiological measures. Within each model, the response variable was transformed as described in the previous section. Potential heteroscedasticity was, once more, accounted for by assigning weights to the response variables based on the number of valid time windows per heifer during the housing period of interest. The hierarchical clustering analyses were conducted using hclust from the stats package (R Core Team, 2021) and specifying an average linkage method. In both the reference and the experimental conditions, the relationships among the different measures were, furthermore, analysed using dendrograms – for which the heigh threshold was arbitrarily set at 0.83 to ensure the identification of a reasonable (and hence interpretable) number of clusters.

Spearman correlations. Relationships between heifers' cognitive biases, physiological measures and personality scores were examined in the reference and in the experimental conditions separately, using Spearman's rank correlations. Spearman's rank tests were performed on the residuals extracted from (1) and (2) in the reference and in the experimental conditions, respectively.

Results

The main effects of housing and personality on the different physiological measures are described in **Table 2**, together with the p-values of their interaction effects.

Hair cortisol

Activity had a greater influence on HC of heifers housed in the positive conditions than on HC of heifers housed in the negative or in the reference conditions, although differences between active and inactive heifers did not significantly differ in any housing conditions. However, HC tended to increase in active heifers moved from the reference to the positive conditions. Means and standard errors of HC concentrations are detailed per personality trait and housing conditions in **Table 3**.

Table 3. Mean ± standard error of hair cortisol concentrations (ng.g-^{1}) per personality class and housing conditions. Differences in letter superscripts indicate statistical differences ($p<0.10$) between housing conditions for heifers of the same personality class. No differences in personality classes for heifers housed within the same conditions were found.

Personality		**Housing conditions**		
Trait	**Class**	**Negative**	**Reference**	**Positive**
Activity	Active	29 ± 4.6^{ab}	**26 ± 3.8^{a}**	**30 ± 3.7^{b}**
	Inactive	36 ± 4.1	36 ± 2.8	33 ± 3.8
Fearfulness	Fearful	30 ± 4.7	30 ± 2.9	33 ± 3.4
	Non-fearful	35 ± 4.2	32 ± 4.0	30 ± 4.3
Sociability	Social	34 ± 4.0	32 ± 3.9	33 ± 5.2
	Non-social	31 ± 5.2	30 ± 2.9	31 ± 2.8

Heart rate variability measures

Means and standard errors of the HRV measures are detailed per personality trait and housing conditions in **Table 4**.

HR. HR tended to increase in fearful heifers moved from the reference to the positive conditions. It also increased in non-fearful heifers moved from the reference to the negative conditions. HR also tended to be higher in non-fearful heifers moved to the negative conditions compared with non-fearful heifers moved to the positive conditions. The influence of HR on HRV measures is detailed in **Table 5.**

Table 2. Effects of housing and personality on the different physiological measures. Hair cortisol (HC) in ng.g^{-1}, heart rate (HR) in beat.min^{-1}, relative power of the high-frequency band (HF) in %, relative power of the low frequency band (LF) in %, standard deviation perpendicular the line of identity in the Pointcaré plot (SD1) in ms, standard deviation along the line of identity in the Pointcaré plot (SD2) in ms, ratio between SD1 and SD2 (SD1.SD2^{-1}), sample entropy (SampEn), milk yield (MY) in kg, fat and protein-corrected milk yield (FPCM) in kg, milk persistency (Persistency) in kg.day^{-1}, daily milk fluctuation (LnVar) in kg, fat content in milk (pFat) in %, protein content in milk (pProt) in %, lactose content in milk (pLac) in %. NA^{1}: for milk persistency only, housing had two levels (i.e. either positive or negative) since one single value of persistency was calculated from wk5-wk8 of the reference period to wk10-wk15 of the experimental period. P-values inferior to 0.01 are written in bold, provided that they reflect either an interaction effect or a main effect that is not engaged in an interaction for which the p-value is already below 0.01. Hous:Act is the interaction effect between housing and activity, Hous:Fear is the interaction effect between housing and fearfulness and Hous:Soc is the interaction effect between housing and sociability.

Measures	**Housing**				**Activity**		
	Negative	Reference	Positive	*p-value*	Active	Inactive	*p-value*
HC	33 ± 3.1	31 ± 2.4	32 ± 2.6	0.487	28 ± 2.4	35 ± 1.9	0.606
HR	82.3 ± 0.84	79.1 ± 0.72	79.5 ± 1.43	<0.001	80.5 ± 0.92	79.5 ± 0.67	0.812
LF	39.2 ± 1.33	40.2 ± 0.92	37.5 ± 1.24	0.996	38.2 ± 0.97	40.3 ± 0.86	0.011
HF	9.6 ± 1.66	9.3 ± 0.97	12.0 ± 1.84	0.398	11.2 ± 1.26	9.0 ± 0.94	0.012
SD1	6.3 ± 0.81	6.8 ± 0.71	7.7 ± 1.52	0.555	7.4± 0.96	6.4 ± 0.60	0.503
SD2	19.7 ± 0.98	21.0 ± 0.99	20.3 ± 1.74	0.193	20.4 ± 1.01	20.6 ± 0.98	0.819
SD1.SD2^{-1}	0.32 ± 0.030	0.31 ± 0.017	0.34 ± 0.0299	0.140	0.34 ± 0.022	0.30 ± 0.016	0.201
SampEn	1.29 ± 0.037	1.34 ± 0.029	1.36 ± 0.038	0.942	1.34 ± 0.025	1.32 ± 0.030	0.698
Milk yield	22.8 ± 0.63	23.7 ± 0.58	23.4 ± 0.88	0.019	23.7 ± 0.66	23.1 ± 0.45	**0.095**
FPCM	25.6 ± 0.65	25.8 ± 0.56	25.9 ± 0.83	0.626	26.3 ± 0.63	25.2 ± 0.44	**0.062**
Persistency	-0.021 ± 0.010	NA^{1}	-0.014 ± 0.002	0.701	-0.031 ± 0.010	-0.004 ± 0.008	**0.090**
LnVar	0.71 ± 0.22	0.35 ± 0.101	0.49 ± 0.159	0.065	0.52 ± 0.131	0.44 ± 0.111	0.967
pFat	4.86 ± 0.099	4.63 ± 0.072	4.78 ± 0.108	0.036	4.79 ± 0.073	4.66 ± 0.073	0.502
pProt	3.76 ± 0.046	3.68 ± 0.042	3.68 ± 0.045	0.659	3.75 ± 0.038	3.65 ± 0.035	0.264
pLac	4.44 ± 0.019	4.50 ± 0.015	4.49 ± 0.022	**0.054**	4.48 ± 0.017	4.48 ± 0.013	0.610

Table 2. (Continued)

Measures	**Fearfulness**			**Sociability**			**Hous:Act**	**Hous:Fear**	**Hous:Soc**
	Fearful	Non-fearful	*p-value*	Social	Non-social	*p-value*	*p-value*	*p-value*	*p-value*
HC	31 ± 2.0	33 ± 2.4	0.780	33 ± 2.5	31 ± 1.9	0.383	**0.026**	0.825	0.843
HR	79.2 ± 0.59	80.8 ± 0.97	0.332	79.4 ± 0.77	80.6 ± 0.83	0.600	0.473	**0.084**	0.308
LF	40.0 ± 0.90	38.5 ± 0.97	**0.090**	38.7 ± 0.94	39.9 ± 0.90	0.406	**0.047**	0.140	0.893
HF	9.1 ± 0.57	11.1 ± 1.49	0.288	11.2 ± 1.36	8.9 ± 0.74	0.199	**0.045**	**0.011**	**0.047**
SD1	6.3 ± 0.36	7.6 ± 1.08	0.592	7.9 ± 1.03	5.9 ± 0.37	0.573	**0.019**	0.542	0.813
SD2	20.3 ± 0.91	20.7 ± 1.09	0.911	21.1 ± 1.12	19.9 ± 0.3	0.997	0.344	**0.011**	**0.014**
$SD1.SD2^{-1}$	0.31 ± 0.010	0.33 ± 0.026	0.473	0.34 ± 0.024	0.30 ± 0.012	0.238	**0.010**	**0.031**	0.153
SampEn	1.37 ± 0.023	1.29 ± 0.032	0.262	1.33 ± 0.030	1.33 ± 0.025	0.815	0.161	**0.036**	0.116
Milk yield	23.8 ± 0.635	23.06 ± 0.468	0.999	22.87 ± 0.527	24.00 ± 0.588	0.422	0.217	**0.066**	0.805
FPCM	26.0 ± 0.620	25.56 ± 0.438	0.811	25.62 ± 0.517	25.95 ± 0.566	0.870	0.207	0.188	0.626
Persistency	-0.021 ± 0.012	-0.0137 ± 0.008	0.531	-0.017 ± 0.008	-0.012 ± 0.012	0.530	0.436	**0.060**	0.673
LnVar	0.63 ± 0.124	0.32 ± 0.114	0.099	0.42 ± 0.110	0.53 ± 0.132	0.169	0.110	**0.001**	0.963
pFat	4.67 ± 0.081	4.79 ± 0.064	0.569	4.86 ± 0.069	4.59 ± 0.073	0.020	0.303	**0.028**	**0.034**
pProt	3.72 ± 0.040	3.68 ± 0.035	0.842	3.76 ± 0.039	3.64 ± 0.034	0.313	0.695	0.529	0.923
pLac	4.49 ± 0.015	4.48 ± 0.015	0.323	4.46 ± 0.014	4.50 ± 0.015	0.695	0.269	0.140	0.502

Table 4. Mean ± standard error of heart rate and heart rate variability measures per personality class and housing conditions. Heart rate (HR) is expressed in beat.min-^{1}, relative power of the high-frequency band (HF) in %, relative power of the low frequency band (LF) in %, standard deviation perpendicular the line of identity in the Pointcaré plot (SD1) in ms, standard deviation along the line of identity in the Pointcaré plot (SD2) in ms. Ratio between SD1 and SD2 (SD1.SD2^{-1}) and sample entropy (SampEn) have no unit. Differences in (a,b) superscripts indicate statistical differences ($p<0.10$) between housing conditions for heifers of the same personality class. Differences in (x,y) superscripts indicate statistical differences ($p<0.10$) between personality classes for heifers housed within the same conditions.

Personality		**HR**			**LF**			**HF**		
Trait	**Class**	**Housing conditions**			**Housing conditions**			**Housing conditions**		
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**
Activity	Active	82.9 ± 1.31	80.2 ± 1.22	78.6 ± 2.47	37.5 ± 1.67	39.3 ± 1.47	36.7 ± 1.93	**12.6 ± 2.91**^{a}	**9.6 ± 1.36**^{b}	12.8 ± 3.21^{ab}
	Inactive	81.7 ± 1.07	78.1 ± 0.77	80.2 ± 1.69	41.1 ± 2.03^{ab}	**41.2 ± 1.12**^{a}	**38.1 ± 1.67**^{b}	6.2 ± 0.62^{ab}	**9.0 ± 1.40**^{b}	**11.3 ± 2.16**^{a}
Fearfulness	Fearful	80.1 ± 1.16^{ab}	**78.3 ± 0.77**^{b}	**80.2 ± 1.33**^{a}	40.8 ± 1.67	40.5 ± 1.42	38.4 ± 1.54	9.6 ± 1.24	9.2 ± 0.91	8.2 ± 0.79
	Non-fearful	**84.4 ± 0.92**^{a}	**80.1 ± 1.23**^{b}	**78.6 ± 2.76**^{b}	37.7 ± 2.01	40.0 ± 1.19	36.3 ± 2.04	**9.5 ± 3.05**^{a}	**9.4 ± 1.76**^{b}	**16.5 ± 3.52**^{a}
Sociability	Social	82.1 ± 0.98	78.5 ± 0.87	78.0 ± 2.55	39.0 ± 2.10	39.7 ± 1.27	35.9 ± 1.42	10.3 ± 2.54	10.2 ± 1.69	14.8 ± 3.72
	Non-social	82.8 ± 1.59	79.8 ± 1.17	80.6 ± 1.67	39.6 ± 1.12	40.8 ± 1.35	38.6 ± 1.86	8.4 ± 1.68^{ab}	**8.4 ± 0.91**^{b}	**10.0 ± 1.64**^{a}

Table 4. (Continued)

Personality		**SD1**			**SD2**		
Trait	**Class**	**Housing conditions**			**Housing conditions**		
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**
Activity	Active	**7.6 ± 1.44^{a}**	7.0 ± 1.18^{ab}	**8.1 ± 2.88^{b}**	19.4 ± 1.11	20.9 ± 1.37	20.3 ± 2.99
	Inactive	4.9 ± 0.44	6.6 ± 0.85	7.4 ± 1.53	20.0 ± 1.70	21.1 ± 1.40	20.3 ± 2.10
Fearfulness	Fearful	6.3 ± 0.59	6.4 ± 0.54	5.8 ± 0.73	19.6 ± 1.53	20.8 ± 1.45	20.0 ± 1.65
	Non-fearful	6.3 ± 1.50	7.2 ± 1.36	9.9 ± 3.16	19.8 ± 1.31^{ab,}	**21.2 ± 1.37^{a}**	**20.6 ± 3.36^{b}**
Sociability	Social	6.5 ± 1.27	7.6 ± 1.29	10.6 ± 3.42	19.1 ± 1.30^{a}	21.6 ± 1.58	22.7 ± 3.43
	Non-social	5.9 ± 0.74	6.0 ± 0.54	5.7 ± 0.69	20.6 ± 1.51^{ab}	**20.4 ± 1.19^{a}**	**18.6 ± 1.67^{b}**
Personality		**SD1.SD2^{-1}**			**SampEn**		
Trait	**Class**	**Housing conditions**			**Housing conditions**		
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**
Activity	Active	**0.38 ± 0.051$_{a,x}$**	**0.32 ± 0.026^{b}**	**0.34 ± 0.053^{b}**	1.33 ± 0.045	1.35 ± 0.037	1.35 ± 0.057
	Inactive	**0.25 ± 0.013^{y}**	0.31 ± 0.023	0.34 ± 0.034	1.23 ± 0.057	1.33 ± 0.046	1.36 ± 0.053
Fearfulness	Fearful	0.32 ± 0.020	0.31 ± 0.015	0.29 ± 0.016	1.39 ± 0.042	1.39 ± 0.035	1.32 ± 0.040
	Non-fearful	0.31 ± 0.056	**0.31 ± 0.032^{b}**	**0.40 ± 0.058^{a}**	1.19 ± 0.046	1.29 ± 0.046	1.40 ± 0.068
Sociability	Social	0.33 ± 0.046	0.33 ± 0.030	0.39 ± 0.066	1.30 ± 0.044	1.34 ± 0.045	1.37 ± 0.078
	Non-social	0.29 ± 0.031	0.29 ± 0.017	0.30 ± 0.018	1.27 ± 0.068	1.34 ± 0.037	1.35 ± 0.037

Table 5. Regression coefficients (ß), standard errors (in brackets) and p-values of the heart rate effects on the different HRV measures. P-values below 0.10 are written in bold.

Response variables	**Heart rate**	
	Estimate (standard error)	p-value
LF	ß = 0.01 (0.006)	**0.075**
HF	ß = - 0.08 (0.012)	**<0.001**
SD1	ß = - 0.05 (0.008)	**<0.001**
SD2	ß = - 0.02 (0.006)	**0.004**
$SD1.SD2^{-1}$	ß = - 0.03 (0.006)	**<0.001**
SampEn	ß = - 0.01 (0.003)	**<0.001**

LF. LF decreased in inactive heifers moved from the reference to the positive conditions, and fearful heifers tended to have higher LF than non-fearful heifers regardless of the housing conditions.

HF. HF increased in active heifers moved from the reference to the negative conditions and tended to increase in non-fearful heifers moved to the negative conditions. HF also increased in inactive heifers, non-fearful heifers and non-social heifers moved from the reference to the positive conditions.

SD1. SD1 was higher in active heifers housed in the negative than in the positive conditions.

SD2. SD2 decreased both in non-fearful and in non-social heifers moved from the reference to the positive conditions.

$SD1.SD2^{-1}$. $SD1.SD2^{-1}$ increased in active heifers moved from the reference to the negative conditions. This resulted further in $SD1.SD2^{-1}$ being higher in active heifers compared with inactive heifers housed in the negative conditions, and in $SD1.SD2^{-1}$ which tended to be higher in active heifers housed in the negative conditions compared with those housed in the positive conditions. Moreover, $SD1.SD2^{-1}$ tended to increase in non-fearful heifers moved from the reference to the positive conditions.

SampEn. Fearfulness had greater influence on SampEn for heifers moved in the positive conditions compared with heifers housed in the reference conditions, although differences between fearful and non-fearful heifers did not significantly differ in any housing conditions.

Milk-derived measures

Means and standard errors of the milk-derived measures are detailed per personality trait and housing conditions in **Table 6**.

Milk yield. MY decreased in fearful heifers moved from the reference to the negative conditions.

FPCM. Active heifers tended to have higher FPCM than inactive heifers.

Persistency. There was a tendency for an interaction effect of housing and fearfulness on milk persistency, with larger variations in milk persistency between fearful and non-fearful heifers in the negative conditions than in the positive conditions. However, pairwise comparisons revealed no significant effect of interest. In addition, milk persistency was lower for active heifers compared with inactive heifers, regardless of the housing conditions.

LnVar. LnVar increased in fearful heifers moved from the reference to the negative conditions. This led to LnVar being higher in fearful heifers in the negative conditions compared to those in the positive conditions, as well as it being higher in fearful heifers compared to non-fearful heifers housed within the negative conditions. Finally, LnVar tended to increase in non-fearful heifers moved from the reference to the positive conditions.

pFat. pFat increased in fearful heifers moved from the reference to the negative conditions, as well as in non-fearful heifers moved from the reference to the positive conditions. The same pattern was observed for the sociability trait: pFat increased in social heifers moved from the reference to the negative conditions, as well as in non-social heifers moved from the reference to the positive conditions.

pProt. There was no evidence for significant effects of housing or personality on pProt.

pLac. There was no evidence for significant interaction effect of housing and personality on pLac. However, there was a tendency for a housing effect on pLac: pLac decreased in heifers moved from the reference to the negative conditions.

Table 6. Mean ± standard error of milk-derived measures per personality class and housing conditions. Milk yield (MY) in kg, fat and protein-corrected milk yield (FPCM) in kg, milk persistency (persistency) in kg.day^{-1}, daily milk fluctuation (LnVar) in kg, fat content in milk (pFat) in %, protein content in milk (pProt) in %, lactose content in milk (pLac) in %. Persistency was calculated from wk5-wk8 of the reference period to wk10-wk15 of the experimental period. Differences in (a,b) superscripts indicate statistical differences (p<0.10) between housing conditions for heifers of the same personality class. Differences in (x,y) superscripts indicate statistical differences (p<0.10) between personality classes for heifers housed within the same conditions.

Personality		**MY**			**FPCM**			**Persistency**	
Trait	**Class**	**Housing conditions**			**Housing conditions**			**Housing conditions**	
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**	**Negative**	**Positive**
Activity	Active	23.0 ± 0.89	24.2 ± 1.00	23.5 ± 1.63	25.7 ± 0.82	26.6 ± 0.95	26.3 ± 1.58	-0.030 ± 0.0153	-0.032 ± 0.0143
	Inactive	22.6 ± 0.92	23.2 ± 0.64	23.4 ± 0.94	25.5 ± 1.1	25.0 ± 0.60	25.5 ± 0.81	-0.009 ± 0.0135	0.001 ± 0.0109
Fearfulness	Fearful	**22.1 ± 0.90**^{a}	**24.1 ± 0.94**^{b}	24.5 ± 1.36^{ab}	25.1 ± 1.0	26.1 ± 0.93	26.5 ± 1.29	-0.039 ± 0.0194	-0.005 ± 0.0125
	Non-fearful	23.5 ± 0.86	23.3 ± 0.68	22.2 ± 0.99	26.1 ± 0.8	25.5 ± 0.61	25.1 ± 1.01	-0.003 ± 0.0059	-0.026 ± 0.0137
Sociability	Social	22.6 ± 0.82	23.2 ± 0.76	22.4 ± 1.40	25.9 ± 0.87	25.7 ± 0.72	25.1 ± 1.39	-0.016 ± 0.0111	-0.018 ± 0.0113
	Non-social	23.2 ± 1.01	24.2 ± 0.89	24.2 ± 1.14	25.2 ± 1.03	25.9 ± 0.87	26.4 ± 1.04	-0.026 ± 0.0205	-0.012 ± 0.0140

Table 6. (Continued)

Personality		**LnVar**			**pFat**		
Trait	**Class**	**Housing conditions**			**Housing conditions**		
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**
Activity	Active	0.75 ± 0.340	0.25 ± 0.119	0.81 ± 0.298	4.80 ± 0.131	4.71 ± 0.109	4.94 ± 0.15
	Inactive	0.66 ± 0.305	0.45 ± 0.162	0.23 ± 0.118	4.92 ± 0.153	4.55 ± 0.093	4.65 ± 0.149
Fearfulness	Fearful	**1.33 ± 0.350**^{a,x}	**0.55 ± 0.136**^{b}	**0.18 ± 0.137**^{b}	**4.96 ± 0.177**^{a}	**4.56 ± 0.103**^{b}	4.60 ± 0.161^{ab}
	Non-fearful	**0.14 ± 0.175**^{ab,y}	**0.15 ± 0.142**^{b}	**0.86 ± 0.274**^{a}	4.77 ± 0.097	**4.70 ± 0.100**^{b}	**4.99 ± 0.117**^{a}
Sociability	Social	0.74 ± 0.267	0.21 ± 0.139	0.47 ± 0.099	**5.03 ± 0.130**^{a}	**4.76 ± 0.100**^{b}	4.85 ± 0.119
	Non-social	0.51 ± 0.268	0.65 ± 0.418	0.50 ± 0.144	4.59 ± 0.102	**4.49 ± 0.097**^{b}	**4.73 ± 0.167**^{a}

Personality		**pProt**			**pLac**		
Trait	**Class**	**Housing conditions**			**Housing conditions**		
		Negative	**Reference**	**Positive**	**Negative**	**Reference**	**Positive**
Activity	Active	3.81 ± 0.073	3.73 ± 0.060	3.73 ± 0.067	4.43 ± 0.033	4.49 ± 0.022	4.53 ± 0.035
	Inactive	3.71 ± 0.055	3.63 ± 0.057	3.65 ± 0.062	4.45 ± 0.019	4.50 ± 0.020	4.46 ± 0.027
Fearfulness	Fearful	3.81 ± 0.087	3.69 ± 0.062	3.68 ± 0.056	4.42 ± 0.031	4.50 ± 0.021	4.51 ± 0.022
	Non-fearful	3.71 ± 0.038	3.67 ± 0.056	3.69 ± 0.077	4.45 ± 0.024	4.49 ± 0.021	4.47 ± 0.040
Sociability	Social	3.81 ± 0.067	3.74 ± 0.061	3.73 ± 0.070	4.44 ± 0.026	4.48 ± 0.021	4.46 ± 0.031
	Non-social	3.68 ± 0.050	3.62 ± 0.055	3.65 ± 0.060	4.44 ± 0.030	4.52 ± 0.020	4.52 ± 0.030

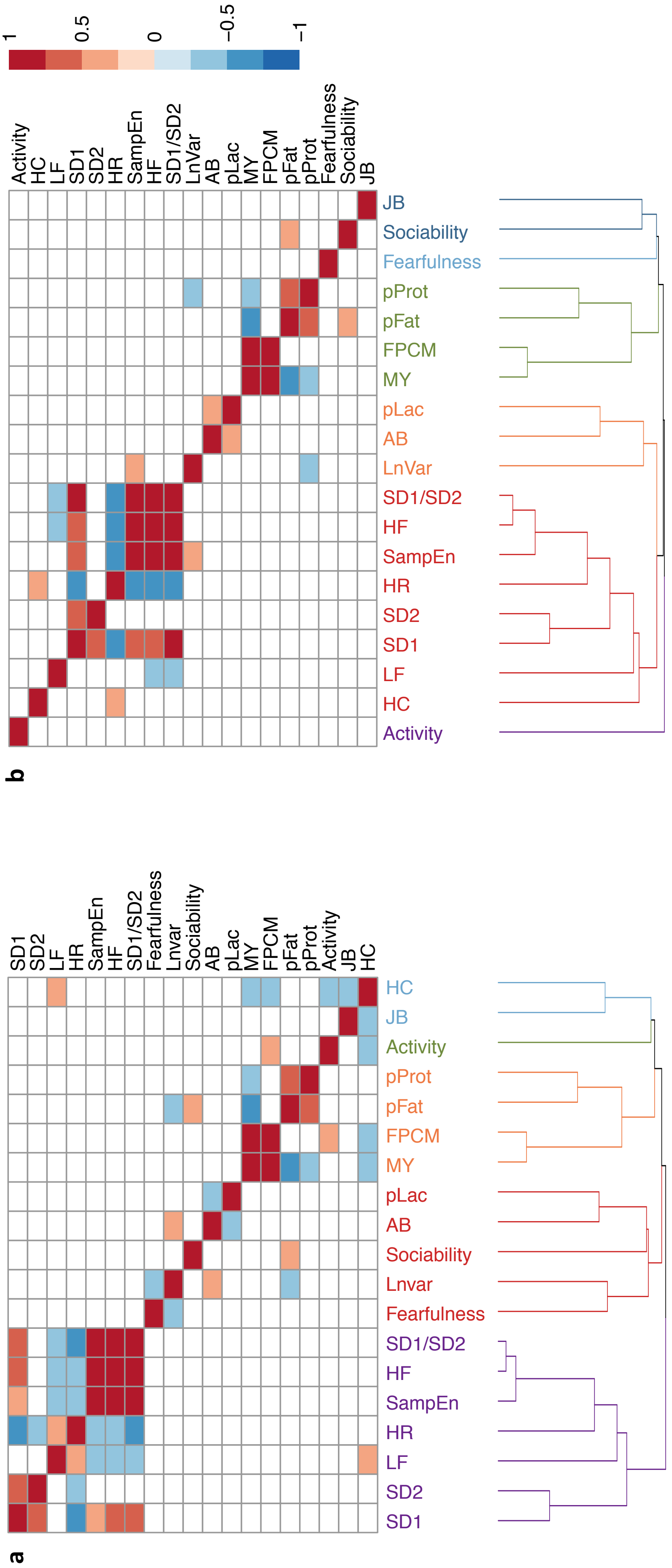
Correlation analyses

Results from the hierarchical clustering are presented in **Figure 1**. The number of identified clusters varied according to the housing conditions. When controlling for treatment differences, five and six clusters were identified when inspecting the correlation structure among 1) the raw physiological measures obtained during the reference conditions and 2) the residuals of the physiological measures obtained during the experimental conditions, respectively. In the experimental conditions, three and five clusters were identified in the positive and in the negative housing conditions, respectively. Similarities in the correlation structure of the different variables were noted between the cluster profiles of the different housing conditions: regardless of the housing conditions, measures of milk production (i.e. MY and FPCM) were positively correlated, as well as measures of milk composition (pFat and pProt). The cluster profiles of the different housing conditions remained, nonetheless, highly heterogeneous: while JB and HC were inter-correlated in both the reference and the negative housing conditions, no evidence of significant association was found between these two measures in the positive conditions. Likewise, AB and pLac were inter-correlated in the reference and the positive housing conditions, but no evidence of significant association was found in the negative conditions. Furthermore, two subgroups of HRV measures, i.e. HF/SampEn/SD1.SD2^{-1} and HR/LF, were merged within the same cluster and negatively correlated with each other in the reference and in the positive conditions, but not in the negative conditions. Finally, measures of cognitive bias were not consistently clustered with the same physiological measures or personality traits across the different housing conditions.

Pairwise Spearman correlation coefficients between the different physiological measures and the cognitive bias measures, specifically, are described according to the housing conditions in **Table 7**.

Table 7. Spearman coefficients (with their respective p-values) between the physiological measures and the cognitive bias measures, i.e. judgement and attention biases. Hair cortisol (HC) in ng.g^{-1}, heart rate (HR), low frequency band (LF) in %, high-frequency band (HF) in %, standard deviation perpendicular the line of identity in the Pointcaré plot (SD1) in ms, standard deviation along the line of identity in the Pointcaré plot (SD2) in ms, ratio between SD1 and SD2 (SD1.SD2^{-1}), sample entropy (SampEn), milk yield (MY) in kg, fat and protein-corrected milk yield (FPCM) in kg, daily milk fluctuation (LnVar) in kg, fat content in milk (pFat) in %, protein content in milk (pProt)in %, lactose content in milk (pLac) in %. Significant p-values ($p<0.05$) are written in bold.

Physiological measures	**Judgement bias**			**Attention bias**		
	Housing conditions			Housing conditions		
	Negative	Reference	Positive	Negative	Reference	Positive
Hair cortisol						
HC	**0.49 (0.028)**	**-0.43 (0.005)**	-0.32 (0.164)	-0.04 (0.879)	-0.22 (0.157)	-0.24 (0.285)
Heart rate measures						
HR	-0.18 (0.438)	-0.08 (0.607)	-0.16 (0.480)	**-0.65 (0.010)**	0.14 (0.363)	-0.19 (0.397)
LF	-0.05 (0.820)	-0.02 (0.911)	-0.16 (0.495)	-0.44 (0.100)	-0.09 (0.570)	-0.32 (0.154)
HF	0.01 (0.955)	0.10 (0.516)	-0.13 (0.577)	0.03 (0.913)	-0.05 (0.747)	0.04 (0.875)
SD1	0.02 (0.930)	0.26 (0.093)	0.001 (0.998)	-0.10 (0.724)	-0.23 (0.152)	-0.22 (0.329)
SD2	-0.15 (0.532)	0.18 (0.245)	0.14 (0.539)	0.04 (0.893)	-0.22 (0.158)	-0.31 (0.166)
SD1.SD2^{-1}	0.20 (0.395)	0.20 (0.207)	0.02 (0.930)	-0.27 (0.327)	-0.01 (0.936)	-0.03 (0.905)
SampEn	0.18 (0.446)	0.16 (0.318)	0.11 (0.625)	-0.31 (0.254)	0.07 (0.640)	-0.004 (0.986)
Milk-derived measures						
Milk yield	-0.04 (0.852)	0.13 (0.412)	0.34 (0.128)	-0.09 (0.411)	0.09 (0.561)	**0.54 (0.009)**
FPCM	-0.06 (0.806)	0.05 (0.775)	0.26 (0.257)	-0.09 (0.753)	0.20 (0.211)	0.28 (0.204)
LnVar	0.22 (0.362)	-0.08 (0.597)	-0.31 (0.176)	-0.20 (0.482)	**0.35 (0.024)**	-0.04 (0.851)
pFat	0.16 (0.507)	-0.13 (0.413)	**-0.50 (0.023)**	0.15 (0.584)	0.12 (0.464)	-0.39 (0.070)
pProt	-0.29 (0.212)	-0.001 (0.996)	-0.03 (0.899)	0.09 (0.743)	0.08 (0.636)	**-0.61 (0.003)**
pLac	0.14 (0.547)	0.02 (0.921)	-0.008 (0.975)	0.21 (0.442)	**-0.46 (0.002)**	**0.67 (0.001)**


1
0.5
0
−0.5
−1
b
Activity
HC
LF
SD1
SD2
HR
SampEn
HF
SD1/SD2
LnVar
AB
pLac
MY
FPCM
pFat
pProt
Fearfulness
Sociability
JB
a
SD1
SD2
LF
HR
SampEn
HF
SD1/SD2
Fearfulness
Lnvar
Sociability
AB
pLac
MY
FPCM
pFat
pProt
Activity
JB
HC

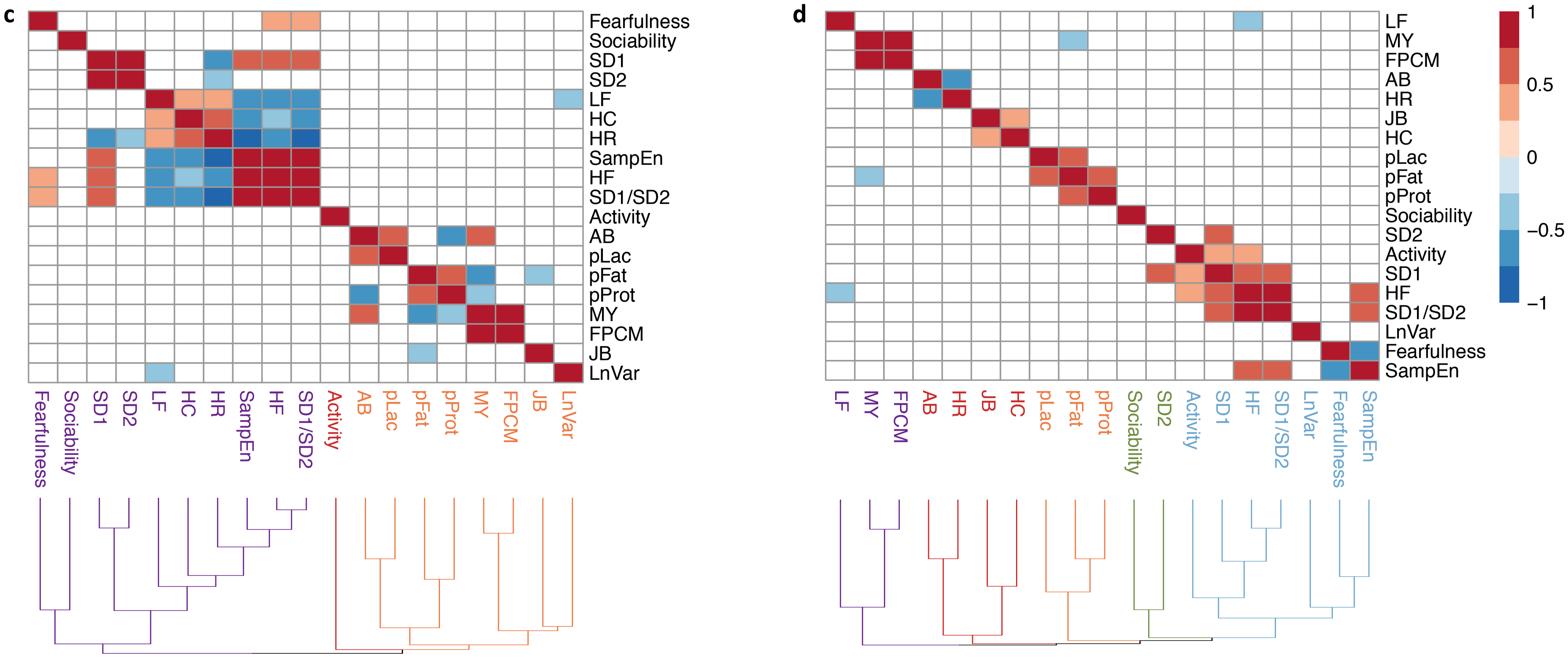


Figure 1. Heatmaps and dendrograms of the physiological measures, cognitive biases and personality scores in (**a**) the reference conditions, (**b**) the experimental conditions after controlling for treatment differences, (**c**) the positive housing conditions and (**d**) the negative housing conditions. Cells in blue (red) indicate negative (positive) correlations among the variables of the corresponding axes. Cells in white represent non-significant correlations ($p < 0.05$) between the two corresponding variables. The dendrograms indicate the degree of dissimilarity between branches: the higher the vertical lines, the less similarity between the variables. Measures of the same colour belong to the same cluster. Judgement bias (JB), attention bias (AB), activity score (Activity), fearfulness score (Fearfulness), sociability score (Sociability), heart rate (HR), sample entropy (SampEn), high-frequency band (HF), low frequency band (LF), standard deviation perpendicular the line of identity in the Pointcaré plot (SD1), standard deviation along the line of identity in the Pointcaré plot (SD2), ratio between SD1 and SD2 (SD1/SD2), milk yield (MY), fat and protein-corrected milk yield (FPCM), daily milk fluctuation (LnVar), fat content in milk (pFat), protein content in milk (pProt), lactose content in milk (pLac).

Discussion

We investigated the potential of three categories of non-invasive physiological measures as indicators of affective states in dairy heifers. To this end, we examined 1) the effect of both personality and contrasted housing conditions (i.e. as a presumed model of affective valence) on HC concentrations, HRV measures and milk-derived measures and 2) the relationships between our physiological measures and two measures of affective biases, i.e. judgement and attention biases. In addition, we explored the correlation structure among our different measures to identify common physiological traits or mechanisms underlying our measures using hierarchical cluster analyses.

Hair cortisol

The influence of the housing conditions on HC depended on personality: differences in HC between active and inactive heifers were greater in the positive conditions than in the reference and negative conditions. This result suggests that individual differences in the heifers' perception of their housing conditions were larger in the positive conditions than in the reference and negative conditions. We speculate that greater inter-individual variations exist within the domain of individual preferences, as opposed to the domain of aversion: from an evolutionary perspective, a shared perception of what is *intrinsically* negative – but not positive – may be beneficial in reducing the risk that individuals expose themselves to danger. Our finding is in contrast with previous studies conducted in cattle, which did not find an effect of personality on HC (Cooke et al. 2017; Lockwood et al. 2017). This discrepancy could arise from methodological differences in personality assessments since the aforementioned studies evaluated cow personality based on chute score and exit velocity – as is typically done in beef cattle to measure 'excitability' – while we assessed activity based on individual locomotion and proximity with the walls/floors of the arena in the context of an Open-Field and a Novel-Object tests. We can, thus, legitimately wonder whether our trait activity reflects the same personality aspect as beef cattle's excitability.

In addition, we found that HC increased in active heifers moved from the reference to the positive conditions. Although unexpected, this finding may indicate that active heifers were more stressed in the positive conditions than in the reference conditions. There are several possible explanations for this finding. First, active heifers may have suffered from the weekly perturbations made in their environment to create the positively contrasted conditions. According to Koolhaas (2008), proactive individuals (here referred to as ‘active’) may perform better under predictable than under unpredictable conditions, as they tend to develop ‘routine’ behavioural patterns. Following this theory, active heifers may have been predisposed to cope better with the stable reference conditions than with the weekly-changing positive housing conditions. Physiologically speaking, the relative inability of active heifers to cope with the positive housing conditions may have resulted in a hyper-activation of the HPA axis due to heightened stress (Ling et al. 2020). Active heifers in the negative conditions, however, were also exposed to successive, unpredictable changes and yet did not have higher HC. Alternatively, the increase in HC of active heifers in the positive conditions may reflect active heifers’ difficulty in asserting their social rank. Social rank is influenced by several individual characteristics, among which the activity trait of the individual and its age. Previous studies demonstrated that active individuals, on the one hand, and older individuals, on the other hand, are more prone to be dominant than inactive and younger individuals (Verbeek et al. 1996; Barton et al. 1973). Therefore, we speculate that active heifers in the positive conditions may have eventually suffered from their relatively low-ranking social status since they were housed with older companion cows who may have occupied the higher ranks. In contrast, active heifers in the negative conditions might have succeeded in maintaining their rank within the group because the older companion cows kept being moved between the negative pens (as part of the mixing treatment), possibly making it more difficult for them to maintain higher ranks. In line with this idea, a recent review in beef cattle revealed that an individual’s time spent within the herd may influence social dominance more than the age itself (Hubbard et al. 2021). Alternatively, and counterintuitively, this result could also suggest that active heifers were in better affective states in the positive than in the reference conditions since one study in humans revealed a positive association between HC and well-being in old women (Smyth et al. 2016). As suggested by Smyth et and colleagues (2016), it is plausible that an increase in HC reflects beneficial energy-mobilizing properties of cortisol. In accordance with this theory, elevated HC of active heifers in our positive conditions may simply reflect an increase in heifers’ physical activity, since heifers may have walked more in their home pens as they were allocated more space. Previous studies have repeatedly shown a positive association between physical activity and HC on the one hand (Gerber et al. 2012) and physical activity and enhanced mood on the other hand (Chan et al. 2019). More research is required to identify the environmental conditions and the individual characteristics for which HC may reflect either positive or negative affective valence.

Furthermore, we found no significant differences in HC between heifers housed under the negative and the positive conditions. This result contrasts with previous studies investigating the link between HC and various facets of animal housing conditions. Both beef cattle and rhesus macaques housed at higher stocking density had higher HC than counterparts housed at lower stocking density (Schubach et al. 2017; Dettmer et al. 2014). Other studies also demonstrated that pigs housed in barren conditions had higher HC than pigs housed in enriched conditions (Casal et al. 2017). In our study, the housing changes in the negative conditions may not have been abrupt enough to effectively induce stress in the heifers, who appeared to have successfully coped with them from a physiological perspective. However, the interpretation of these results requires caution since we cannot rule out the possibility that season had an effect on our measure of HC. As we opted for a longitudinal approach, the seasonal effect on HC is confounded with our treatment effect (Heimbürge et al. 2020a); which may have masked housing-induced effects on HC. Finally, it is worth noting that our HC may not accurately reflect the expected housing-induced effects due to inadequate hair sampling moments. In our study, hair samples in the experimental conditions were collected at the end of a 5-wk housing treatment, whereas a recent study revealed that the optimal time for hair sampling in cows would be within a 4-wk period *after* the end of the stressor to ensure that the section of the hair containing stress-induced cortisol has actively regrown (Heimbürge et al. 2020b).

Heart rate variability measures

The present study also investigated the effects of housing and personality on HR and HRV measures in dairy heifers. Concerning HR, we found that fearful heifers became more aroused when moved from the reference to the positive conditions, while non-fearful heifers became more aroused when moved from the reference to the negative conditions. Heifers' autonomic responses to weekly changes are, therefore, both context- and personality-dependent. Fearful heifers may have been more responsive to daily events occurring in the positive conditions, while non-fearful heifers may have been more responsive to daily events occurring specifically under the negative conditions. Whether heifers perceived these events positively or negatively remains impossible to determine on the sole basis of HR analyses. Numerous studies have demonstrated that HR can increase in response to both positive and negative affective states (Briefer et al. 2015; Reefmann et al. 2009a), but also in response to non-affectively charged events like locomotion – as a result of increased arousal. Therefore, the HR increase in both experimental conditions could arise from a more frequent exposure of individuals to both positive *and* negative experiences of high arousal, as well as changes in individual locomotion. However, considering that we measured resting heart rate (i.e. when the heifers were lying down), heightened daily exercise is unlikely to have caused increased heart rate as previous studies reported that regular training is associated with lower resting heart rate (Barnard et al. 1976). Instead, we speculate that HR increase in the positive conditions reflects heightened anxiety states in fearful heifers. Supporting this assumption, one study in humans revealed that resting heart rate tended to be higher in patients suffering from generalized anxiety compared with healthy controls (Kemp et al. 2014). Although non-significant, the relative decrease of SampEn in fearful heifers housed under the positive conditions also supports this idea since studies in humans have demonstrated that SampEn decreases when individuals are anxious (Dimitriev et al. 2016) or exposed to exam stress (Bakhchina et al. 2018). It is, therefore, plausible that fearful heifers may have suffered from the positive conditions – a theory in line with our exploratory analyses demonstrating that fearful heifers were more pessimistic than non-fearful heifers in the positive conditions (Kremer et al. 2021a). Furthermore, we hypothesise that an increase in HR in the negative conditions may be due to a disturbance in sleeping patterns as a result of limited access to cubicles. In line with this idea, a recent study in humans demonstrated that deviations from sleeping habits were associated with increased resting HR (Faust et al. 2020). Provided that this hypothesis is true, more research would, nonetheless, be necessary to determine why non-fearful heifers' cardiovascular health may be more sensitive to sleeping disturbances than that of fearful heifers. Finally, this explanation also highlights a potential limitation of our study: while we did control for the effect of heifers' posture on heart rate data by focusing on recordings taken when individuals were lying down, we did not control for the possible effects of sleep on our measurements. This may have influenced our results as HR was recently found to decrease with deepening sleeping stages in cows (Hunter et al. 2021). Future studies should therefore account for sleeping stages in addition to body postures when investigating the effects of affective treatment on heart rate.

When controlling for arousal, we found personality-dependent effects of housing conditions on HRV measures. Inactive heifers moved from the reference to the positive conditions had lower relative power of LF combined with higher relative power of HF. From a psychological perspective, this finding suggests that inactive heifers were in better affective states when housed in the positive conditions since positive affect has been associated with lower relative power of LF coupled with higher relative power of

HF in humans (Bhattacharyya et al. 2008). Non-fearful and non-social heifers also had higher relative power of HF, a result indicative of increased parasympathetic control (Shaffer and Ginsberg, 2017). In non-humans animals, higher parasympathetic activity has been repeatedly associated with positive affective states (Kowalik et al. 2017; Reefmann et al. 2009b), while decreased parasympathetic activity has been associated with negative affective states (Kovács et al. 2015; Kowalik et al. 2017). Physiologically, specific changes in the positive conditions may thus have had relaxing effects that were positively perceived by certain heifers depending on their personality. In particular, increased pro-social behaviours – as a result of enhanced social familiarity (Rault 2019) – and more frequent brushing may have led to an elevated parasympathetic activity. In line with this idea, previous studies in humans and non-human animals have demonstrated that massage, grooming and positive social connection were associated with increased parasympathetic activity (Kok et al. 2013; Grandi and Ishida 2015; Kowalik et al. 2017; Field 2019). Additional analyses of heifers' physiology and behaviour would, nonetheless, be necessary to validate our assumptions. Analyses of plasma oxytocin in dairy heifers, for instance, could be used to validate our theory since positive contacts appear to result in long-lasting oxytocin release (Rault, 2016, Faraji et al. 2018). Moreover, analyses of time-budget and activity patterns of heifers inside the home pen would help us determine the extent to which our housing changes influenced the daily routine of the heifers. We could, for instance, assess whether the allocation of extra-space in the positive conditions resulted in increased locomotion among heifers, and determine whether all heifers (e.g. fearful and non-fearful) made use of the new brush. Such analyses could help us pinpoint events or housing changes that effectively challenged heifers in a personality-dependent manner, positively or negatively. Finally, and unexpectedly, we found that active heifers in the positive conditions had lower SD1 – an index of parasympathetic HR modulation (Shaffer et al. 2014) – than active heifers in the negative conditions. Hence, active heifers in the positive conditions appeared to have an altered parasympathetic control compared with active heifers in the negative conditions. This physiological contrast may result from affective differences between the two subpopulations, since low SD1 measures have typically been associated with a wide range of psychological conditions such as depression, social anxiety disorders and post-traumatic stress disorders in humans (Kemp et al. 2012; Alvares et al. 2013; Meyer et al. 2016). Once more, we speculate that active heifers in the positive conditions may have been experiencing worse affective states than active heifers in the negative conditions, potentially due to their inability to occupy the higher social ranks within their groups.

Milk-derived measures

The present study also investigated the effects of housing and personality on milk-derived measures in dairy heifers. In contrast with HRV analyses which mostly revealed personality-based differences in heifers moved from the reference to the positive conditions, milk-derived measures mostly revealed personality-based differences between heifers moved from the reference to the positive conditions. We found that, unlike other heifers, fearful heifers showed a decrease in MY when moved to the negative conditions. While a progressive decrease in MY is expected over the lactation period (Wilmink 1987), such a drop may also be indicative of stress (Rushen et al. 2001). Therefore, we hypothesise that fearful heifers experienced the negative conditions as aversive. In line with this idea, fearful heifers in the negative conditions also had greater day-to-day milk fluctuations (as reflected by increased Lnvar) than fearful heifers in the reference or in the positive conditions. In addition, fearful heifers also had greater milk fluctuation than non-fearful heifers in the negative conditions. These results suggest that fearful heifers were less resilient when moved to the negative conditions, further suggesting that fearful heifers were the most impacted by the negative changes. These findings are in line with a previous study indicating that farm management (i.e. the environmental context) – and not only individual traits – can influence resilience in heifers (Poppe et al. 2021). The milk composition of fearful heifers also differed according to the housing conditions. The milk of fearful heifers moved to the negative conditions increased in relative fat content. This result is in contrast with previous studies, which revealed a negative relationship between cow resilience and relative fat content in milk (Poppe et al. 2021) and a drop in absolute fat content following acute stress (Hong et al. 2019). The discrepancy between our results and other findings may arise from underlying differences in terms of total MY and absolute milk composition. Based on measures of relative fat content solely, we cannot determine whether our relative fat increase originates from greater milk fat production or simply reflects a more concentrated milk (i.e. with less water). In the future, we recommend that researchers focus on absolute rather than relative measures of milk composition in order to identify whether housing conditions, and potential affective state differences, influence specific milk constituents. Relative lactose content in dairy heifers' milk was also affected by the housing conditions, whereby heifers moved from the reference to the negative conditions had a decrease in pLac. To our knowledge, only one study has investigated the effect of psychological stressors on lactose content in milk, in this case in humans, without finding any significant association between these two parameters (Ziomkiewicz et al. 2021). Research into the association between psychological factors and milk composition is still in its infancy, hence we cannot confirm or exclude the possibility that the pLac decrease observed in heifers moved to the negative conditions is linked to an increase in psychological stress. Aside from psychological causes, it is also plausible that the drop in milk pLac solely reflects a decrease in the energy balance since overstocking has been associated with increased negative energy balance in primiparous cows (Huzzey et al. 2012). This idea is, furthermore, supported by previous research demonstrating a negative relationship between pLac and β-hydroxybutyrate (Belay et al. 2017), a marker of negative energy balance in cows.

Correlation analyses

Finally, we investigated the correlation structure among personality scores, physiological measures, and cognitive bias measures across different housing contexts in dairy heifers. By carrying out a hierarchical clustering analysis, we aimed at identifying common physiological traits or mechanisms underlying our different measures. In doing so, our end goal was to determine a subset of inter-related measures that would potentially reflect previously suggested affective dimensions (i.e. arousal, valence: Mendl et al. 2010b). The cluster profiles between the different housing conditions, however, were too heterogeneous to allow for the identification of overarching physiological traits. Our goal could only have been attained if consistent associations among the same set of measures would have been found across all three housing conditions – traits being, by definition, consistent over time and across contexts (Koolhaas et al. 1999). Considering that none of the physiological measures was consistently associated with judgement or attention bias measures – either positively or negatively – across all three housing conditions, it is plausible that none of the physiological measures analysed in this study directly relates to the affective experiences of heifers. Valid physiological indicators of affect should have covaried with JB and AB, provided that these measures constitute themselves valid measures of affective states. Following this line of reasoning, we initially hypothesised that HC – a suggested yet controversial measure of chronic stress (Meyer and Novak, 2012; Schaafsma et al. 2021) – would be positively correlated with JB whichever the housing context. However, we found that HC was negatively correlated with JB under the reference conditions, while it was positively correlated with JB under the negative conditions. More pessimistic heifers thus had lower HC in the reference conditions, while they had higher cortisol in the negative conditions. While such a finding may indicate that HC does not constitute a valid indicator of affective states, it may also indicate that HC increases are emotion-specific rather than valence-specific. In line with this idea, previous studies in humans have demonstrated that both hypo- and hypercortisolism could be linked to negative affective states, i.e. with anxiety and depression, respectively (Steudte et al. 2011; Dettenborn et al. 2012). Therefore, we hypothesise that the inconsistent associations between JB and HC arose from differences in heifers' affective states – pessimistic heifers being predominantly anxious in the reference conditions while being depressed in the negative conditions. Associations between HC and other physiological measures which are specific to the positive housing conditions were also noted. In the positive conditions, HC concentrations were positively correlated with the relative power of LF and HR, while all three measures were also negatively correlated with SampEn, HF and SD1/SD2. Therefore, we speculate that HC and the frequency-domain measures of HRV reflected the same primary physiological mechanism when heifers were housed in the positive housing conditions, i.e. when they were in supposedly better affective states. In contrast, we hypothesise that HC and the frequency-domain measures of HRV were influenced by additional and distinct physiological mechanisms when heifers were housed under the reference or the negative conditions, i.e. when they were in supposedly worse affective states. If valid, such a theory would explain why no significant monotonic relationship between HC and the aforementioned HRV measures was found in the reference or in the negative housing conditions. In addition, mixed associations between AB and pLac were identified across housing contexts: heifers who looked at the threat later had lower pLac in the reference conditions, but higher pLac in the positive conditions. This inconsistent finding may be due to certain limitations of our study. It could be that AB did not reflect heifers' affective states in both the reference and the positive conditions as a result of ABT's potential lack of repeatability. In our experiment, each heifer was exposed twice to the ABT, although the effects of prior exposure to ABT on subsequent responses to the test are largely unknown (Kremer et al. 2021a). Factors other than affective states may, thus, have influenced heifers' responses to the ABT in the positive conditions – thereby altering the relationship observed between AB and pLac under the reference conditions. Additional studies examining the validity and the repeatability of the ABT as a suitable, and repeatable assessment tool for affective states appear necessary before researchers can identify more practical indicators of affective states based on their associations with AB. Alternatively, the significant relationships found between AB and pLac (as well as any other significant relationships identified here) may result from a type I error since our correlation analyses were conducted on a rather limited number of animals (Knudson and Lindsey 2014). Correcting for multiple correlations would have, nonetheless, been excessively conservative considering the exploratory nature of our study.

Conclusion

In conclusion, we found that contrasted housing conditions, which putatively induced a difference in heifers' affective states, had personality-dependent effects on frequency-domain measures of heart rate variability and on daily milk fluctuation. Inactive heifers, non-fearful heifers and non-social heifers moved from the reference to the positive conditions had higher relative power of the high-frequency band, suggesting that the high-frequency band may constitute an adequate indicator of positive shifts in the affective states of dairy heifers. In addition, fearful heifers moved from the reference to the negative conditions had greater milk fluctuation, suggesting that milk fluctuation could be a valid indicator of negative shifts in the affective states of dairy heifers. Although useful to detect similarities between various measures, our hierarchical clustering analysis did not allow for the identification of common mechanisms between our physiological measures and two suggested indicators of affective states, i.e. judgement and attention biases, across the different housing contexts. More research is, therefore, required to further validate our physiological measures as indicators of affective states.

Supplementary materials

Supplementary Table 1. Detailed treatment applied every Friday during the experimental conditions to design the positive and negative housing conditions (from Kremer et al. 2021a).

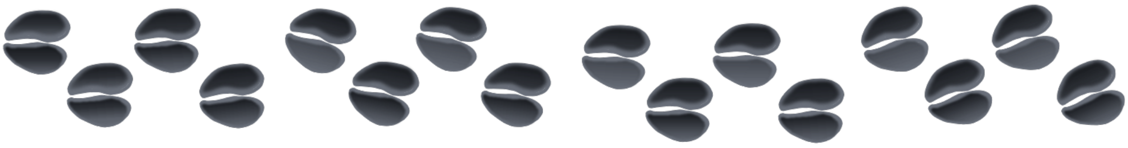
Week	Levers of actions	Positive housing	Negative housing
Wk10	Crowding conditions	Increase space allowance Open 1 cubicle and 1 feeding gate	Decrease space allowance Close 2 cubicles and 2 feeding gates
	Social stability	Add feeding partitions	
Wk11	Enrichment	Add a fixed brush or replace a fixed brush by a rotating one	Remove the fixed brush
	Social stability	Keep stable groups	Mix two companion animals
Wk12	Crowding conditions	Open 1 cubicle and 1 feeding gate	Close 2 cubicles and 2 feeding gates
Wk13	Enrichment	Add a fixed brush or replace a fixed brush by a rotating one	Switch to another home pen
	Social stability	Keep stable groups	Mix two companion animals
Wk14	Crowding conditions	Open 1 cubicle and 1 feeding gate	Close 1 cubicle and 1 feeding gate
	Social stability	Keep stable groups	Mix two companion animals

Supplementary Table 2. Number of heifers in each housing conditions per personality trait. Heifers were divided in two classes per personality trait based on their behavioural scores on the related personality trait in comparison to the median score. Superior (Sup.), Inferior (Inf.), Rotated component (RC), n number of heifers per personality class in the reference conditions, n_1 number of heifers per personality class in the positive conditions, n_2 number of heifers per personality class in the negative conditions (from Kremer et al. 2021a).

Personality			**Housing conditions (number)**		
Trait (median score)	**Class**	**Definition**	**Reference ($n = n_1+n_2$)**	**Positive (n_1)**	**Negative (n_2)**
RC1: Activity (-0.08)	Active	Sup. to -0.08	23	11	12
	Inactive	Inf. or equal to -0.08	24	13	11
RC2: Fearfulness (0.23)	Fearful	Inf. or equal to 0.23	24	13	11
	Non-fearful	Sup. to 0.23	23	11	12
RC3: Sociability (0.26)	Social	Sup. to 0.26	24	10	14
	Non-social	Inf. or equal to 0.26	23	14	9

Chapter 6

General discussion

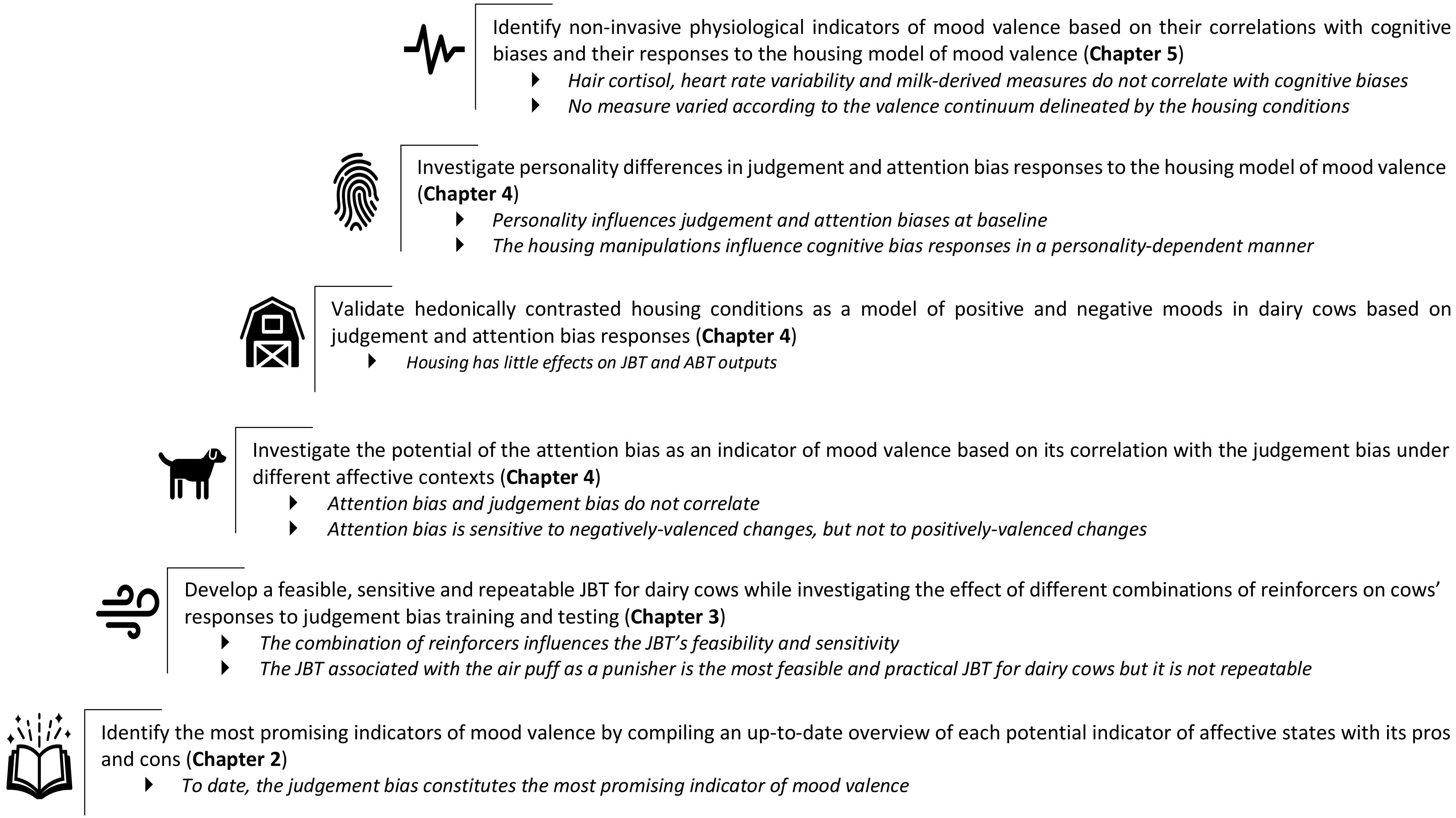


Introduction

Valid assessment tools of dairy cow welfare are necessary to ensure that cow welfare is optimal and adequality monitored. Proper assessment of dairy cow welfare must now integrate an adequate evaluation of affective states (Watanabe, 2007), since welfare is considered optimal when the balance between positive and negative affective states is positive (Green and Mellor, 2011). The final goal this PhD thesis was to identify valid and practical physiological indicators of long-term affective states (i.e. of mood) in dairy cows, to come closer to providing farmers with a tool allowing them to monitor the affective valence of the animals in their care. To meet this goal, the stepwise approach introduced in Chapter 1 was followed. The major findings to emerge from this project are summarized in **Figure 1**. In this final chapter, I will discuss these findings in a larger context focusing on (i) the use of cognitive biases as gold standard indicators of mood valence, (ii) the development of effective models of mood valence, and (iii) the arduous quest of identifying an indicator of valence, which combines all the expected theoretical and practical requirements.

Cognitive biases as gold-standard indicators of mood valence

In this PhD study, the validity of the cognitive bias measures was a prerequisite to the development of an adequate model of mood valence in dairy cows (**Chapter 4**), and to the identification of practical indicators of mood valence (**Chapter 5**). The idea that judgement bias, in particular, would constitute a valid indicator of mood valence was based on the numerous studies demonstrating that the manipulation of mood biases one’s judgement of ambiguous events, i.e. with positive treatments inducing more optimistic responses in animals than negative treatments (Harding et al. 2004; Douglas et al. 2012; Lagisz et al. 2020; Neville et al. 2020). Attention bias, which can be assessed faster than judgement bias, was also selected as a potential reference measure of mood valence since it has previously been validated as an indicator of anxiety states in both sheep and beef cattle (Lee et al. 2017; Monk et al. 2020). Both methods of cognitive bias assessment are still, nonetheless, fraught with several limitations that have been highlighted in this thesis (**Chapter 3, Chapter 4, Chapter 5**) and will be further discussed here.



Identify non-invasive physiological indicators of mood valence based on their correlations with cognitive biases and their responses to the housing model of mood valence (**Chapter 5**)

- *Hair cortisol, heart rate variability and milk-derived measures do not correlate with cognitive biases*
- *No measure varied according to the valence continuum delineated by the housing conditions*

Investigate personality differences in judgement and attention bias responses to the housing model of mood valence (**Chapter 4**)

- *Personality influences judgement and attention biases at baseline*
- *The housing manipulations influence cognitive bias responses in a personality-dependent manner*

Validate hedonically contrasted housing conditions as a model of positive and negative moods in dairy cows based on judgement and attention bias responses (**Chapter 4**)

- *Housing has little effects on JBT and ABT outputs*

Investigate the potential of the attention bias as an indicator of mood valence based on its correlation with the judgement bias under different affective contexts (**Chapter 4**)

- *Attention bias and judgement bias do not correlate*
- *Attention bias is sensitive to negatively-valenced changes, but not to positively-valenced changes*

Develop a feasible, sensitive and repeatable JBT for dairy cows while investigating the effect of different combinations of reinforcers on cows' responses to judgement bias training and testing (**Chapter 3**)

- *The combination of reinforcers influences the JBT's feasibility and sensitivity*
- *The JBT associated with the air puff as a punisher is the most feasible and practical JBT for dairy cows but it is not repeatable*

Identify the most promising indicators of mood valence by compiling an up-to-date overview of each potential indicator of affective states with its pros and cons (**Chapter 2**)

- *To date, the judgement bias constitutes the most promising indicator of mood valence*

Figure 1. Objectives and major findings of this PhD study following the stepwise approach introduced in Chapter 1. The objectives are written in roman style, while the results are specified in italic. JBT: Judgement Bias Task. ABT: Attention Bias Task.

Methodological considerations

As seen in **Chapter 3** and **Chapter 4**, the validity of both the Judgement Bias Task (JBT) and the Attention Bias Task (ABT) in providing an assessment of mood is dependent upon the set-up of the tasks themselves. Several methodological aspects must, hence, be carefully considered before implementing these two tasks.

The specific choice of the stimuli and of the reinforcers used during the cognitive bias tasks (CBTs) are of high relevance. Consistent with previous theories (Mendl et al. 2009; Roelofs et al. 2016), the combination of the two reinforcers – in this case, the reward and the punisher – was shown to influence the JBT's baseline sensitivity to assessing affective states, be it mood or emotion (**Chapter 3**). The term baseline sensitivity refers to the sensitivity of the task when animals' affective states are not yet subjected to voluntarily manipulation. In an ideal JBT, the response pattern of latencies across the different cues (ordered according to their valence from the positive cue P to the negative cue N) should be linear at baseline to maximise the task's ability to detect both positive and negative treatment-induced shifts in mood. In our experiment, all cows were exposed to the same positive reinforcer (i.e. 150 g of concentrates), but the type of negative reinforcers varied among the treatment groups. Cows exposed to an electric shock or to a 'no-reward' as punishers were found to exhibit negatively biased baselines, whereas cows exposed to an air puff were found to exhibit a positively biased baseline. It was, thus, hypothesised that the effect of the reinforcer combination on JBT's baseline sensitivity is determined by the imbalance in affective states elicited by both reinforcers (**Chapter 3**). To validate this hypothesis, future studies should examine whether supposedly contrasted punishers in terms of affective valence effectively induce differences in animal affective responses – by investigating individual arousal during the exposure to the punisher. Provided that the selected punishers effectively elicit avoidance behaviours in animals, and are effectively perceived as negative (Elliot et al. 2013), punishers of greater aversiveness will induce greater arousal than punishers of lower aversiveness as a result of the V-shaped relationship between arousal and valence (Kuppens et al. 2013).

Besides altering the JBT's sensitivity, we have also demonstrated that the combination of reinforcers may influence the feasibility of the task – an aspect of the JBT that has been heavily criticised (Roelofs et al. 2016; Hintze et al. 2018). In contrast with the learning goals inherent to Go/NoGo paradigms, cows exposed to the electric shock (i.e. the supposedly most aversive punisher) were the least likely to reach the positive cue to receive the concentrates, while cows exposed to the ‘no-reward’ punisher (i.e. the supposedly least aversive punisher) were the most likely to reach the negative cue (**Chapter 3**). Learning to display a NoGo response to N, therefore, appears to be altered when using a punisher where the aversiveness outweighs the rewarding properties of the positive reinforcer; whereas learning to display a Go response to P is affected when using a positive reinforcer where the rewarding properties outweighs the aversiveness of the punisher. By dedicating more research and effort into the selection of appropriate reinforcers, researchers may develop JBTs where sensitivity to affective states would be enhanced (thereby limiting the risk of type II error) and where feasibility would be improved.

In the ABT, the choice of the stimuli also matters, in particular the choice of the stimuli used to represent the threat since the ABT’s principle relies on the idea that anxious animals look at the threat faster than less anxious animals (Lee et al. 2016, 2017). In the ABT designed specifically for dairy cows (**Chapter 4**), the threat consisted of a dog model made from an association of visual, olfactory and auditory cues (i.e. a dog statue combined with dog urinary smell and a growling sound). The choice of using such a threat was based on previous studies conducted in ruminants, which successfully validated their ABT set-ups as tools for the assessments of anxiety states when using a live dog as a threat (e.g. Lee et al. 2017) and demonstrated the influence of visual, olfactory and auditory models of canine predators on fear, vigilance and feeding behaviours (Pfister et al. 1990; Terlouw et al. 1998; Kluever et al. 2009). In our study, however, the fact that heifers effectively experienced the dog model as a threat was not ensured prior to the test. Thus, it is not certain that the behavioural responses observed in the ABT adequately reflected the influence of mood (and in particular of anxiety states) on heifers’ allocation of attention toward one cue over another. Instead, the differences in heifers’ behavioural responses during the test may have simply reflected personality differences in terms of neophobia since heifers were unfamiliar with the dog model when exposed to the first attention bias test. This idea is in line with one study demonstrating that a simulated predator encounter elicited fear responses in predator-experienced cows but not in predator-naïve cows (Cooke et al. 2013). Before implementing our ABT paradigm in future studies, it should thus be determined whether the dog model is effectively perceived as threatening. One strategy to achieve this goal would be to pharmacologically validate the dog model as a threat following a strategy similar to that applied in sheep by Monk et al. (2018b) and using a control group where a more neutral stimulus, looking like the dog model (e.g. a lamb model) would be used instead of the dog model.

The repeatability of both CBTs must also be ensured, especially within the context of longitudinal studies, the principle of which relies on testing repetitions. In the second experiment, the goal was to validate the housing contrasts as an appropriate model of mood valence by investigating housing-induced changes in cognitive biases at the individual level (**Chapter 4**). Because personality can modulate one's subjective experience of events (Roelofs et al. 2016) and can influence individual responses to the cognitive bias tests (Lecorps et al. 2018; Cussen and Mench 2014), controlling for personality was necessary to properly assess the validity of our housing contrasts as a model of valence. Following the same approach as Neave et al. (2013) and Daros et al. (2014) in calves, we, hence, opted for a longitudinal study where heifers were exposed twice to each cognitive bias test: at baseline, when heifers were housed under the reference conditions, and at the end of the experimental conditions, when heifers were housed under either supposedly more negative or more positive conditions. Our experimental design was based on the postulate that both CBTs were repeatable, meaning that the heifers' experience of the first cognitive bias tests did not alter, in any way, individual responses to the second cognitive bias tests. In a previous experiment, the repeatability of our JBT was assessed based on cows' responses to the task over two testing periods (**Chapter 3**). At that time, each testing period was composed of two distinct testing sessions, during which all cows were exposed to three consecutive ambiguous cues (i.e. six per testing period). Both testing periods were, furthermore, separated by a wash-out period of two weeks, which consisted of three training sessions and aimed at reducing the likelihood of cows' remembering their first encounter with the ambiguous cues (Doyle et al. 2010b). Considering that these conditions were not sufficient to ensure the repeatability of the JBT in our first experiment (**Chapter 3**), it was decided in the second experiment to 1) reduce the individual's exposure to three ambiguous cues per testing period and to 2) increase the duration of the wash-out period to 5 weeks (**Chapter 4**). Whether these adjustments were effective remains, however, purely speculative since the effects of housing and testing sessions were confounded in the experiment. Likewise, it cannot be confirmed whether our ABT was repeatable either. Although one study in macaques showed that a week interval between two testing periods was sufficient to ensure the ABT's repeatability (Howarth et al. 2021), heifers spent more time eating and focused their attention more toward the positive cue during the second attention bias test compared with the first, regardless of the treatment (**Chapter 4**). Based on these results, it can be hypothesised that heifers were less scared of the dog model during the second compared to the first attention bias test, potentially due to the fact that no negative event resulted from their initial encounter with the dog model. Hence, the possibility that heifers' affective experience of the cognitive bias tests during the first testing session influenced their responses during the second testing session, potentially due to a ‘first impression’ bias, cannot be excluded. Future studies should, therefore, investigate whether animals’ first impression of a task influences how animals perceive – and thus react to – this task when re-exposed to it a second time around.

The significance of personality

One major finding of this PhD work is the importance of personality on responses to the cognitive bias tests, in particular at baseline level (**Chapter 4**). To embrace the complexity of personality in heifers (**Chapter 4**), we did not restrict ourselves to the study of personality traits, but, instead, we focused on personality profile. Here, the term personality profile refers to the two or three-dimensional space delineated by a combination of two or three personality traits, respectively. This decision was based on the fact that cattle personality is pluri-dimensional (van Reenen et al. 2005; Graunke et al. 2013; Koolhaas and van Reenen 2016; Lecorps et al. 2018). In **Chapter 4**, at least three personality traits were identified and labelled as *Activity*, *Fearfulnes*s and *Sociability*. Each of these personality traits, and some of their respective combinations (i.e. personality profile), were found to influence heifers’ responses to the judgement and attention bias tests (**Chapter 4**). For instance, *inactive–fearful* heifers were slower to reach the ambiguous cues, and were hence characterised as more pessimistic, than *inactive-non-fearful* heifers. Similarly, heifers that were both *non-fearful* and *non-social* spent more time eating in the ABT than *fearful-non-social*, *fearful-social* and *non-fearful-social* heifers. These results have several implications, which should not be overlooked in future studies.

First, these findings suggest that accounting for single personality traits when investigating cognitive biases is not sufficient to properly assess treatment-induced mood shifts, at least in cattle. When balancing treatment groups based on single personality traits, researchers run the risk of having an imbalance in personality profile which may lead to the over-representation of a certain personality profile in one group compared with the others. Provided that this over-represented personality profile (e.g. *inactive-fearful*) is predisposed to certain cognitive biases (e.g. pessimistic judgement), the effects of the treatment and of personality profile on cognitive biases may be confounded – potentially leading to erroneous interpretations and conclusions. In practice, such mistakes should be avoided as they hamper progress in the field of animal welfare and may have dramatic consequences in practice with, for instance, the implementation on-farm of aversive treatments that are thought to be pleasurable to the animals. Although laborious as it requires the assessment of personality for each individual animal before the start of the experiment, researchers should, thus, balance personality profiles between treatment groups to avoid erroneous interpretations, especially within the context of cross-sectional studies. Alternatively, researchers may opt for longitudinal studies to account for personality differences between treatment groups, provided that their CBTs are repeatable.

Second, these findings highlight the possibility that the CBTs *themselves* are triggering differences in affective states. This hypothesis was already formulated in **Chapters 2 and 3** as a potential explanation for the differences observed in responses to the ambiguous cues between cows exposed to JBTs differing in terms of punisher – cows exposed to an electric shock as a punisher being more pessimistic than cows exposed to an air puff or to ‘no-reward’. Provided that cows effectively perceived the electric shock more negatively than the air puff and the ‘no-reward’, it is likely that the perception of the task itself influenced cows’ transient mood and their subsequent responses to the test. This idea is, moreover, in line with a previous study demonstrating that an affective mismatch between the procedure of JBT testing and the intended mood manipulation can alter (and even reverse) the expected JBT outcomes (Raoult et al. 2017). Likewise, differences in heifers’ perception of the JBT and ABT set-ups in **Chapter 4** may explain the personality-based differences in responses to the cognitive bias tests. Because *fearful* individuals are more prone to neophobia (Boissy and Bouissou 1995), it is, for instance, likely that *inactive-fearful* heifers were more scared of the ambiguous, and intrinsically novel, cues compared with *active-non-fearful* heifers. Likewise, *fearful* heifers may have spent more time looking at the threat in the first ABT compared with *non-fearful* heifers because they were more scared of the unknown dog model. The set-up of the cognitive bias tests itself may, therefore, influence animals’ ongoing mood and bias animals’ responses to the tests. As far as possible, researchers should aim at developing CBTs that minimise personality-based differences in responses to the tests as a way to optimise population validity.

Interpretative limitations

Although JBT and ABT constitute unique and promising tools allowing inferences into animal affective experiences (Lagisz et al. 2020; Crump et al. 2018), the interpretation of animal responses to these CBTs in terms of valence is sometimes arduous.

First, the interpretation of CBT outputs with regard to mood valence is challenged by the fact that cognitive biases can be modulated by processes that are affect-unrelated. While it is plausible that personality-based differences in CBT outputs reflect variations in animal affective states due to differences 1) in baseline affective states (Watson and Tellegen 1985; Winter and Kuiper 1997) or 2) in the perception of the tasks themselves, differences in CBT outputs may also simply reflect variations in behavioural predispositions. For instance, heifers characterised as *active* are, by definition, more likely to walk than *inactive* heifers, whatever the context and affective states. Thus, the fact that *active* heifers spent more time in locomotion than *inactive* heifers in response to the first ABT (**Chapter 4**) does not necessarily indicate that *active* heifers were more anxious (Lee et al. 2017; Campbell et al. 2019), but may instead reflect their natural predispositions for walking. Researchers should, therefore, refrain from interpreting any behavioural differences in response to the CBTs as evidence for affective states differences; and combine the results obtained from different measures of cognitive bias to make inferences about animal transient mood (**Chapter 2**).

Second, our insufficient knowledge on the nature of the affective states (e.g. emotion or mood) that effectively influences animals' cognitive biases also limit our interpretation of cognitive bias shifts. As discussed in **Chapter 2**, judgement bias appears to reflect a combination of both emotion and mood, since it is influenced by short-term events (Sanger et al. 2011; Neave et al. 2013) as well as longer-term environmental manipulations (Harding et al. 2004; Douglas et al. 2012). The same holds for attention bias which seems to be modulated both by acute procedures (Bethell et al. 2016) and longer-term manipulations (Brilot and Bateson 2012; Raoult and Gygax 2019). In any experiments, the effect of treatment-induced affective states on cognitive biases may thus be overruled by a wide range of uncontrolled and unintended short-term emotions that impede the ability of a research study to adequality assess the effect of a treatment on animal affective states (Luo et al. 2019; Verbeek et al. 2019). Thus, it can be speculated that – despite our effort to standardise the cognitive bias tests – heifers in the positive housing conditions may have experienced negative events (e.g. agonistic interactions in the home pen) shortly before the test. Such events may have buffered the effect of housing-induced positive mood on judgement biases, thereby preventing us from detecting positive judgement bias in response to the positive housing conditions (**Chapter 4**). These potential confounding effects could have possibly been minimised by enrolling a larger number of heifers in the experiment. Following the same line of reasoning, the fact that heifers looked at the threat later when housed under the negative conditions compared with the reference conditions could either indicate that 1) heifers were in a more positive mood when housed under the negative conditions or 2) heifers were temporarily relieved from exiting their home pen when housed under the negative conditions, therefore suggesting that they were in a more negative mood when housed under the negative conditions (**Chapter 4**). Regrettably, several plausible, yet contradictory, interpretations of the same cognitive bias results are therefore possible. This latter notion ultimately questions the reliability of the conclusions drawn by researchers when assessing affective states from cognitive biases only, and warrants caution against the use of a single indicator to describe affective states in animals (**Chapter 2**).

Third, the interpretation of attention biases in terms of general affective valence may be inappropriate. In Chapter 4, no significant correlation was found between judgement and attention biases. This result suggests that both cognitive biases are mediated by distinct mechanisms. As postulated by Crump and colleagues (2018), while the judgement bias may constitute a valid indicator of valence, attention biases may only be indicative of specific affective states (i.e. anxiety) – the nature of which remains to be determined.

Building-up models of mood valence

The ability to generate, and *a fortiori* to detect, positive and negative moods in animals is a prerequisite to the identification of reliable indicators of mood, in particular of affective valence. Thorough consideration must, hence, be given to the conception and the development of valid models of mood, which should elicit the intended affective responses in the population of focus. As introduced in Chapter 1, one major aim of this PhD work was to establish a model paradigm able to induce mood shifts in dairy heifers based on a combination of various housing changes.

The validity of our model paradigm was investigated based on heifers' responses to the cognitive bias tests (**Chapter 4**). The model paradigm was expected to effectively induce shifts in heifers' judgement and attention biases – provided that the latter were valid indicators of mood valence. Thus, it was hypothesised that heifers moved from the reference to the positive conditions would become more optimistic and would look at the threat later and for a shorter duration during the ABT, and vice versa for heifers moved from the reference to the negative conditions. Our housing contrasts, however, yielded mixed results in terms of cognitive bias responses. Heifers did not become more optimistic and did not look at the threat sooner when moved from the reference to the positive conditions. They did, however, spend less time looking at the threat compared to heifers housed in the reference conditions – but not compared to heifers housed in the negative conditions. Finally, heifers did not become more pessimistic and did not look at the threat later when housed from the reference to the negative conditions. Contrary to expectations, heifers even looked at the threat later when moved from the reference to the negative conditions. Such unexpected results inevitably challenge the validity of our model paradigm and question the reliability of our validation method – two aspects which will be further discussed here.

The validity of our model paradigm

The development of a model paradigm efficient at inducing mood shifts was challenged by the absence of validated theories of mood induction (Eldar et al. 2016; Raoult et al. 2017). As introduced in Chapter 1, our model paradigm consisted of a combination of housing changes, which were successively applied on a weekly basis (**Chapters 4 and 5**). The model paradigm was based on previous studies suggesting that applying a combination of various stimuli was effective at inducing mood (Westermann et al. 1996; Mendl et al. 2009). However, although congruent with other research conducted in the field of animal welfare that claims to have effectively induced mood shifts (Destrez et al. 2013, 2014; Bethell and Koyama 2015), our housing model to induce mood shifts could have possibly failed since prior knowledge on how to develop model paradigms effective at inducing mood is currently lacking (Raoult et al. 2017). To date, research efforts have mainly focused on the elaboration and the validation of models of moods disorders (e.g. depression-like states: Novak et al. 2016) at the expense of research on adaptive mood processes – be it positive or negative. Following Raoult et al. (2017), who investigated the validity of cumulative expectation mismatch as a model of mood, future research should, therefore, aim at conceptualising and validating methodological recipes for the induction of (adaptive) mood states.

Our model paradigm may also have failed at inducing the expected mood shifts in heifers due to an inadequate application of the housing manipulations (**Chapters 4 and 5**). Heifers housed under the supposedly positive conditions were, for instance, exposed to unfamiliar feeding partitions and a new fixed brush for a maximum period of five and four weeks (35 and 28 d), respectively. However, one study published in nutcrackers suggested that enrichment may be perceived as stressful depending on its duration of exposure – as nutcrackers showed greater physiological stress following a 31d-period of enrichment while they showed lower physiological stress following a 67d-period of enrichment (Fairhurst et al. 2011). Therefore, it could be speculated that the duration of heifers' exposure to our structural enrichment was not sufficient to already be experienced as beneficial. This idea, however, can be questioned as heifers also had the opportunity to learn from their pen mates that the enrichment objects were harmless – or even interesting (e.g. brush) – provided that cows are able of observational learning (Rørvang and Nawroth 2021). Further research is warranted to investigate the optimal application, in terms of type, frequency, duration and intensity, of species-specific stimuli for the induction of the intended mood states.

The validity of our model paradigm may also be limited to certain animal models of specific characteristics. Only heifers with a particular personality may have been sensitive to the mood treatment, at least in the expected direction. In the second experiment, the same mood treatment was applied to all heifers – regardless of their personality. Our results, nonetheless, suggest that individual preferences and aversiveness for certain stimuli are personality-dependent (**Chapters 4**). The relative lack of consistency in pessimism between heifers housed in the reference and the positive conditions may, for instance, indicate that our positive treatment induced different moods of contrasted valence in heifers with different personalities, an assumption supported by the fact that *fearful* heifers appeared more pessimistic than *non-fearful* heifers in the positive conditions (**Chapter 4**). Similarly, the fact that *non-social* heifers looked at the threat later than *social* heifers in the negative conditions suggests that personality modulated heifers' perception of the negative housing conditions. However, considering the multifactorial nature of our mood treatment, the specific housing changes behind these personality-driven mood divergences cannot be identified with certainty. It is, nonetheless, reasonable to assume that housing facets congruent with heifers' personality traits are likely to exacerbate mood differences. As proposed in Chapter 4, *fearful* heifers may have experienced the repeated introductions of enrichment within their home pen more negatively than *non-fearful* heifers as a result of differences in neophobia, since *Fearfulness* was assessed based on individual latency to reach a novel object. This idea is in line with a study conducted in parrots which demonstrated that structural enrichment increased fear-related behaviours among the most fearful individuals while reducing them in less fearful individuals (Fox and Millam 2007). Considering that neophobia is a widespread phenomenon across taxa and species

(Crane and Ferrari 2017), it is likely that personality-based differences in perception of new enrichment exist both in birds and cattle – although further studies are warranted to confirm this theory. Likewise, it is plausible that *social* and *non-social* heifers had contrasted experiences of the social manipulations performed in the negative housing conditions, such as the frequent regrouping (Nogues et al. 2020). Based on these considerations, the existence of a 'one size fits all' housing environment, which would be equally liked or disliked by all animals, seems improbable (Coleman 2012). When focusing on non-pharmacological affective treatments, I, therefore, encourage researchers to first investigate personality differences in individual perception of their treatment, before including personality as an input to their affective model (i.e. Personality x Treatment) in order to elicit the intended affective states.

The state of the focal heifers in the study may also have altered the validity of the model paradigm. The limited evidence for housing-induced cognitive biases (**Chapter 4**) may indicate that the model paradigm was inefficient at inducing the expected mood shifts in heifers, who might have had the unforeseen ability to cope with the successive housing changes. As mentioned in Chapter 1, we decided to focus on heifers (here primiparous cows) in the second experiment in an attempt to standardise individual background. However, numerous studies have shown that cognitive and behavioural flexibility decline over age in animals – with older individuals demonstrating increased difficulty at set-shifting, learning speed, information gathering and reverse learning than younger ones (Manrique and Call 2015; Wallis et al. 2016; Lacreuse et al. 2018; Mazza et al. 2018). Heifers' ability to adapt to successive changes may thus have exceeded our expectations, which allowed them to successfully cope with our housing manipulations. To validate this theory, researchers should determine whether and at what age cognitive decline may occur in dairy cows.

The pregnancy status of our animal model may also have overridden the potential effects of the housing contrasts on heifers' mood. Considering that mood disorders and mood variations are prevalent during pregnancy in humans (Bowen et al. 2012), it is reasonable to assume that similar affective dysregulations may also occur in other mammals – potentially owing to common hormonal fluctuations (Kindahl et al. 2002; Bloch et al. 2003). Provided that this theory holds, the effects of pregnancy and housing changes on heifers' mood may thus have been confounded. On days surrounding the judgement bias test, certain heifers in the positive conditions may, for instance, have been more depressed or anxious as a result of pregnancy-related mood variation than they would have been without being pregnant. So far, however, the psychological impact of pregnancy on cows has largely been ignored; and additional research is required to investigate the link between physiological changes during pregnancy and mood variations in dairy cows.

Lastly, heifers' initial psychological state may have buffered the effect of our housing contrasts on heifers' mood. Considering that dairy cows are exposed to various stressors in their everyday life (e.g. restrained pasture access: Crump et al. 2021, regrouping: Schirmann et al. 2011, veterinarian procedures: Cingi et al. 2012), the possibility that our heifers were already depressed from the start of our experiment cannot be excluded, as chronic stress contributes to the psychopathology of depression (Ross et al. 2017). If heifers were effectively initially depressed, their responsiveness to our mood treatment might have been limited, potentially due to apathy – which is a possible symptom of depression (Marin et al. 1994). Additionally, heifers in the positive conditions may have lacked reactivity to pleasurable stimuli as a result of anhedonia, a core feature of depression (Pizzagalli et al. 2008). Finally, our negative treatment may have been ineffective at further deteriorating heifers' mood, depressive states already being at the negative end of the valence continuum (Mendl et al. 2010b). In practice, it could have been determined whether heifers initially suffered from apathy or anhedonia by investing their behavioural responses to different stimuli. Following the line of reasoning of Meagher and Mason (2012), apathetic individuals would have exhibited a decreased interest in all kinds of stimuli (from negative to positive), while anhedonic individuals would have demonstrated a reduced interest specifically to positive stimuli. In a broader context, combining behavioural assessments with cognitive evaluations of mood valence would have also strengthened the reliability of our validation method of the housing model.

The reliability of our validation method

Rather than challenging the validity of our model paradigm, these unexpected findings may also question the reliability of our validation method. As discussed in the first part of this chapter, the validity of the model paradigm was assessed based on measures of judgement and attention biases, for which the methods of assessments are not free of limitations. These limitations, notably the potential lack of repeatability of the CBTs and their possible confounding effects on individual affective states, constrain the interpretative value of the judgement and attention biases, and consequently cast doubt upon the reliability of our validation method. Our validation method should, hence, not have been limited to the sole assessment of cognitive biases, but also have included additional measures influenced by mood processes.

As mentioned in Chapter 2, moods are multi-component psychological processes that are associated with measurable changes, especially in cognition and behaviour (Désiré et al. 2002; Mendl et al. 2010b). The assessment of housing-induced behavioural changes, on top of cognitive ones, might have shed light on the real influence of our housing contrasts on dairy heifers' moods. Even though the choice to manipulate each housing aspect was motivated by previous studies associating the corresponding housing manipulation with improved or worsened welfare (**Chapter 4**), additional behavioural analyses would have allowed us to assess the emotion-eliciting quality of our housing conditions on the focal heifers. In particular, behavioural analyses would have permitted the identification of the housing changes that effectively induced the occurrence of supposedly positive or negative events, as well as the detection of individuals that were the most sensitive (positively or negatively) to these housing changes. For instance, by investigating heifers' behavioural responses to the introduction of a new brush, one could determine whether the theory that *fearful* heifers were scared of the new brush is valid. If this hypothesis were true, *fearful* heifers would be expected to avoid contact with the brush – as postulated by the approach-avoidance theory of emotion (Elliot et al. 2013).

The analysis of heifers' behaviours within the different housing conditions would have also provided a clearer picture of the social climate within each group. When developing the model paradigm, the positive treatment was expected to promote affiliative behaviours as a result of long-term social familiarity (Gutmann et al. 2015), while the negative treatment was expected to accentuate agonistic interactions due to the increased competition for the cubicles and the feeding gates (Huzzey et al. 2006). Against expectations, however, it appeared that agonistic interactions occurred more frequently in the groups subjected to the positive treatment than in the groups subjected to the negative treatment (personal observation). Provided that analyses of heifers' behaviours within the home pens confirm this observation, several explanations for this finding should be explored. As suggested in Chapter 5, one explanation could be that *active* heifers in the positive conditions repeatedly engaged in agonistic interactions with older companion cows, in an attempt to assert their social ranks – *Activity* and age being two determinants of dominance (Verbeek et al. 1996; Barton et al. 1973). Additionally, it is possible that heifers in the negative conditions formed tight social bonds, as a result of shared negative experiences – thus leading to reduced agonistic interactions. This theory is coherent with previous studies demonstrating that shared adversity strengthens social bonds in humans and that moderate stressors promote social support behaviours and resource-sharing in rodents (Peng et al. 2021; Muroy et al. 2016). To provide cows with optimal housing facilities tailored to their behavioural needs, further research is thus warranted to investigate the effect of housing conditions on the social dynamics of the herd, a crucial determinant of affective states which is too often overlooked in research. Alternatively, but not mutually exclusively, groups in the positive conditions might have been – involuntarily – composed of individuals with divergent personality, while groups in the negative conditions might have been predominantly composed of individuals with similar personality. Such group compositions could have led to an increased occurrence of agonistic behaviours in the former case, and to an increased occurrence of affiliative behaviours in the latter case – as previously shown in non-human primates (Coleman 2012). After preliminary investigations, however, this theory appears unlikely since ten heifers shared a personality profile (*Activity* x *Fearfulness* x *Sociability*) with one of their pen mate heifers in the positive as well as in the negative treatment groups. The degree of personality dissimilarity between heifers and companion cows from the same group remains, however, unknown and may have significantly differed between the positive and the negative treatment groups. This consideration once more highlights the need to account for individual personality when constituting social groups, especially in the context of animal welfare studies.

In addition, the reliability of our validation method could have been improved by including a physiological assessment of heifers' responses to acute stress. Considering that positive and negative moods, respectively, buffer and worsen the individual experience of negative events (Raoult et al. 2017; van Steenbergen et al. 2021), one would expect heifers moved to the positive conditions to cope better with an acute-stress test (e.g. a hoof-trimming procedure) than they would do at baseline – assuming that the positive treatment was effective. Likewise, one would expect heifers moved to the negative conditions to struggle more with the acute-stress test than they would normally do at baseline – assuming that the negative treatment was effective. Physiologically, this could translate into a smaller peak of serum cortisol and a faster return to initial concentrations for heifers moved from the reference to the positive conditions (Speer and Mauricio 2017), and into a higher peak of serum cortisol with potentially longer return to initial concentrations for heifers moved from the reference to the negative conditions (Mendonça-De-Souza et al. 2007). Of note, the negative housing conditions could also induce different types of maladaptive cortisol responses with, for instance, lower peak of cortisol as a result of hypocortisolism (Juster et al. 2010) – a physiological dysregulation caused by chronic stress (Guilliams and Edwards 2010). To ensure a proper interpretation of the physiological read-outs obtained in response to the acute-stress test, I would, therefore, recommend combining cortisol measures with other physiological measures. The assessment of interleukin-6 (IL-6) could, for instance, help disentangle lower cortisol responses induced by positive mood from hypocortisolism-driven blunted cortisol responses, hypocortisolism being associated with increased levels of IL-6 (Papanicolaou et al. 1996). Following this idea, I developed an acute-stress test procedure during the second experiment of my PhD thesis to investigate the effects of the housing manipulations on heifers' cortisol and IL-6 responses. Measures of serum cortisol and IL-6 were, moreover, combined with measures of HRV data in order to obtain an overall assessment of the housing-induced physiological responses to acute-stress, i.e. at the neuroendocrine, immune and autonomic levels. The acute-stress test consisted of a standardised hoof-trimming procedure which lasted exactly 20 min per heifer. All heifers were exposed twice to the hoof-trimming procedure, both before and after the housing manipulations. For each hoof-trimming, blood samples and HRV recordings were collected 1) in the home pen at baseline (approximatively 10 min before the onset of the stressor), 2) in the hoof-trimming apparatus at the end of the stressor (20 min after the onset of the stressor) and 3) in the home pen when heifers were released from the stressor (50 min after the onset of the stressor). The results are not yet available but will hopefully help us determine whether the housing paradigm was effective. It is worth mentioning, nonetheless, that such physiological assessment does not (yet) constitute an adequate validation method on its own since research on the link between mood induction (in particular positive mood) and subsequent physiological responses to challenges is still in its infancy.

Finally, caution against circular reasoning is required when attempting to validate a model paradigm of mood induction – and that whichever the chosen validation method. As tempting as it may be, results obtained from a presumed model of affective states cannot, in turn, be used as proof of the model validity. Conforming to this idea, I refrained here from discussing the potential validity of our housing paradigm based on the results obtained from hair cortisol, heart rate variability and milk data presented in Chapter 5. The potential validity of these measures as indicators of affective states is discussed in the following section, under the assumption that the model paradigm was effective.

The quest for the ideal indicator of mood valence – a fantasy?

As introduced in Chapter 1, the final goal of this PhD study was to identify valid physiological indicators of mood – and more specifically of mood valence – which could be of practical relevance for on-farm assessments of cattle welfare. To this end, I investigated the validity of different physiological measures that could be non-invasively collected as indicators of valence: hair cortisol, heart rate variability and milk-derived measures (**Chapter 5**). The validity of these physiological measures was examined following two lines of reasoning. First, the influence of the housing contrasts and of heifers' personality on each physiological measure was examined under the assumption that the model paradigm effectively induced positive or negative shifts in heifers' mood – possibly in a personality-dependent manner. Second, the correlation structure between the physiological measures and heifers' judgement and attention biases was examined under the assumptions that judgement and attention biases constitute valid indicators of valence, and that valid indicators of valence share common underlying - traits (e.g. physiological mechanisms). The effects of the housing-induced mood shifts on each physiological measure are detailed according to heifers' personality traits in **Table 1.** Of particular interest, *inactive* heifers, *non-fearful* heifers and *non-social* heifers were found to have a higher relative high-frequency band (HF) – an indicator of parasympathetic activity (Shaffer and Ginsberg 2017) – when moved from the reference to the positive conditions. Provided that *inactive*, *non-fearful* and *non-social* heifers were more receptive to the positive conditions than *active*, *fearful* and *social* heifers, HF may constitute a promising and non-invasive indicator of mood shift in cattle, albeit restricted to positive shifts. Whether HF would constitute an absolute or a relative measure of positive valence is, however, impossible to determine. In addition, *fearful* heifers were found to have greater milk fluctuations – a possible indicator of impaired resilience (Poppe et al. 2020) – when moved from the reference to the negative conditions. Provided that *fearful* heifers were more vulnerable to the negative conditions than other heifers, milk fluctuations may constitute a valid – and readily available on-farm – indicator of mood shifts, this time restricted to negative shifts.

Against expectations, however, none of the physiological measures varied according to the valence continuum delineated by our housing conditions (i.e. from negative to positive) or consistently correlated with measures of judgement and attention biases across housing contexts. Based on the conjecture that a single valence continuum exists, these findings question the validity of our physiological measures as indicators of mood valence and highlight the complexity of identifying indicators of valence that would both be theoretically valid and of practical relevance. Note that the concept of a continuum for affective valence ranging from positive to negative is not a given, and researchers in the field of psychology have previously postulated that positive and negative valences may, in fact, constitute two separate dimensions (**Chapter 2**).

I would like to conclude this chapter by presenting the characteristics that one would expect from an ideal indicator of valence, and by discussing each of these characteristics with regard to the cognitive and physiological measures investigated – but not all yet necessarily described – during this PhD project. I will start by focusing on the theoretical requirements that an ideal indicator of valence should fulfil, before discussing some practical aspects that would facilitate its implementation on-farm.

Table 1. Mood-induced physiological changes according to heifers' classes of personality (**Chapter 5**). Up-arrows indicate a mood-induced physiological increase, and down-arrows indicate a mood-induced physiological decrease, with corresponding p-values inferior to 0.05. Hair cortisol (HC) in ng.g^{-1}, heart rate (HR) in beat.min^{-1}, relative power of the low frequency band (LF) in %, relative power of the high-frequency band (HF) in %, standard deviation perpendicular the line of identity in the Pointcaré plot (SD1) in ms, standard deviation along the line of identity in the Pointcaré plot (SD2) in ms, ratio between SD1 and SD2 (SD1.SD2^{-1}), sample entropy (SampEn), milk yield (MY) in kg, fat and protein-corrected milk yield (FPCM) in kg, daily milk fluctuation (LnVar) in kg, relative fat content in milk (pFat) in %, relative protein content in milk (pProt) in %, relative lactose content in milk (pLac) in %.

Positive mood shift		**Hair cortisol**	**Heart rate variability measures**							**Milk-derived measures**					
Personality trait	Personality class	HC	HR	LF	HF	SD1	SD2	SD1.SD2^{-1}	SampEn	MY	FPCM	LnVar	pFat	pProt	pLac
Activity	Active	=	=	=	=	=	=	=	=	=	=	=	=	=	=
	Inactive	=	=	↓	↑	=	=	=	=	=	=	=	=	=	=
Fearfulness	Fearful	=	=	=	=	=	=	=	=	=	=	=	=	=	=
	Non-fearful	=	=	=	↑	=	↓	=	=	=	=	=	↑	=	=
Sociability	Social	=	=	=	=	=	=	=	=	=	=	=	=	=	=
	Non-social	=	=	=	↑	=	↓	=	=	=	=	=	↑	=	=
Negative mood shift		**Hair cortisol**	**Heart rate variability measures**							**Milk-derived measures**					
Personality trait	Personality class	HC	HR	LF	HF	SD1	SD2	SD1.SD2^{-1}	SampEn	MY	FPCM	LnVar	pFat	pProt	pLac
Activity	Active	=	=	=	↑	=	=	↑	=	=	=	=	=	=	=
	Inactive	=	=	=	=	=	=	=	=	=	=	=	=	=	=
Fearfulness	Fearful	=	=	=	=	=	=	=	=	↓	=	↑	↑	=	=
	Non-fearful	=	↑	=	=	=	=	=	=	=	=	=	=	=	=
Sociability	Social	=	=	=	=	=	=	=	=	=	=	=	↑	=	=
	Non-social	=	=	=	=	=	=	=	=	=	=	=	=	=	=

The ideal indicator of valence in theory

First and foremost, an ideal indicator of mood valence must allow for clear predictions in terms of valence. In other words, the interpretation of an indicator's change in response to variations in animal affective valence should be unequivocal. This is, for instance, the case for judgement biases – for which decreased latencies reflect better affective states, while increased latencies reflect worsened affective states (Mendl et al. 2009). If valid, the HRV measure of SD1 (most commonly referred to as root mean square of successive inter heartbeat interval differences RMSSD) also allows for clear predictions in terms of valence, as increased SD1 has most often been associated with improved affective states while decreased SD1 has been associated with worsened affective states (**Chapter 1**). To put it another way, these indicators reflect specific changes both in the positive and negative ranges, hence without overlap. On the contrary, measures of attention biases do not allow for the formulation of *a priori* hypotheses in response to changes in affective states, potentially owing to the complexity of the relationship (i.e. not monotonic) between attention biases and affective valence. In practice, positive shifts in attention to threat are not necessarily associated with either positive or negative shifts in affective valence, nor are negative shifts in valence necessarily associated with either positive or negative shifts in attention to threat. Sustained attention to threat can, for instance, both reflect vigilance and curiosity (Raoult et al. 2017), while negative affective changes can be associated both with positive or negative shifts in individuals' attention to threat – as a result of attentional facilitation or attentional avoidance (Cisler and Koster 2010; Bethell et al. 2012). Likewise, measures of hair cortisol do not allow for clear predictions in terms of valence since both hypo- and hypercortisolism can result from negative affective states (Packer et al. 2019). Hence, measures of attention to threat and of cortisol cannot be used on their own to infer animal affective valence and may only be used as complementary indicators. Ideally, the relationships between indicators of affective valence and affective valence itself should, thus, be monotonic to allow for unique and straightforward interpretations.

Furthermore, an ideal indicator of valence should be valence-specific, meaning that it reliably tracks changes in valence-related processes but not changes in valence-unrelated processes. While, at first sight, it may seem absurd to mention such an evident characteristic of an indicator of valence, identifying indicators that actually are *uniquely* valence-specific appears to be one of the researchers' greatest challenges. The fact that not a single measure assessed within this PhD project (**Chapters 4 and 5**) appears to meet this characteristic confirms this idea, considering that our cognitive and physiological outcomes could all be attributed to valence-unrelated factors rather than housing-induced mood shifts. Arousal, in particular, has been identified as one of the main confounding factors in studies of affective valence. The influence of arousal on physiological measures, especially on measures of the autonomic activity, has repeatedly been demonstrated (e.g. Briefer et al. 2015). In our study, arousal was, hence, controlled for when studying the effect of housing-induced mood shifts on HRV parameters (**Chapter 5**). However, its effect was not accounted for when measuring hair cortisol or cognitive biases (**Chapters 4 and 5**). An influence of arousal on hair cortisol is highly plausible considering that higher arousal levels are associated with higher salivary cortisol (Rief et al. 1998) – a measure strongly correlated with hair cortisol (Xie et al. 2011). It is plausible that, in our conditions, arousal masked or distorted the effect of valence on hair cortisol – thus potentially leading to the lack of significant differences of hair cortisol observed between our different housing conditions (**Chapter 5**). The same assumption holds for our cognitive bias results. To my knowledge, no study in animals has ever investigated the *sole* effect of arousal on judgement and attention processes, even though this knowledge gap has already been pointed out (Baciadonna and McElligott 2015; Monk et al. 2018b). To ensure the validity of our cognitive measures as indicators of valence, the initial postulate that cognitive biases would reflect valence over arousal processes (Mendl et al. 2009) should be examined. Besides arousal, many other confounding variables may have influenced our physiological measures. While some of these have already been identified (e.g. season and hair cortisol: Heimbürge et al. (2020a), heat stress and milk yield: Ogundeji et al. (2021), water consumption and HRV: Heathers et al. (unpublished)), many more confounding factors will most likely emerge in the coming years considering the rapid pace at which research on affective sciences is evolving. My intention here, however, is not to provide an exhaustive list of such potential confounding factors, but rather to emphasize the idea that – to date – no measure provides an unequivocal assessment of affective valence. Provided that no single indicator can reliably assess valence, researchers should, thus, aim at developing composite indicators of valence in order to dilute the effect of unwanted – yet not easily controllable – processes on their indicator of valence, as well as control (where possible) for arousal.

The development of an allostatic load index (ALI) could serve this purpose. ALI is a multisystemic measure reflecting the cumulative burden of physiological dysregulations that occur in response to stress (McEwen and Stellar 1993). ALI focuses, in particular, on four physiological systems: the neuroendocrine, immune, autonomic and metabolic systems (McEwen 2012). Classically, each physiological system comprises a set of parameters relevant to disease risks, which often combines primary mediators and secondary outcomes of stress. Primary mediators of stress (e.g. cortisol, DHEA-S and IL-6) have both protective and damaging effects against stress since they maintain homeostasis in response to acute stress but also quicken disease progress in response to cumulative stress (McEwen 2006). Secondary outcomes of stress (e.g. blood pressure, total cholesterol and high-density lipoprotein) reflect sub-clinical disturbances which result from the sustained activation of the primary mediators occurring in response to cumulative stress (McEwen et al. 2012). Taken together, primary mediators and secondary outcomes of stress provide a more comprehensive picture of the health status than an independent parameter would (Schenk et al. 2018). Considering that positive and negative affective states contribute to individual health (Kubzansky and Kawachi 2000; Davidson et al. 2010), research on the associations between allostatic load, on the one hand, and positive or negative affective states, on the other hand, have recently emerged (Archer 2017; Schenk et al. 2018). Preliminary evidence indicates that the ALI increases in response to negative affective states and decreases in response to positive affective states (Schenk et al. 2018) – therefore suggesting that ALI may constitute a valid indicator of affective valence in humans. Since all mammals share common physiological stress responses, the ALI may also be used as an indicator of valence in other mammalian species – provided that some species-specific adjustments to the set of investigated parameters are made. For instance, while salivary alpha-amylase is often investigated in ALI applied to humans (e.g. Egorov et al. 2017), it does not constitute a suitable parameter for an ALI in dairy cows since the latter have no alpha-amylase in their saliva (Boehlke et al. 2015). In cows, salivary alpha-amylase could, instead, be replaced by another marker of chronic conditions, e.g. serum amyloid A or haptoglobin (Grönlund et al. 2005). Furthermore, in order to develop an ALI especially targeted at affective processes, one might consider including parameters that have previously been associated with distinct and complementary facets of affective states. Oxytocin, for instance, might serve as a proxy for social well-being (Rault et al. 2017), while ghrelin could be used as a measure of mood disorder (Wittekind and Kluge 2015). Based on these considerations, I propose an ALI specifically for dairy cows (**Figure 2**), whose parameters were collected weekly during the second experiment of my PhD. Its validity as an indicator of affective valence remains to be explored based on its association with our measures of cognitive bias and its response to our housing contrasts.

NEUROENDOCRINE	IMMUNE	AUTONOMIC	METABOLIC
Cortisol DHEA-s Ghrelin Oxytocin	Amyloid A Albumin Total protein Fibrinogen IL-6	HRV parameters	BHBA Insulin Glucose

Figure 2. Proposed set of parameters for an allostatic load index sensitive to affective processes in dairy cows. Dehydroepiandrosterone-sulfate (DHEA-s), interleukin-6 (IL-6), heart rate variability (HRV), β-hydroxybutyric acid (BHBA).

The metabolomic profile of affective states could also be used as a composite, albeit unisystemic, indicator of affective valence. Nowadays, metabolomic techniques constitute hypothesis-free methodologies that provide the simultaneous and quantitative measurements of hundreds of metabolites within a single biofluid sample (Hivert et al. 2015). Metabolomic techniques allow for the identification of the metabolic signature of diverse phenomena – including affective processes. To date, metabolomics has mostly been used in affective sciences to identify the metabolomic profile of affective disorders such as depression (Ding et al. 2014; Yu et al. 2016), but no study has investigated its potential as an indicator of non-clinical psychological conditions (Trudel-Fitzgerald et al. 2017). Further studies are warranted to investigate whether individuals under affective states of contrasted valence differ in their metabolomic profile. In our second experiment, blood samples were collected both at the end of the reference and of the experimental housing conditions to assess whether the housing conditions influenced heifers' metabolite profiles. Blood samples were analysed using non-targeted global metabolite profiling (Evans et al. 2014; Ford et al. 2020). Our preliminary results, however, did not reveal any significant differences between the metabolome profile of heifers housed in the positive conditions and the metabolome profile of heifers housed in the negative conditions. Distinct changes in heifers' metabolomes were, nonetheless, noticed between individuals moved from the reference to the positive conditions and individuals moved from the reference to the negative conditions. The exact causes behind these housing-specific physiological changes remain to be determined.

Finally, an ideal indicator of mood valence should provide an absolute measure of valence. While humans can convey the intrinsic value of how well they are doing (i.e. either ‘good’ or ‘bad’) using verbal reports, proxy measures in animals can only provide relative measures of affective experiences. Current measures of affective valence can simply be used to compare animals’ affective experiences in response to different alternatives (Bateson and Nettle 2015). Consequently, relative measures of valence allow the investigation of whether one alternative improves or worsens animals’ affective states compared with other alternatives, but they do not permit to establish whether the animals are, *per se*, doing well (Ahloy-Dallaire et al. 2018). In other terms, animals may still be under negative affective states, even though they become more optimistic, have higher HF and lower milk fluctuations after being subjected to a supposedly positive treatment. Based on this premise, animal welfare scientists may never know when their ultimate goal to provide animals with ‘a good life’ is achieved (Green and Mellor 2011). From an ethical standpoint, this argument should incite researchers and farmers to never stop looking for better alternatives in the hope of eventually meeting this goal.

The ideal indicator of valence in practice

To ensure its implementation on-farm, an ideal indicator of affective valence should also meet several practical constraints.

First, the amount of preparation required for the indicator’s collection should be minimal. No equipment, other than those readily available on-farm, should be required. That is, for instance, the case for hair cortisol – whose collection solely necessitates the acquisition of a hair clipper and a disinfectant (e.g. Sharma et al. 2019). In contrast, heart rate data collection requires the procurement of heart rate belts designed for humans, which must be manually adjusted to fit the animal of focus (**Chapter 5**). Similarly, video recordings for analyses of animal behaviours during JBT and ABT often necessitate to invest in high-quality cameras, which must furthermore be properly installed before the beginning of the tests. If specific materials are necessary for the indicator’s collection, their installation on-farm should at least be trivial for the assessment of animal affective valence to be considered practical.

For ease of preparation, the level of expertise required for the indicator's collection should also be reasonable. Farmers interested in monitoring their animals' affective states should not have to undergo extensive training periods to obtain the minimal set of skills or knowledge needed for the indicator's assessment. Hair samples of good quality can, for instance, be obtained following simple instructions that require little to no training. Extensive periods of training under the supervision of an experimented colleague are, however, often required in order to succeed in performing relatively painless and stress-free blood samples. For an indicator of valence to be ideal, the training period required for its assessment should, hence, be minimal.

Second, the collection of the indicator itself should be straightforward. In practice, one person (ideally the farmer themselves) should be able to collect the indicator. To this end, animal handling should be limited – in particular when working with large animals where handling can be hazardous (Douphrate et al. 2009). In our experiment (**Chapter 5**), for instance, hair samples required little animal handling as heifers could be easily restrained at the feeding gates when fresh feed was provided. In contrast, the collection of judgement and attention biases required regular handling, before the start of the CBTs in order to bring heifers into the waiting arena and in-between the judgement bias trials (**Chapters 3 and 4**). As far as possible, the procedure of the indicator's collection should, thus, be automated to improve the operator's safety and smoothen the collection process. In this regard, automated JBTs have been recently proposed (Jones et al. 2018).

The collection of an ideal indicator of valence should also be relatively fast. Provided that such an indicator must be animal-based as each individual's affective experience is unique (Keeling et al. 2021), the time input required per animal for the indicator's assessment constrains the overall number of individuals for which the affective experience can be monitored. The collection of indicators directly derived from milk production data, for instance, does not require any extra workload for dairy farmers, since milk yield is automatically recorded for all lactating cows in modern farms. The collection of basal HRV data, on the other hand, can be time-consuming as extensive periods of recordings (e.g. several hours) per animal are sometimes required in order to obtain valid but short periods of measurements (e.g. several minutes). Furthermore, the collection of HRV data necessitates regular surveillance of the effective recordings of the HRV, as the sensors may fall when animals engage in routine behaviours like scratching against a brush. To ensure that the maximum number of individuals' affective valence can be assessed within a reasonable timeframe, the collection process of an indicator of valence should be easy to integrate into the farmers' working routine.

In addition, the collection of an ideal indicator of valence should be repeatable. Considering that mistakes occur in practice, the operator should be able to reiterate the collection procedure without affecting the validity of the results. This implies that 1) the timing of collection has no significant influence on the indicator of interest, and that 2) the material of collection (e.g. blood volume, hair quantity) is available in sufficient amount to allow for a second harvest. In a scenario where a hair sample is blown away due to heavy wind, for instance, a second chunk of hair can immediately be collected without impacting the cortisol concentrations – as the amount of hair cortisol produced in-between the two sampling moments will not yet have been incorporated into the hair shaft, i.e. the part of the hair above the skin that can be shaved (Heimbürge et al. 2020b). However, the quantity of hair collected on the second occasion may be insufficient for cortisol assays. To avoid the potential confounding effect of the body location on hair cortisol concentrations (Heimbürge et al. 2020a), the operator should collect the hair from the same body zone on both attempts. The remaining hair available for sampling will nonetheless be limited by the size of the spot which had initially been shaved, weeks or months before the first attempt of hair collection. Whenever possible, the protocol for the indicator's collection should, hence, be developed based on the possibility that the sampling procedure may have to be repeated.

Third, an ideal indicator of valence should be easy to analyse. In practice, the tools (i.e. software or equipment) required for the analysis of such an indicator should be easy-to-access. Multiple measures of HRV can, for instance, be obtained using the open-access version of Kubios, i.e. Kubios Standard, which can be downloaded online. Anyone with a computer and Internet access has the opportunity to analyse recordings of heart rate data – if they are willing to. In contrast, public access to metabolomic platforms is not permitted, considering that the technology required for metabolomic analyses is restricted to a limited number of private companies and to some universities. Ideally, the tools necessary for the indicator's analysis should, hence, be available at any given time – at the user's convenience.

In line with this idea, the analysis of an ideal indicator of valence should also be achievable in complete independence – without having to rely on a third party to obtain the expected readouts. Given the ease of access to open-source multimedia players, analyses of behavioural recordings (e.g. latency to reach an ambiguous cue) can, for instance, be self-performed, at the user's own pace. In contrast, blood sampling analyses often require materials and expertise that necessitate the services of a laboratory. As a result, one becomes dependent upon others' availabilities, and must be prepared to invest time into negotiations and energy into the preservation of good relationships (Fox and Faver 1984). Therefore, to ensure that farmers obtain their results at the deadline they expect, the analyses of the indicator should be feasible without the need for external assistance.

Furthermore, the analysis of the indicator of valence should be fast. Ideally, farmers should not lose extra time on the indicator's analysis, once it has been collected (Battini et al. 2014). Indicators derived from routinely collected data meet this criterion. In our study, the milk fluctuations could be rapidly calculated from individual milk yield using a simple mathematical formula. In practice, such an algorithm should, thus, be implemented into farmers' most popular executive programs to ensure the automatic and immediate transmission of the indicator of valence.

Finally, the preparation, the collection and the analysis of an ideal indicator of valence should be cost-effective. Ideally, the total cost of 1) assessing animal affective valence and 2) taking the subsequent actions to improve animal welfare (that is, if necessary) should be inferior to the economic gain derived from animals' enhanced affective states – as farmers' perception of animal welfare is influenced by economic values (Becker et al. 2014; Balzani and Hanlon 2020). Further studies are required to investigate the economic benefits of improving animals' affective states. Two main lines of research may be of particular interest. First, more studies should investigate the influence of improved affective states on long-term health-related costs. Provided that positive affective states effectively contribute to individual health status (Richman et al. 2005; Davidson et al. 2010), it is reasonable to expect that veterinarian costs would decrease as animals' affective states improve. Second, more research should aim at modelling the economic gain that farmers could obtain from selling products derived from certified 'happier animals'. In practice, consumers are willing to pay more for products of more humanely raised animals, provided that they trust the label claims (Spain et al. 2018). It is, therefore, plausible that farmers could sell their products at a better price if the latter meet the scientifically defined requirements for better animal affective experiences.

Overall, numerous qualities – both theoretical and practical – are expected from an indicator of mood valence to ensure its reliability and its implementation on-farm. At this point, however, most research efforts should focus on developing valid assessments of affective valence. While various measures have been suggested as promising indicators of animal affective valence, no unique measure – to date – allows for an unequivocal and comprehensive assessment of animal affective valence. Future studies should, thus, aim at developing composite indicators of affective valence based on measures that convey complementary information with regard to the animal affective experience. Once such an indicator has been developed, studies should then work at improving the feasibility of the indicator's assessment to ensure its widespread use on farms. This could possibly be achieved by assessing, in a non-invasive manner, the different measures constitutive of the validated indicator using 1) technology sensors (Neethirajan et al. 2021) and 2) routinely collected media samples (e.g. milk in dairy cows) – provided that the studied measures remain indicative of the affective processes of interest in the newly considered media sample.

Conclusion

This PhD thesis contributes to our understanding of affective states in dairy cows in several ways. First, it proposes a judgement bias task (JBT) set-up suitable for adult dairy cows, which is relatively feasible and sensitive to affective shifts. Up to date, the JBT appears to represent the most promising methodology to infer animal mood valence. To ensure its validity, however, a certain number of fundamental questions remain to be answered: which kind of affective states (e.g. mood, emotion or both) does JBT assess? To which extent does personality influence judgement bias?

In addition, this thesis investigated the potential of an attention bias task (ABT) designed specifically for dairy cows as an alternative to the JBT for assessing affective valence. Although more practical than the JBT, it appears that the ABT is fraught with several limitations – which question the validity of attention bias as an indicator of valence. To date, it remains unclear whether individual responses to the attention bias test are sensitive to positive shifts in affective valence, or if their sensitivity is restricted to shifts in negative valence. Furthermore, the interpretation value of the ABT appears limited since responses to the attention bias test may lead to distinct, and even opposed, interpretations in terms of affective valence.

Furthermore, this thesis attempted to develop a valid model of mood valence for dairy cows by investigating its influence on cows' responses to both judgement and attention bias tests. The model consisted of housing conditions of supposedly contrasted hedonic values (e.g. positive or negative), which differed in stocking density, social stability and enrichment level. In practice, housing conditions had little effect on cows' cognitive biases. These results suggest that the housing model was ineffective at inducing shifts in mood valence – at least in our animal model. In particular, the pregnancy status, the age and the initial psychological states of the cows may have influenced their sensitivity to the housing changes. Alternatively, these results may echo the failure of the cognitive bias responses to reflect mood changes. In this PhD work, the effects of the housing contrasts on cognitive biases were investigated in a longitudinal fashion, when individuals were first housed under reference housing conditions before being brought under either more positive or negative housing conditions. However, the effect of the repeated exposures to ABT and JBT on individual cognitive responses was not controlled for, and may have altered the cognitive bias outcomes.

Finally, this thesis explored the validity of non-invasive physiological measures as more practical substitutes to the JBT for assessing mood valence. To this end, the effects of the housing contrasts and of personality – two possible determinants of affective states – on cows' physiological responses were explored. It was demonstrated that the housing conditions influenced the physiological outcomes in a personality-dependent manner. Such findings emphasize the need to account for personality differences when assessing valence, and challenge the idea of a 'one-fit-for-all' type of housing conditions. HF and milk fluctuations were, furthermore, identified as promising indicators of positive valence and negative valence, respectively. However, no single physiological measure investigated in this PhD project was responsive to both positive and negative housing changes – thus questioning their validity as indicators of affective valence. Overall, my work strengthens the idea that no single measure of affective states currently allows for unequivocal assessment of affective valence, and highlights the need for future research to combine complementary measures of affective processes in order to reliably infer animal affective valence.

Miscellaneous

Summary

Addressing the welfare issues inherent to dairy cow farming has now become an ethical priority since dairy cow welfare is considered the second greatest animal welfare problem in Europe. To ensure that dairy cows are raised under optimal housing conditions in commercial farms, valid assessments of dairy cow welfare that are compatible with on-farm monitoring and that align with the current understanding and definition of animal welfare are warranted. Nowadays, animal welfare revolves around the concept of 'affective states', which are subjective experiences that vary in valence – from pleasant/positive to unpleasant/negative – and in arousal. Affective states are linked to measurable changes in cognition, behaviour, and physiology; and they typically include both emotions and moods. Emotions are defined as intense but short-lived adaptive psychological processes that occur in response to specific stimuli, while moods represent longer lasting states which would result from an accumulation of emotions. In the context of animal welfare, mood valence seems of primary importance since it codes information on the intrinsic pleasantness or aversiveness of long-term, integrated animal experiences. Over the last decade, several cognitive and behavioural indicators of mood valence have been proposed, the suitability of which is hardly compatible with on-farm assessment. The potential of physiological measures as valid indicators of mood valence, however, has been overlooked while physiological assessments could be suitable with on-farm monitoring of dairy cow welfare. ***The aim of this thesis was to identify physiological indicators of mood valence in dairy cows that would be compatible with on-farm assessments.***

Based on an up-to-date overview of the field of animal affective states (**Chapter 2**), the judgement bias – which reflects mood-driven shifts in the interpretation of ambiguous stimuli – was identified as the gold-standard indicator of mood valence in non-human animals. Although incompatible with on-farm assessment as its evaluation requires animal training, the judgement bias allows to make inferences about animal mood. As in humans, animals in a positive mood interpret ambiguous information more positively, i.e. are more optimistic and display more positive judgement biases, than animals in a more negative mood – and vice versa. If properly assessed, judgement bias

may thus serve as a validation measure for further identification of more practical physiological indicators of mood valence.

To evaluate judgement biases in cows, a method of judgement bias assessment, namely the Judgement Bias Task (JBT), was designed specifically for dairy cows (**Chapters 3 and 4**). Three JBTs were developed in which cows were trained to discriminate between two locations of a feeder, which signalled either a reward or a punisher. All JBTs were associated with the same reward – i.e. concentrates – but differed in terms of punishers, cows being allocated either to 'no-reward', an air puff or an electric shock. The objective was to identify the most feasible and sensitive JBT to ensure that subsequent judgement bias assessments would require reasonable training periods and detect both positive and negative shifts in dairy cow mood. The repeatability of each JBT was also assessed separately to investigate their suitability in the context of longitudinal studies. Cows were exposed to ambiguous stimuli on two distinct occasions – which were separated by a wash-out period of ten days that exclusively consisted of extra training. The JBT associated with the air puff was the most feasible and sensitive method of judgement bias assessment – therefore suggesting that researchers should opt for a combination of rewards and punishers of equal strength when developing JBTs. However, the repeatability of none of the JBTs could be supported. In future studies, longer wash-out periods between consecutive testing may thus be necessary to ensure the appropriate assessment of judgement bias in longitudinal designs.

A putative model of mood valence based on housing manipulations was also developed in an attempt to induce positive and negative shifts in dairy cow mood (**Chapter 4**). The model was inspired by previous models designed for various farm species, which succeeded at inducing mood shifts based on housing manipulations. The housing treatment involved changes in three housing elements previously found to influence cow welfare: the stocking density, the social stability, and the level of enrichment. After being housed under stable reference conditions for 9 weeks, cows were subjected either to weekly positive or weekly negative changes in their housing conditions for 5 weeks. To validate the housing contrasts as a model of mood valence, judgement biases before and after the application of the housing treatment were compared. Against expectations, housing did not influence judgement biases – therefore suggesting either that the housing model failed at inducing mood shifts in dairy cows or that the JBT was unable to detect valence differences in dairy cow mood.

The influence of personality on judgement biases both before and after the housing manipulations was also investigated (**Chapter 4**). Considering that personality – i.e. a correlated set of individual behavioural and physiological traits that are consistent over time and across contexts – modulates the subjective experience of events, the objective was to determine whether personality influences the cow perception of the JBT *itself* as well as the cow perception of the housing conditions. Cow personality was investigated using a battery of previously used challenge tests: Open-Field, Novel-Object and Runway tests. Results demonstrated that cow personality was multifaceted and could be described according to three traits labelled Activity, Fearfulness and Sociability. Personality influenced responses to the JBT before the application of the housing treatment. Provided that cow moods were effectively standardized at the end of the reference period, these findings question the task's generalizability across individuals and emphasize the need to account for personality when assessing the effects of treatments on judgement biases. Personality also appeared to influence housing-induced judgement biases. This result could indicate that cows experienced the housing conditions differently according to their personality – thus challenging the existence of a 'one-fit-for-all' type of optimal housing conditions.

Besides, the validity of three categories of non-invasive physiological measures as indicators of mood valence in dairy cows was examined (**Chapter 5**): hair cortisol, heart rate variability (HRV) and milk-derived measures. Once again, these physiological measures were assessed – alongside judgement biases – when cows were first kept under the reference housing conditions and once brought under either supposedly more positive or negative housing conditions. Their validity was investigated following two lines of reasoning. First, the effects of the housing contrasts and cow personality on the different physiological measures were assessed on the assumption that the housing model was effective at inducing mood shifts. Second, the correlations between the physiological measures and the judgement biases were examined following the idea that the judgement bias was able to detect housing-induced mood shifts in cows. Against expectations, none of the physiological measures varied according to the valence continuum delineated by the housing conditions or were consistently correlated with measures of judgement bias across housing contexts. These results suggest that none of the studied physiological measures constitutes a valid indicator of mood valence. The housing conditions, however, influenced the relative power of the high-frequency band (HF – a measure of HRV) and milk fluctuations in a personality-dependent manner. Inactive heifers, non-fearful heifers and non-social heifers moved to the positive conditions had higher HF, while fearful heifers moved to the negative conditions had greater milk fluctuations. HF and milk fluctuations may thus constitute promising indicators of positive and negative mood valence, respectively, but further research is warranted to confirm this idea.

Finally, the findings of this study were discussed focusing on three main points: (i) the use of judgement biases as gold standard indicators of mood valence, (ii) the development of effective models of mood valence, and (iii) the arduous quest of identifying an ideal indicator of valence (**Chapter 6**). First, methods of assessment of judgement biases are still fraught with several limitations which impede their appropriate use in the context of mood studies. In practice, the repeatability of the JBT is not guaranteed, and set-ups suitable for the longitudinal assessments of judgement biases should be developed. Furthermore, the interpretation of judgement biases in terms of affective processes remains equivocal. Animal responses to the JBT may not only be sensitive to the intended mood treatment but may also be influenced by uncontrollable environmental stimuli and by individual personality – in an affect-related or in an affect-unrelated manner.

Second, the development of effective models of mood valence is currently hampered by the absence of validated theories of mood induction. To ensure that indicators of mood valence are identified in the future, fundamental research is required to provide a functional recipe for mood induction in animals. The generalizability of models of mood valence should also be guaranteed – whichever the status of the animals of focus (i.e. regardless of the pregnancy stage, age and initial psychological state). Considering the limitations inherent to the JBT, the validation method of such a model should integrate behavioural and physiological assessments – in addition to cognitive bias assessments – to ensure its own reliability.

Third, the quest for the identification of an indicator of mood valence that would combine all the practical and theoretical qualities expected from an ideal indicator remains arduous. Considering that no indicator to date allows for an unequivocal assessment of mood valence, future research should aim at developing a composite indicator of mood valence. The latter should rely on complementary measures of mood processes to ensure the provision of an accurate and comprehensive picture of animal mood. The development of an allostatic load index and the use of -omics techniques may serve this purpose.

Résumé (French summary)

Traiter la question du bien-être animal en élevage laitier est devenu une priorité éthique puisque le bien-être des vaches laitières est considéré comme le deuxième plus grand problème de bien-être animal en Europe. Pour s'assurer que les vaches laitières sont élevées dans des conditions optimales, il est nécessaire de développer des indicateurs de bien-être animal qui soient 1) compatibles avec des évaluations de bien-être menées directement dans les élevages et 2) qui résonnent avec la définition scientifique actuelle du bien-être animal. De nos jours, le principe de bien-être animal repose sur le concept d'états affectifs. Le terme état affectif désigne une expérience subjective vécue par un individu, qui se définit selon sa valence (d'agréable/positif à désagréable/négatif) et son niveau d'excitation. Les états affectifs sont associés à des changements mesurables dans la cognition, le comportement et la physiologie des individus, et sont communément apparentés aux notions d'émotions et d'humeurs. Les émotions sont définies comme des processus psychologiques intenses et de courtes durées en réponse à des stimuli spécifiques, tandis que les humeurs représentent des états psychologiques de plus longues durées qui découleraient de l'accumulation d'émotions ressenties antérieurement. Dans le cadre d'études du bien-être animal, la valence thymique (c'est-à-dire relative à l'humeur) revêt une importance capitale puisqu'elle caractérise le plaisir ou le désagrément global ressenti par les animaux sur le long terme. Au cours de la dernière décennie, plusieurs indicateurs cognitifs et comportementaux de valence thymique ont été proposés, mais leur détermination n'autorise pas une évaluation pratique du bien-être animal qui serait compatible avec les conditions actuelles d'élevage. En revanche, le potentiel des mesures physiologiques en tant qu'indicateurs de la valence thymique a été négligé, alors que des évaluations physiologiques pourraient permettre un suivi aisé et fréquent du bien-être des vaches laitières en élevage. L'objectif de cette thèse était donc d'identifier des indicateurs physiologiques de valence thymique chez les vaches laitières qui seraient compatibles avec des évaluations pratiques du bien-être animal au sein des exploitations agricoles.

Notre revue des connaissances scientifiques des états affectifs chez les animaux a démontré que le biais de jugement constitue l'indicateur de référence de valence thymique chez les animaux (**Chapitre 2**). Bien qu'incompatible avec une évaluation du bien-être animal en exploitation agricole en raison du dressage animal qu'il requiert, le biais de jugement permet de déduire l'humeur des animaux à partir des

interprétations que ces derniers font de situations ambigües. Comme chez les humains, les animaux de bonne humeur interprètent les informations ambiguës de manière plus positive, c'est-à-dire qu'ils sont plus optimistes, que les animaux de plus mauvaise humeur – et vice versa. S'il est correctement évalué, le biais de jugement peut donc servir de mesure de validation d'autres indicateurs de valence thymique, potentiellement plus pratiques à déterminer.

Afin de mesurer les biais de jugement chez les vaches laitières, une méthode de leur évaluation a été conçue spécifiquement pour elles : le Judgement Bias Task (JBT) (**Chapitres 3 et 4**). Au total, trois JBTs ont été développés durant lesquels des vaches furent entrainées à différencier deux positions d'une mangeoire en fonction de l'évènement qu'elles signalaient. L'une des positions indiquait la distribution d'une récompense (150 g de concentrés) alors que l'autre indiquait la distribution d'une punition. Les punitions délivrées variaient en fonction du JBT, les vaches recevant soit « aucune récompense », un jet d'air, ou un choc électrique. L'objectif était double. Tout d'abord, il s'agissait d'identifier le JBT le plus aisé à mettre en place afin de s'assurer, qu'à l'avenir, les mesures de biais de jugement ne nécessitent pas de périodes de dressage trop chronophages. De plus, il s'agissait d'identifier le JBT le plus sensible aux variations d'humeur chez les vaches de façon à garantir la détection de changements d'humeur à la fois positifs et négatifs lors d'expériences futures. La répétabilité de chaque JBT fut aussi évaluée afin d'étudier leur potentiel en tant qu'outil de mesure de biais de jugement dans le contexte d'études longitudinales. Pour ce faire, les vaches furent exposées à des situations ambiguës à deux occasions, séparées l'une de l'autre par une période de dressage de 10 jours. Le JBT associé avec le jet d'air fut identifié comme l'outil d'évaluation de biais de jugement le plus aisé à mettre en place et le plus sensible. Ce résultat devrait inciter les chercheurs à utiliser des combinaisons de récompenses et de punitions de forces égales lorsqu'ils élaborent des JBTs. Cependant, aucun JBT ne s'avéra répétable. Il se pourrait donc que les mesures de biais de jugement doivent être séparées par de plus longues périodes de dressage pour être correctement utilisées dans le cadre d'études longitudinales.

Un modèle présumé de valence thymique basé sur une modification des conditions de logement a également été développé dans le but d'induire des changements positifs et négatifs d'humeur chez les vaches laitières (**Chapitre 4**). Ce modèle a été conçu à partir de précédents modèles parvenus à entrainer des changements d'humeur chez d'autres animaux d'élevage en modifiant leurs conditions de logement. Dans le cadre de notre expérience, ces modifications concernaient trois aspects de logement connus pour influencer le bien-être des vaches : le chargement, la stabilité sociale et le niveau d'enrichissement. Après avoir été hébergées dans un environnement stable pendant 9 semaines (conditions de référence), les vaches furent soumises à des changements hebdomadaires, soit positifs soit négatifs, de leurs conditions de logement pendant 5 semaines (conditions expérimentales). Afin de valider nos conditions de logement comme modèle adéquat de valence thymique, les biais de jugement des vaches furent comparés avant et après la modification des conditions de logement. Contre toute attente, les conditions de logement n'influencèrent pas les biais de jugement des vaches. Deux interprétations de ce résultat sont plausibles. Premièrement, il se peut que les modifications des conditions de logement n'aient pas engendré les changements d'humeur prédits chez les vaches. Deuxièmement, il est possible que le JBT lui-même n'ait pas été en mesure de détecter les variations de valence thymique des vaches laitières.

L'influence de la personnalité des vaches sur les biais de jugement fut aussi déterminée avant et après manipulation des conditions de logement (**Chapitre 4**). La personnalité des individus consiste en un recueil de traits comportementaux et physiologiques qui demeurent constants quels que soient les périodes ou les contextes d'étude. Puisque la personnalité des individus module l'expérience ressentie par chacun dans une situation donnée, l'objectif fut d'établir si la personnalité des vaches influence leur perception du JBT ainsi que leur perception des conditions de logement. La personnalité des vaches fut analysée à l'aide d'une série de tests précédemment validés : l'Open-Field test, le Novel-Object test, et le Runway test. Les résultats prouvent que la personnalité des vaches a plusieurs facettes et qu'elle peut être décrite selon trois traits que sont l'Activité, la Crainte et la Sociabilité des individus. Il fut démontré que la personnalité des vaches influence leurs réponses au JBT avant toute manipulation des conditions de logement. Sous réserve que toutes les vaches étaient d'humeur semblable dans les conditions de logement de référence, ce constat questionne la possibilité de généraliser les résultats du JBT à différents individus et souligne l'importance de tenir compte de la personnalité des animaux lors de l'évaluation des effets de certains traitements sur les biais de jugement. En outre, il semblerait que la personnalité des vaches influence aussi les biais de jugement induits par les modifications des conditions de logement. En conséquence, les vaches percevraient différemment les changements de leurs conditions de logement en fonction de leur personnalité – une idée qui remet en cause l'existence de conditions de logement optimales et uniques pour toutes les vaches.

Par ailleurs, la validité de trois catégories de marqueurs physiologiques pouvant être mesurés de façon non-invasive en tant qu'indicateurs de valence thymique chez les vaches laitières fut examinée (**Chapitre 5**) : le cortisol capillaire, des mesures de variabilité du rythme cardiaque et des mesures dérivées du lait. De même que les biais de jugement, ces marqueurs physiologiques furent étudiés quand les vaches étaient hébergées dans les conditions de logement de référence et dans les conditions de logement expérimentales positives ou négatives. Leur validité fut examinée suivant deux axes de raisonnement. En premier lieu, les effets de conditions de logement et de la personnalité des vaches sur les différents marqueurs physiologiques furent examinés, en partant du principe que nos conditions de logement suscitèrent les changements d'humeur désirés chez les vaches. En second lieu, les corrélations entre les différents marqueurs physiologiques et les biais de jugement furent examinées, en partant du postulat que les biais de jugement reflétaient bien les changements d'humeur induits chez les vaches par la modification des conditions de logement. Cependant, aucun des marqueurs physiologiques ne varia en fonction du continuum de valence délimité par les conditions de logement ni ne corréla de façon constante avec les biais de jugement mesurés dans les différentes conditions de logement. Ces résultats indiquent qu'aucun des marqueurs physiologiques étudiés ne constitue un indicateur valide de valence thymique. Néanmoins, les conditions de logement influencèrent différemment les hautes bandes de fréquence cardiaques (HF, une mesure de variabilité du rythme cardiaque) et les fluctuations de productions laitières, selon la personnalité des vaches. A la suite de leur déplacement vers les conditions de logement favorables, HF augmenta chez les vaches inactives, les vaches non-craintives et les vaches peu sociales. De même, les fluctuations de productions laitières augmentèrent chez les vaches craintives déplacées vers les conditions de logement défavorables. HF et les fluctuations de productions laitières pourraient donc constituer, respectivement, des indicateurs de valence positive et négative. Davantage d'études sur le sujet sont néanmoins requises pour confirmer cette théorie.

Enfin, les résultats de cette étude furent discutés selon trois axes : (i) l'utilisation des biais de jugement comme indicateurs de référence de la valence thymique, (ii) le développement de modèles valides de valence thymique, et (iii) la difficulté à identifier un indicateur idéal de valence thymique (**Chapitre 6**). Tout d'abord, les méthodes d'évaluation actuelles des biais de jugement présentent encore certaines limites qui freinent leur utilisation adéquate dans le cadre d'études sur les humeurs animales. En pratique, la répétabilité des JBTs n'est pas garantie, et les prochaines recherches devraient s'attacher à développer des JBTs dont la configuration est compatible avec des études longitudinales. De plus, l'interprétation des biais de jugement en termes de processus affectifs demeure équivoque. Il semblerait que la réponse des animaux au JBT ne reflète pas uniquement l'effet d'un traitement affectif sur leur humeur mais qu'elle soit également influencée par des facteurs environnementaux incontrôlables ainsi que la personnalité de chaque individu – indépendamment ou non de tout processus affectif.

En second lieu, le développement de véritables modèles de valence thymique est actuellement freiné par l'absence de théorie valide sur la façon dont les humeurs sont générées. Afin de permettre l'identification d'indicateurs de valence thymique, des études visant à établir un protocole clair d'induction d'humeur sont nécessaires. La capacité de généralisation de tels modèles de valence thymique à différents individus devrait aussi être garantie – peu importe le statut de l'animal étudié (quel que soit son stade de grossesse, son âge ou son état psychologique initial). Compte tenu des limitations actuelles inhérentes au JBT, la méthode de validation de ces modèles devrait par ailleurs reposer sur des mesures comportementales et physiologiques, en plus des mesures de biais de jugement, afin de s'assurer de leur fiabilité.

En dernier lieu, la quête d'identification d'un indicateur de valence thymique dont les caractéristiques répondraient à tous les critères théoriques et pratiques attendus d'un indicateur idéal s'annonce ardue. Puisqu'aucun indicateur étudié ne permet actuellement d'identifier l'humeur des animaux de façon univoque, les prochaines études devraient chercher à développer un indicateur composite de valence thymique. Pour offrir une image précise et complète des humeurs animales, cet indicateur devrait reposer sur des mesures de processus affectifs complémentaires. A l'avenir, le développement d'index de charge allostatique et l'utilisation de techniques -omiques pourraient permettre d'atteindre cet objectif.

References

Adriaense JE, Martin JS, Schiestl M et al (2019) Negative emotional contagion and cognitive bias in common ravens (Corvus corax). Proc. Natl. Acad. Sci. 116(23):11547-11552. https://doi.org/10.1073/pnas.1817066116

Ahloy-Dallaire J, Espinosa J,Mason G (2018) Play and optimal welfare: Does play indicate the presence of positive affective states? Behav. Process. 156:3-15. doi: 10.1016/j.beproc.2017.11.011

Alonso ME, González-Montaña JR, Lomillos JM (2020) Consumers' concerns and perceptions of farm animal welfare. Animals 10:1–13. https://doi.org/10.3390/ani10030385

Alvares GA, Quintana DS, Kemp AH, et al (2013) Reduced Heart Rate Variability in Social Anxiety Disorder: Associations with Gender and Symptom Severity. PLoS One 8:1–8. https://doi.org/10.1371/journal.pone.0070468

Amsel A (1992) Frustration theory. An analysis of dispositional learning and memory. Camb. Univ. Press. https://doi: 10.1037/0033-2909.112.3.396

Andersen AH, Zhang Z, Barber T et al (2002) Functional MRI studies in awake rhesus monkeys: methodological and analytical strategies. J. Neurosci. Methods 118(2):141-152. doi: 10.1016/s0165-0270(02)00123-1

Anderson DJ, Adolphs R (2014) A framework for studying emotions across species. Cell 157(1):187-200. doi: 10.1016/j.cell.2014.03.003

Andrew RJ (1963) Evolution of facial expression. Science 142(3595):1034-1041. doi: 10.1126/science.142.3595.1034

Archer J (1973) Tests for emotionality in rats and mice: a review. Anim. Behav. 21(2):205-235. https://doi.org/10.1016/S0003-3472(73)80065-X

Archer T (2017) Health benefits for ageing : positive affect and life satisfaction , exercise and coping , and telomere length. J Ment Heal Aging 1:13:17. https://doi.org/10.35841/mental-health-aging.1.1.13-17

Asher L, Friel M, Griffin K,Collins, LM (2016) Mood and personality interact to determine cognitive biases in pigs. Biol. Lett. 12(11):1-4

Avarguès-Weber A, de Brito Sanchez MG, Giurfa M, Dyer AG (2010) Aversive reinforcement improves visual discrimination learning in free-flying honeybees. PLoS One 5:. https://doi.org/10.1371/journal.pone.0015370

Baciadonna L, McElligott AG (2015) The use of judgement bias to assess welfare in farm livestock. Anim. Welf. 24:81-91. https://doir.org/ 10.7120/09627286.24.1.081

Baciadonna L, Nawroth C, McElligott AG (2016) Judgement bias in goats (*Capra hircus*): investigating the effects of human grooming. PeerJ 4:e2485. https://doi.org/10.7717/peerj.2485

Bahi A, Dreyer J-L (2019) Dopamine transporter (DAT) knockdown in the nucleus accumbens improves anxiety-and depression-related behaviors in adult mice. Behav. Brain Res. 359:104-115. doi: 10.1016/j.bbr.2018.10.028

Baker KC, Aureli F (1997) Behavioural indicators of anxiety: an empirical test in chimpanzees. Behaviour 134(13-14):1031-1050. https://doi.org/10.1163/156853997X00386

Bakhchina A V., Arutyunova KR, Sozinov AA, et al (2018) Sample entropy of the heart rate reflects properties of the system organization of behaviour. Entropy 20:. https://doi.org/10.3390/e20060449

Balzani A, Hanlon A (2020) Factors that Influence Farmers' Views on Farm Animal Welfare: A Semi-Systematic Review and Thematic Analysis Agnese. Animals 10:1524. https://doi.org/10.3390/ani10091524

Barnard RJ, Corre K, Cho H (1976) Effect of training on the resting heart rate of rats. Eur J Appl Physiol Occup Physiol 35:285–289. https://doi.org/10.1007/BF00423288

Barnard S, Wells DL, Milligan AD et al (2018) Personality traits affecting judgement bias task performance in dogs (Canis familiaris). Sci. Rep. 8:6660. https://doir.org/10.1038/s41598-018-25224-y

Barrett LF (2006) Solving the emotion paradox: Categorization and the experience of emotion. Pers. Soc. Psy. Rev. 10(1):20-46. doi: 10.1207/s15327957pspr1001_2

Barrett LF (2017) The theory of constructed emotion: an active inference account of interoception and categorization. Soc. Cogn. Affect. Neurosci. 12(1):1-23. https://doi.org/10.1093/scan/nsw154

Barton DA, Esler MD, Dawood T et al (2008) Elevated brain serotonin turnover in patients with depression: effect of genotype and therapy. Arch. Gen. Psychiatry 65(1):38-46. doi: 10.1001/archgenpsychiatry.2007.11

Barton EP, Donaldson SL, Ross M, Albright JL (1973) Social rank and social index as related to age, body weight and milk production in dairy cows. Proc Indiana Acad Sci 83:473–477. https://doi.org/10.1016/s0002-8223(99)00212-6

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:. https://doi.org/10.18637/jss.v067.i01

Bateson M, Matheson SM (2007) Performance on a categorisation task suggests that removal of environmental enrichment induces "pessimism" in captive European starlings (Sturnus vulgaris). Anim Welf Potters Bar then Wheathampstead 16:33–36

Bateson M, Nettle D (2015) Development of a cognitive bias methodology for measuring low mood in chimpanzees. PeerJ 3:e998. https://doi.org/10.7717/peerj.998. eCollection 2015

Bateson M (2016) Cumulative stress in research animals: telomere attrition as a biomarker in a welfare context? BioEssays 38(2):201-212. https://doi.org/10.1002/bies.201500127

Battini M, Vieira A, Barbieri S, et al (2014) Invited review: Animal-based indicators for on-farm welfare assessment for dairy goats. J Dairy Sci 97:6625–6648. https://doi.org/10.3168/jds.2013-7493

Becker J, Reist M, Steiner A (2014) Factors influencing the attitudes of cattle veterinarians, farmers, and claw trimmers towards the pain associated with the treatment of sole ulcers and the sensitivity to pain of dairy cows. Vet J 200:38–43. https://doi.org/10.1016/j.tvjl.2014.01.016

Bekoff M (2000) Animal Emotions: Exploring Passionate NaturesCurrent interdisciplinary research provides compelling evidence that many animals experience such emotions as joy, fear, love, despair, and grief—we are not alone. BioScience 50(10):861-870

Belay TK, Svendsen M, Kowalski ZM, Ådnøy T (2017) Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows. J Dairy Sci 100:6298–6311. https://doi.org/10.3168/jds.2016-12458

Berkowitz L (2000) Causes and consequences of feelings. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9780511606106

Berridge KC (1999) Pleasure, pain, desire, and dread: Hidden core processes of emotion.

Berridge KC (2018) Evolving concepts of emotion and motivation. Front. Psychol. 9. https://doi.org/10.3389/fpsyg.2018.01647

Berridge KC, Kringelbach ML (2008) Affective neuroscience of pleasure: reward in humans and animals. Psychopharmacology 199(3):457-480. https://doi.org/ 10.1007/s00213-008-1099-6

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Rev. 28(3):309-369. https://doi.org/10.1016/S0165-0173(98)00019-8

Bethell EJ (2015) A “how-to” guide for designing judgement bias studies to assess captive animal welfare. J. Appl. Anim. Welf. Sci. 18(sup1):S18-S42. https://doi.org/ 10.1080/10888705.2015.1075833

Bethell EJ, Holmes A, MacLarnon A, Semple, S (2012) Evidence that emotion mediates social attention in rhesus macaques. PLoS One 7(8):e44387. https://doi.org/10.1371/journal.pone.0044387

Bethell EJ, Holmes A, MacLarnon A, Semple S (2016) Emotion evaluation and response slowing in a non-human primate: New directions for cognitive bias measures of animal emotion? Behav Sci (Basel) 6:. https://doi.org/10.3390/bs6010002

Bethell EJ, Koyama NF (2015) Happy hamsters? Enrichment induces positive judgement bias for mildly (but not truly) ambiguous cues to reward and punishment in *Mesocricetus auratus*. R Soc Open Sci 2:140399. https://doi.org/10.1098/rsos.140399

Bhattacharyya MR, Whitehead DL, Rakhit R, Steptoe A (2008) Depressed mood, positive affect, and heart rate variability in patients with suspected coronary artery disease. Psychosom Med 70:1020–1027. https://doi.org/10.1097/PSY.0b013e318189afcc

Blanchette I, Richards A (2010) The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. Cogn Emot 24:561–595. https://doi.org/10.1080/02699930903132496

Bliss-Moreau E (2017) Constructing nonhuman animal emotion. Curr. Opin. Psychol. 17:184-188. doi: 10.1016/j.copsyc.2017.07.011

Bliss-Moreau E, Rudebeck PH (2021) Animal models of human mood. Neurosci Biobehav Rev 120:574–582. https://doi.org/10.1016/j.neubiorev.2020.06.024

Bloch M, Daly RC, Rubinow DR (2003) Endocrine factors in the etiology of postpartum depression. Compr Psychiatry 44:234–246. https://doi.org/10.1016/S0010-440X(03)00034-8

Blokhuis HJ, Veissier I, Miele M, Jones B (2019) Safeguarding farm animal welfare. In: Sustainability Certification Schemes in the Agricultural and Natural Resource Sectors. pp 137–153. https://doi.org/ 10.4324/9780203701737

Boehlke C, Zierau O, Hannig C (2015) Salivary amylase - The enzyme of unspecialized euryphagous animals. Arch Oral Biol 60:1162–1176. https://doi.org/10.1016/j.archoralbio.2015.05.008

Boissy A, Aubert A, Désiré L et al (2011) Cognitive sciences to relate ear postures to emotions in sheep. Anim. Welf. 20(1):47

Boissy A, Bouissou MF (1995) Assessment of individual differences in behavioural reactions of heifers exposed to various fear-eliciting situations. Appl Anim Behav Sci 46:17–31. https://doi.org/10.1016/0168-1591(95)00633-8

Boissy A, Manteuffel G, Jensen MB et al (2007) Assessment of positive emotions in animals to improve their welfare. Physiol. Behav. 92(3):375-397. https://doi.org/10.1016/j.physbeh.2007.02.003

Bonafos L, Simonin D, Gavinelli A (2010) Animal welfare: European legislation and future perspectives. J Vet Med Educ 37:26–29. https://doi.org/10.3138/jvme.37.1.26

Bowen A, Bowen R, Balbuena L, Muhajarine N (2012) Are Pregnant and Postpartum Women Moodier? Understanding Perinatal Mood Instability. J Obstet Gynaecol Canada 34:1038–1042.

https://doi.org/10.1016/S1701-2163(16)35433-0

Breslow NE, Clayton DG (1993) Approximate Inference in Generalized Linear Mixed Models. J Am Stat Assoc 88:9–25. https://doi.org/10.1080/01621459.1993.10594284

Briefer E (2012) Vocal expression of emotions in mammals: mechanisms of production and evidence. J. Zool. 288(1):1-20.

Briefer EF (2018) Vocal contagion of emotions in non-human animals. Proceedings of the Royal Society B: Biological Sciences 285(1873):20172783. https://doi.org/10.1098/rspb.2017.2783

Briefer EF, Tettamanti F, McElligott AG (2015) Emotions in goats: mapping physiological, behavioural and vocal profiles. Anim. Behav. 99:131-143. https://doi.org/10.1016/j.anbehav.2014.11.002

Briefer Freymond S, Briefer EF, Zollinger A, et al (2014) Behaviour of horses in a judgement bias test associated with positive or negative reinforcement. Appl Anim Behav Sci 158:34–45. https://doi.org/10.1016/j.applanim.2014.06.006

Brilot BO, Asher L, Bateson M (2010) Stereotyping starlings are more “pessimistic.” Anim Cogn 13:721–731. https://doi.org/10.1007/s10071-010-0323-z

Brilot BO, Bateson M (2012) Water bathing alters threat perception in starlings. Biol. Lett.:rsbl20111200. https://doi.org/10.1098/rsbl.2011.1200

Brooks ME, Kristensen K, van Benthem KJ, et al (2017) Modeling zero-inflated count data with glmmTMB. bioRxiv 132753. https://doi.org/10.1101/132753

Broom (2017) Animal Welfare in European Union. Anim Welf Eur Union 4–78

Broom DM, Johnson KG (1993) Assessing welfare: Short-term responses. in Stress and Animal Welfare. D. M. Broom and K. G. Johnson, ed. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/978-3-030-32153-6_5

Brudzynski SM, Pniak A (2002) Social contacts and production of 50-kHz short ultrasonic calls in adult rats. J. Comp. Psychol. 116(1):73. doi: 10.1037/0735-7036.116.1.73

Brydges NM, Leach M, Nicol K, et al (2011) Environmental enrichment induces optimistic cognitive bias in rats. Anim Behav 81:169–175. https://doi.org/10.1016/j.anbehav.2010.09.030

Bučková K, Špinka M, Hintze S (2019) Pair housing makes calves more optimistic. Sci Rep 9:1–9. https://doi.org/10.1038/s41598-019-56798-w

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

Burgdorf J, Kroes RA, Moskal JR et al (2008) Ultrasonic vocalizations of rats (Rattus norvegicus) during mating, play, and aggression: Behavioral concomitants, relationship to reward, and self-administration of playback. J. Comp. Psychol. 122(4):357. doi: 10.1037/a0012889

Burgdorf J, Panksepp J,Moskal JR (2011) Frequency-modulated 50kHz ultrasonic vocalizations: A tool for uncovering the molecular substrates of positive affect. Neurosci. Biobehav. Rev. 35(9):1831-1836. doi: 10.1016/j.neubiorev.2010.11.011

Burman O, McGowan R, Mendl M et al (2011) Using judgement bias to measure positive affective state in dogs. Appl. Anim. Behav. Sci. 132(3):160-168. https://doi.org/10.1016/j.applanim.20011.04.001

Burman OH, Mendl MT (2018) A novel task to assess mood congruent memory bias in non-human animals. J. Neurosci. Methods 308:269-275. https://doi/org/10.1016/j.jneumeth.2018.07.003

Burman OHP, Parker RMA, Paul ES, Mendl M (2008) Sensitivity to reward loss as an indicator of animal emotion and welfare. Biol. Lett. 4(4):330-333. https://doi.org/ 10.1098/rsbl.2008.0113

Burman OHP, Parker RMA, Paul ES, Mendl MT (2009) Anxiety-induced cognitive bias in non-human animals. Physiol Behav 98:345–350. https://doi.org/10.1016/j.physbeh.2009.06.012

Burnett TA, Madureira AML, Silper BF, et al (2014) Short communication: Factors affecting hair cortisol concentrations in lactating dairy cows. J Dairy Sci 97:7685–7690. https://doi.org/10.3168/jds.2014-8444

Cabanac M, Gosselin F (1993) Emotional fever in the lizard Callopistes maculatus (Teiidae). Anim. Behav. 46:200-202. https://doi.org/ 10.1006/anbe.1993.1178

Cacioppo JT, Tassinary LG, Berntson G (2007) Handbook of Psychophysiology. Cambridge University Press, Cambridge, UK

Campbell DLM, Taylor PS, Hernandez CE, et al (2019) An attention bias test to assess anxiety states in laying hens. PeerJ 2019:1–27. https://doi.org/10.7717/peerj.7303

Carreras R, Mainau E, Rodriguez P, et al (2015) Cognitive bias in pigs: Individual classification and consistency over time. J Vet Behav 10:577–581. https://doi.org/10.1016/j.jveb.2015.09.001

Casal N, Manteca X, Escribano D, et al (2017) Effect of environmental enrichment and herbal compound supplementation on physiological stress indicators (chromogranin A, cortisol and tumour necrosis factor-α) in growing pigs. Animal 11:1228–1236. https://doi.org/10.1017/S1751731116002561

Chalmers DJ (1995) The puzzle of conscious experience. Sci. Am. 273(6):80-86

Chan JSY, Liu G, Liang D, et al (2019) Special Issue–Therapeutic Benefits of Physical Activity for Mood: A Systematic Review on the Effects of Exercise Intensity, Duration, and Modality. J Psychol Interdiscip Appl 153:102–125. https://doi.org/10.1080/00223980.2018.1470487

Chittka L, Skorupski P, Raine NE (2009) Speed-accuracy tradeoffs in animal decision making. Trends Ecol Evol 24:400–407. https://doi.org/10.1016/j.tree.2009.02.010

Cingi CC, Baser DF, Karafakioglu YS, Fidan AF (2012) Stress Response in Dairy Cows Relatd to Rectal Examination. Acta Sci Vet 40:1–7

Cisler JM, Koster EHW (2010) Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. Clin Psychol Rev 30:203–216. https://doi.org/10.1016/j.cpr.2009.11.003

Clarkson JM, Leach MC, Flecknell PA, Rowe C (2020) Negative mood affects the expression of negative but not positive emotions in mice: Mood affects discrete emotions in mice. Proc R Soc B Biol Sci 287:1–9. https://doi.org/10.1098/rspb.2020.1636rspb20201636

Clay-Warner J, Robinson DT (2015) Infrared thermography as a measure of emotion response. Emot Rev 7:157–162. https://doi.org/10.1177/1754073914554783

Clore GL, Schwarz N, Conway M (1994) Affective causes and consequences of social information processing. in Handbook of Social Cognition. Volume 1: Basic Processes. R. Wyer and T. Srull, ed. Psychology Press, London, UK

Coleman K (2012) Individual differences in temperament and behavioral management practices for nonhuman primates. Appl Anim Behav Sci 138:106–113. https://doi.org/10.1016/j.applanim.2011.08.002

Coleman K, Pierre PJ (2014) Assessing anxiety in nonhuman primates. Ilar J. 55(2):333-346. https://doi.org/10.1093/ilar/ilu019

Comrey AL, Lee HB (1992) A first course in factor analysis, 2nd ed. Psychology Press

Cooke RF, Bohnert DW, Reis MM, Cappellozza BI (2013) Wolf presence in the ranch of origin: Impacts on temperament and physiological responses of beef cattle following a simulated wolf encounter. J Anim Sci 91:5905–5911. https://doi.org/10.2527/jas2013-6777

Cooke RF, Schubach KM, Marques RS, et al (2017) Effects of temperament on physiological, productive, and reproductive responses in beef cows. J Anim Sci 95:1. https://doi.org/10.2527/jas2016.1098

Cornish A, Raubenheimer D, McGreevy P (2016) What we know about the public's level of concern for farm animal welfare in food production in developed countries. Animals 6(11):74. https://doi.org/10.3390/ani6110074

Crane AL, Ferrari MCO (2017) Patterns of predator neophobia: A meta-analytic review. Proc R Soc B Biol Sci 284:. https://doi.org/10.1098/rspb.2017.0583

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

Crump A, Arnott G, Bethell E (2018) Affect-Driven Attention Biases as Animal Welfare Indicators: Review and Methods. Animals 8:136. https://doi.org/10.3390/ani8080136

Crump A, Jenkins K, Bethell EJ, et al (2021) Optimism and pasture access in dairy cows. Sci Rep 11:4882. https://doi.org/10.1038/s41598-021-84371-x

Cussen VA, Mench JA (2014) Personality predicts cognitive bias in captive psittacines, Amazona amazonica. Anim. Behav. 89:123-130. https://doi.org/10.1016/j.anbehav.2013.12.022

CVB (2012). CVB. Cvb tabellenboek veevoeding [feedstuff table 2012]. Centraal Veevoeder Bureau Lelystad, the Netherlands (2012)

Da Estrela C, McGrath J, Booij L, Gouin JP (2021) Heart Rate Variability, Sleep Quality, and Depression in the Context of Chronic Stress. Ann Behav Med 55:155–164. https://doi.org/10.1093/abm/kaaa039

Dalla Costa E, Minero M, Lebelt D et al (2014) Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS One 9(3):e92281. https://doi.org/10.1371/journal.pone.0092281

Damasio A, Carvalho GB (2013) The nature of feelings: evolutionary and neurobiological origins. Nature Rev. Neurosci. 14(2):143. https://doi.org/10.1038/nrn3403

Damasio A, Stella F, Helderman S (2010) Self Comes to Mind: Constructing the Conscious Brain. Pantheon Books, New York, USA

Damasio AR, Grabowski TJ, Bechara A et al (2000) Subcortical and cortical brain activity during the feeling of self-generated emotions. Nature Neurosci. 3(10):1049. https://doi.org/10.1038/79871

Dantzer R, Arnone M, Mormède P (1980) Effects of frustration on behaviour and plasma corticosteroid levels in pigs. Physiol. Behav. 24:1 https://doi.org/10.1016/0031-9384(80)90005-0

Daros RR, Costa JHC, von Keyserlingk MAG, et al (2014) Separation from the dam causes negative judgement bias in dairy calves. PLoS One 9:5–9. https://doi.org/10.1371/journal.pone.0098429

Darwin C (1872) The expression of the emotions in man and animals. 3rd ed. John Murray, London, UK

Davidson KW, Mostofsky E, Whang W (2010) Don't worry, be happy: Positive affect and reduced 10-year incident coronary heart disease: The Canadian nova scotia health survey. Eur Heart J 31:1065–1070. https://doi.org/10.1093/eurheartj/ehp603

Dawkins MS (1988) Behavioural deprivation: a central problem in animal welfare. Appl. Anim. Behav. Sci. 20(3-4):209-225. https://doi.org/10.1016/0168-1591(88)90047-0

Dawkins MS (1990) From an animal's point of view: motivation, fitness, and animal welfare. Behav. Brain Sci. 13(1):1-9. https://doi.org/10.1017/S0140525X00077104

Dawkins MS (2000) Animal minds and animal emotions. Am. Zool. 40(6):883-888. https://doi.org/10.1093/icb/40.6.883

Dawkins MS (2006) Through animal eyes: what behaviour tells us. Appl. Anim. Behav. Sci. 100(1):4-10. https://doi.org/10.1016/j.applanim.2006.04.010

De la Fuente J, González-Torres MC, Artuch-Garde R, et al (2021) Resilience as a Buffering Variable Between the Big Five Components and Factors and Symptoms of Academic Stress at University. Front Psychiatry 12:. https://doi.org/10.3389/fpsyt.2021.600240

Denton D, McKinley M, Farrell M, Egan G (2009) The role of primordial emotions in the evolutionary origin of consciousness. Conscious. Cogn. 18(2):500-514. https://doi.org/10.1016/j.concog.2008.06.009

Des Roches ADB, Richard-Yris M-A, Henry S et al (2008) Laterality and emotions: Visual laterality in the domestic horse (Equus caballus) differs with objects' emotional value. Physiol. Behav. 94(3):487-490. https://doi.org/10.1016/j.physbeh.2008.03.002

Descovich K, Wathan JW, Leach MC et al (2017) Facial expression: An under-utilized tool for the assessment of welfare in mammals. ALTEX-Altern. Anim. Ex. 34(3):409-429. https://doi.org/10.14573/altex.1607161

Désiré L, Boissy A, Veissier I (2002) Emotions in farm animals: a new approach to animal welfare in applied ethology. Behav. Process. 60(2):165-180. https://doi.org/10.1016/S0376-6357(02)00081-5

Désiré L, Veissier I, Després G, Boissy A (2004) On the Way to Assess Emotions in Animals: Do Lambs (Ovis aries) Evaluate an Event Through Its Suddenness, Novelty, or Unpredictability? J Comp Psychol Comp 118:363–374. https://doi.org/10.1037/0735-7036.118.4.363

Destrez A, Deiss V, Belzung C, et al (2012) Does reduction of fearfulness tend to reduce pessimistic-like judgement in lambs? Appl Anim Behav Sci 139:233–241. https://doi.org/10.1016/j.applanim.2012.04.006

Destrez A, Deiss V, Leterrier C, et al (2014) Repeated exposure to positive events induces optimistic-like judgment and enhances fearfulness in chronically stressed sheep. Appl Anim Behav Sci 154:30–38. https://doi.org/10.1016/j.applanim.2014.01.005

Destrez A, Deiss V, Lévy F, et al (2013) Chronic stress induces pessimistic-like judgement and learning deficits in sheep. Appl Anim Behav Sci 148:28–36. https://doi.org/10.1016/j.applanim.2013.07.016

Dettenborn L, Muhtz C, Skoluda N, et al (2012) Introducing a novel method to assess cumulative steroid concentrations: Increased hair cortisol concentrations over 6 months in medicated patients with depression. Stress 15:348–353. https://doi.org/10.3109/10253890.2011.619239

Dettmer AM, Novak MA, Meyer JS, Suomi SJ (2014) Population density-dependent hair cortisol concentrations in rhesus monkeys (Macaca mulatta). Psychoneuroendocrinology 42:59–67. https://doi.org/10.1016/j.psyneuen.2014.01.002

D'Ettorre P, Carere C, Demora L et al (2017) Individual differences in exploratory activity relate to cognitive judgement bias in carpenter ants. Behav. Process. 134:63-69. https://doi.org/10.1016/j.beproc.2016.09.008

De Vere AJ, Kuczaj SA (2016) Where are we in the study of animal emotions? Wiley Interdiscip. Rev. Cogn. Sci. 7(5):354-362. https://doi.org/ 10.1002/wcs.1399

Devries TJ, Keyserlingk MAG Von (2006) Feed Stalls Affect the Social and Feeding Behavior of Lactating Dairy Cows Feed Stalls Affect the Social and Feeding Behavior. J Dairy Sci 89:3522–3531. https://doi.org/10.3168/jds.S0022-0302(06)72392-X

De Waal FB (1999) Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. Philosophical Topics 27(1):255-280. https://www.jstor.org/stable/43154308

de Waal FB (2008) Putting the altruism back into altruism: the evolution of empathy. Annu. Rev. Psychol. 59:279-300. https://doi.org/10.1146/annurev.psych.59.103006.093625

de Waal F (2011) What is an animal emotion? Ann. N. Y. Acad. Sci. 1224(1):191-206. https://doi.org/ 10.1111/j.1749-6632.2010.05912.x

Dich N, Doan SN, Kivimäki M et al (2014) A non-linear association between self-reported negative emotional response to stress and subsequent allostatic load: Prospective results from the Whitehall II cohort study. Psychoneuroendocrinology 49:54-61. https://doi.org/ 10.1016/j.psyneuen.2014.07.001

Diener E (2000) Subjective well-being: The science of happiness and a proposal for a national index. Am. Psychol. 55(1):34-43. https://doi.org/10.1037/0003-066X.55.1.34

Diener E, Sandvik E, Pavot W (2009) Happiness is the Frequency, Not the Intensity, of Positive Versus Negative Affect. in Assessing Well-Being: The Collected Works of Ed Diener. E. Diener, ed. Springer, Dordrecht, Netherlands

Di Giminiani P, Brierley VL, Scollo A et al (2016) The assessment of facial expressions in piglets undergoing tail docking and castration: toward the development of the piglet grimace scale. Frontiers in veterinary science 3:100. https://doi.org/10.3389/fvets.2016.00100

Dimitriev DA, Saperova E V., Dimitriev AD (2016) State Anxiety and Nonlinear Dynamics of Heart Rate Variability in Students. PLoS One 11:e0146131. https://doi.org/10.1371/journal.pone.0146131

Ding X, Yang S, Li W, et al (2014) The potential biomarker panels for identification of Major Depressive Disorder (MDD) patients with and without Early Life Stress (ELS) by metabonomic analysis. PLoS One 9:e97479. https://doi.org/10.1371/journal.pone.0097479

Ditzen B, Heinrichs M (2014) Psychobiology of social support: the social dimension of stress buffering. Restorative neurology and neuroscience 32(1):149-162. https://doi.org/ 10.3233/RNN-139008

Dolan RJ (2002) Neuroscience and psychology: Emotion, cognition, and behavior. Science (80-) 298:1191–1194. https://doi.org/10.1126/science.1076358

Douglas C, Bateson M, Walsh C et al (2012) Environmental enrichment induces optimistic cognitive biases in pigs. Appl. Anim. Behav. Sci. 139(1-2):65-73. https://doi.org/10.1016/j.applanim.2012.02.018

Douphrate DI, Rosecrance JC, Stallones L, et al (2009) Livestock-handling injuries in agriculture: An analysis of colorado workers compensation data. Am J Ind Med 52:391–407. https://doi.org/10.1002/ajim.20686

Doyle RE, Fisher AD, Hinch GN et al (2010a) Release from restraint generates a positive judgement bias in sheep. Appl. Anim. Behav. Sci. 122(1):28-34. https://doi.org/ 10.1016/j.applanim.2009.11.003

Doyle RE, Hinch GN, Fisher AD et al (2011) Administration of serotonin inhibitor p-Chlorophenylalanine induces pessimistic-like judgement bias in sheep. Psychoneuroendocrinology 36(2):279-288. https://doi.org/10.1016/j.psyneuen.2010.07.018

Doyle RE, Vidal S, Hinch GN et al (2010b) The effect of repeated testing on judgement biases in sheep. Behav. Process. 83(3):349-352. https://doi.org/ 10.1016/j.beproc.2010.01.019

Duncan IJH (1996) Animal welfare defined in terms of feelings. Acta Agric. Scand. Sect. A-Anim. Sci. 27:29-35

Duncan IJH (2004) A Concept of Welfare Based on Feelings. In: The Well-Being of Farm Animals. pp 85–101. https://doi.org/10.1002/9780470344859.ch5

Duncan IJH (2006) The changing concept of animal sentience. Appl. Anim. Behav. Sci. 100(1–2):11-19. https://doi.org/10.1016/j.applanim.2006.04.011

Dunsmoor JE, Kroes MCW, Braren SH, Phelps EA (2017) Threat intensity widens fear generalization gradients. Behav Neurosci 131:168–175. https://doi.org/10.1037/bne0000186

Edgar J, Held S, Paul E et al (2015) Social buffering in a bird. Anim. Behav. 105:11-19. https://doi.org/ 10.1016/j.anbehav.2015.04.007

Edgar JL, Lowe JC, Paul ES, Nicol CJ (2011) Avian maternal response to chick distress. Proc. R. Soc. B-Biol. Sci.:rspb20102701. https://doi.org/10.1098/rspb.2010.2701

Egorov AI, Griffin SM, Converse RR, et al (2017) Vegetated land cover near residence is associated with reduced allostatic load and improved biomarkers of neuroendocrine, metabolic and immune functions. Environ Res 158:508–521. https://doi.org/10.1016/j.envres.2017.07.009

Ekman P (1992) Are there basic emotions? Psychology Review 99:550-553. https://doi.org/ 10.1037/0033-295x.99.3.550

Ekman P (2016) What scientists who study emotion agree about. Perspect. Psychol. Sci. 11(1):31-34. https://doi.org/ 10.1177/1745691615596992

Eldar E, Rutledge RB, Dolan RJ, Niv Y (2016) Mood as Representation of Momentum. Trends Cogn Sci 20:15–24. https://doi.org/10.1016/j.tics.2015.07.010

Elliot AJ, Eder AB, Harmon-Jones E (2013) Approach-avoidance motivation and emotion: Convergence and divergence. Emot Rev 5:308–311. https://doi.org/10.1177/1754073913477517

Ellsworth PC (2013) Appraisal theory: Old and new questions. Emot. Rev. 5(2):125-131. https://doi.org/10.1177/1754073912463617

Enkel T, Gholizadeh D, von Bohlen Und Halbach O, et al (2010) Ambiguous-cue interpretation is biased under stress- and depression-like states in rats. Neuropsychopharmacology 35:1008–15. https://doi.org/10.1038/npp.2009.204

European Commission (2019) EU Agricultural Outlook For Markets and Income 2019-2030. Publications Office of the European Union

Evans A, Bridgewater B, Liu Q, et al (2014) High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics. J Postgenomics Drug Biomark Dev 04: https://doi.org/10.4172/2153-0769.1000132

Eysenck MW, Mogg K, May J, et al (1991) Bias in Interpretation of Ambiguous Sentences Related to Threat in Anxiety. J Abnorm Psychol 100:144–150. https://doi.org/10.1037/0021-843X.100.2.144

Fabbro F, Aglioti SM, Bergamasco M et al (2015) Evolutionary aspects of self-and world consciousness in vertebrates. Front. Hum. Neurosci. 9:1-16. https://doi.org/10.3389/fnhum.2015.00157

Fairhurst GD, Frey MD, Reichert JF, et al (2011) Does environmental enrichment reduce stress? An integrated measure of corticosterone from feathers provides a novel perspective. PLoS One 6:1–10. https://doi.org/10.1371/journal.pone.0017663

Fanselow MS (2018) Emotion, motivation and function. Curr. Opin. Behav. Sci. 19:105-109. https://doi.org/10.1016/j.cobeha.2017.12.013

Faragó T, Takács N, Miklósi Á,Pongrácz P (2017) Dog growls express various contextual and affective content for human listeners. R. Soc. Open Sci. 4(5):170134. https://doi.org/10.1098/rsos.170134

Faraji J, Karimi M, Soltanpour N, et al (2018) Oxytocin-mediated social enrichment promotes longer telomeres and novelty seeking. Elife 1–19. https://doi.org/ 10.7554/eLife.40262

Faust L, Feldman K, Mattingly SM, et al (2020) Deviations from normal bedtimes are associated with short-term increases in resting heart rate. npj Digit Med 3:1–9. https://doi.org/10.1038/s41746-020-0250-6

Faustino AI, Oliveira GA, Oliveira RF (2015) Linking appraisal to behavioral flexibility in animals: Implications for stress research. Front Behav Neurosci 9:1–7. https://doi.org/10.3389/fnbeh.2015.00104

Feigley DA, Spear NE (1970) Effect of age and punishment condition on long-term retention by the rat of active- and passive-avoidance learning. J Comp Physiol Psychol 73:515–526. https://doi.org/10.1037/h0030234

Field T (2019) Pediatric Massage Therapy Research: A Narrative Review. Children 6:78. https://doi.org/10.3390/children6060078

Finkemeier MA, Langbein J, Puppe B (2018) Personality research in mammalian farm animals: Concepts, measures, and relationship to welfare. Front Vet Sci 5:1–15. https://doi.org/10.3389/fvets.2018.00131

Finkenwirth C, van Schaik C, Ziegler TE, Burkart JM (2015) Strongly bonded family members in common marmosets show synchronized fluctuations in oxytocin. Physiol. Behav. 151:246-251. https://doi.org/ 10.1016/j.physbeh.2015.07.034

Finlayson K, Lampe JF, Hintze S et al (2016) Facial Indicators of Positive Emotions in Rats. PLoS One 11(11):e0166446. https://doi.org/10.1371/journal.pone.0166446

Fitness J, Fletcher GJ (1993) Love, hate, anger, and jealousy in close relationships: A prototype and cognitive appraisal analysis. J. Pers. Soc. Psychol. 65(5):942. https://doi.org/ 10.1037//0022-3514.65.5.942

Flaherty CF (1999) Incentive relativity. Vol. 15. Cambridge University Press, Cambridge, UK

Ford L, Kennedy AD, Goodman KD, et al (2020) Precision of a Clinical Metabolomics Profiling Platform for Use in the Identification of Inborn Errors of Metabolism. J Appl Lab Med 5:342–356. https://doi.org/10.1093/jalm/jfz026

Forkman B, Boissy A, Meunier-Salauen MC et al (2007) A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiol. Behav. 92(3):340-374. https://doi.org/10.1016/j.physbeh.2007.03.016

Foris B, Zebunke M, Id JL, Id NM (2018) Evaluating the temporal and situational consistency of personality traits in adult dairy cattle. PLoS One 1–16. https://doi.org/10.1371/journal.pone.0204619

Fox MF, Faver CA (1984) Independance and Cooperation in Research. The Motivations and Costs of Collaboration. J Higher Educ 55:347–359. https://doi.org/10.2307/1981888

Fox RA, Millam JR (2007) Novelty and individual differences influence neophobia in orange-winged Amazon parrots (Amazona amazonica). Appl Anim Behav Sci 104:107–115. https://doi.org/10.1016/j.applanim.2006.04.033

Fraser D (2008) Understanding animal welfare. Acta Vet. Scand. 50(1):S1

Fraser D (2009) Animal behaviour, animal welfare and the scientific study of affect. Appl Anim Behav Sci 118:108–117. https://doi.org/10.1016/j.applanim.2009.02.020

Fregonesi JA, Tucker CB, Weary DM (2007) Overstocking Reduces Lying Time in Dairy Cows. J Dairy Sci 90:3349–3354. https://doi.org/10.3168/jds.2006-794

Frijda NH (1986) The emotions. Cambridge University Press, Cambridge, UK

Fureix C, Meagher RK (2015) What can inactivity (in its various forms) reveal about affective states in non-human animals? A review. Appl. Anim. Behav. Sci. 171:8-24. https://doi.org/10.1016/j.applanim.2015.08.036

Fustini M, Galeati G, Gabai G, et al (2017) Overstocking dairy cows during the dry period affects dehydroepiandrosterone and cortisol secretion. J Dairy Sci 100:620–628. https://doi.org/10.3168/jds.2016-11293

Garcia J, Koelling RA (1966) Relation of cue to consequence in avoidance learning. Psychon Sci 4:123–124. https://doi.org/10.3758/bf03342209.

Geisler FCM, Vennewald N, Kubiak T, Weber H (2010) The impact of heart rate variability on subjective well-being is mediated by emotion regulation. Pers Individ Dif 49:723–728. https://doi.org/10.1016/j.paid.2010.06.015

Gerber M, Brand S, Lindwall M, et al (2012) Concerns regarding hair cortisol as a biomarker of chronic stress in exercise and sport science. J Sport Sci Med 11:571–581

Ghassemi Nejad J, Kim BW, Lee BH, Sung K Il (2017) Coat and hair color: hair cortisol and serotonin levels in lactating Holstein cows under heat stress conditions. Anim Sci J 88:190–194. https://doi.org/10.1111/asj.12662

Gibbons JM, Lawrence AB, Haskell MJ (2010) Measuring sociability in dairy cows. Appl Anim Behav Sci 122:84–91. https://doi.org/10.1016/j.applanim.2009.11.011

Gilboa E, Revelle W (1994) Personality and the structure of affective responses. Pages 135-159 in Emotions: Essays on emotion theory. S. H. M. van Goozen, N. E. Van de Poll, and J. A. Sergeant, ed. Psychology Press, New York, USA

Gill NP, Bos EH, Wit EC, de Jonge P (2017) The association between positive and negative affect at the inter-and intra-individual level. Pers. Individ. Dif. 105:252-256. https://doi.org/10.1016/j.paid.2016.10.002

Glatzer W, Gulyas J (2014) Affect Balance Scale. in Encyclopedia of Quality of Life and Well-Being Research. A. C. Michalos, ed. Springer, Netherlands

Goldansaz SA, Guo AC, Sajed T et al (2017) Livestock metabolomics and the livestock metabolome: A systematic review. PLoS One 12(5):e0177675. https://doi.org/10.1371/journal.pone.0177675

Goldsmith HH, Buss AH, Plomin R, et al (1987) Roundtable Four Approaches. Child Dev 5:505–29

Goumon S, Špinka M (2016) Emotional contagion of distress in young pigs is potentiated by previous exposure to the same stressor. Anim. Cogn. 19(3):501-511. https://doi.org/ 10.1007/s10071-015-0950-5.

Graham-Engeland JE, Sin NL, Smyth JM et al (2018) Negative and positive affect as predictors of inflammation: Timing matters. Brain Behav. Immun 74:222-230. https://doi.org/10.1016/j.bbi.2018.09.011

Grandi LC, Ishida H (2015) The physiological effect of human grooming on the heart rate and the heart rate variability of laboratory non-human primates: A pilot study in male rhesus monkeys. Front Vet Sci 2:1–9. https://doi.org/10.3389/fvets.2015.00050

Graulich DM, Kaiser S, Sachser N,Richter SH (2016) Looking on the bright side of bias—Validation of an affective bias test for laboratory mice. Appl. Anim. Behav. Sci. 181:173-181. https://doi.org/ 10.1016/j.applanim.2016.05.011

Graunke KL, Nürnberg G, Repsilber D, et al (2013) Describing Temperament in an Ungulate: A Multidimensional Approach. PLoS One 8:1–12. https://doi.org/10.1371/journal.pone.0074579

Gray JA (1987) Perspectives on anxiety and impulsivity: A commentary. J. Res. Person. 21(4):493-509. https://doi.org/10.1016/0092-6566(87)90036-5

Gray JA (1990) Brain systems that mediate both emotion and cognition. Cogn. Emot. 4(3):269-288. https://doi.org/10.1080/02699939008410799

Gray JA, McNaughton N (2000) The Neuropsychology of Anxiety, 2nd ed. Oxford University Press, New York

Green T, Mellor D (2011) Extending ideas about animal welfare assessment to include 'quality of life' and related concepts. N Z Vet J 59:263–271. https://doi.org/10.1080/00480169.2011.610283

Grigoriev I, Nikolaeva L, Artamonov I (2003) Protein content of human saliva in various psycho-emotional states. Biochemistry (moscow) 68(4):405-406. https://doi.org/ 10.1023/A:1023695729019

Grippo AJ, Johnson AK (2009) Stress, depression and cardiovascular dysregulation: a review of neurobiological mechanisms and the integration of research from preclinical disease models. Stress 12(1):1-21. https://doi.org/10.1080/10253890802046281

Groenewold NA, Opmeer EM, de Jonge P, et al (2013) Emotional valence modulates brain functional abnormalities in depression: Evidence from a meta-analysis of fMRI studies. Neurosci Biobehav Rev 37:152–163. https://doi.org/10.1016/j.neubiorev.2012.11.015

Grönlund U, Hallén Sandgren C, Waller KP (2005) Haptoglobin and serum amyloid A in milk from dairy cows with chronic sub-clinical mastitis. Vet Res 36(2), 191-198 36:191–198. https://doi.org/10.1051/vetres

Guilliams TG, Edwards L (2010) Chronic stress and the HPA axis: Clinical assessment and therapeutic considerations. Stand 9:1–12

Gutmann AK, Spinka M, Winckler C (2015) Long-term familiarity creates preferred social partners in dairy cows. Appl Anim Behav Sci 169:1–8. https://doi.org/10.1016/j.applanim.2015.05.007

Guttman N, Kalish HI (1956). Discriminability and stimulus generalization. Journal of Experimental Psychology, 51 : 79–88. https://doi.org/10.1037/h0046219

Gygax L (2014) The A to Z of statistics for testing cognitive judgement bias. Anim Behav 95:59–69. https://doi.org/10.1016/j.anbehav.2014.06.013

Hagen K, Langbein J, Schmied C, et al (2005) Heart rate variability in dairy cows—influences of breed and milking system. Physiol Behav 85:195–204. https://doi.org/10.1016/J.PHYSBEH.2005.03.019

Handlin L, Hydbring-Sandberg E, Nilsson A et al (2011) Short-term interaction between dogs and their owners: effects on oxytocin, cortisol, insulin and heart rate—an exploratory study. Anthrozoös 24(3):301-315. https://doi.org/10.2752/175303711X13045914865385

Harding EJ, Paul ES, Mendl M (2004) Cognitive bias and affective state. Nature 427(6972):312. https://doi.org/10.1038/427312a

Hatfield E, Cacioppo JT, Rapson RL (1994) Emotional contagion. in Studies in Emotion and Social Interaction. Cambridge University Press, Cambridge

Heathers JAJ, Quintana DS, Angus DJ, et al (unpublished) Water consumption as a source of error in the measurement of heart rate variability

Hebb DO (1946) Emotion in man and animal: an analysis of the intuitive processes of recognition. Psychol. Rev. 53(2):88. https://doi.org/10.1037/h0063033

Heffner, R. S., & Heffner, H. E. (1983). Hearing in large mammals: Horses (Equus caballus) and cattle (Bos taurus). Behavioral Neuroscience, 97(2), 299–309. https://doi.org/10.1037/0735-7044.97.2.299

Heimbürge S, Kanitz E, Tuchscherer A, Otten W (2020a) Within a hair's breadth – Factors influencing hair cortisol levels in pigs and cattle. Gen Comp Endocrinol 288:113359. https://doi.org/10.1016/j.ygcen.2019.113359

Heimbürge S, Kanitz E, Tuchscherer A, Otten W (2020b) Is it getting in the hair? – Cortisol concentrations in native, regrown and segmented hairs of cattle and pigs after repeated ACTH administrations. Gen Comp Endocrinol 295:113534. https://doi.org/10.1016/j.ygcen.2020.113534

Held SDE, Špinka M (2011) Animal play and animal welfare. Anim. Behav. 81(5):891-899. https://doi.org/10.1016/j.anbehav.2011.01.007

Hemsworth P, Mellor D, Cronin G, Tilbrook A (2015) Scientific assessment of animal welfare. N Z Vet J 63:24–30. https://doi.org/10.1080/00480169.2014.966167

Henry S, Fureix C, Rowberry R et al (2017) Do horses with poor welfare show 'pessimistic'cognitive biases? Sci. Nat. 104(1-2):1-15. https://doi.org/ 10.1007/s00114-016-1429-1

Hewson CJ (2003) Bien-être des animaux: Quelques définitions et courantes et leurs incidences. Can Vet J 44:496–499

Hinchcliffe JK, Stuart SA, Mendl M, Robinson ES (2017) Further validation of the affective bias test for predicting antidepressant and pro-depressant risk: effects of pharmacological and social manipulations in male and female rats. Psychopharmacology 234(20):3105-3116

Hinde RA (1985) Was 'The expression of the emotions'a misleading phrase? Anim. Behav. 33(3):985-992. https://doi.org/10.1016/S0003-3472(85)80032-4

Hintze S, Melotti L, Colosio S et al (2018) A cross-species judgement bias task: integrating active trial initiation into a spatial Go/No-go task. Sci. Rep. 8(1):5104

Hivert MF, Perng W, Watkins SM, et al (2015) Metabolomics in the developmental origins of obesity and its cardiometabolic consequences. J Dev Orig Health Dis 6:65–78. https://doi.org/10.1017/S204017441500001X

Hong H, Lee E, Lee IH, Lee SR (2019) Effects of transport stress on physiological responses and milk production in lactating dairy cows. Asian-Australasian J Anim Sci 32:442–451. https://doi.org/10.5713/ajas.18.0108

Horváth M, Pichová K, Košťál Ľ (2016) The effects of housing conditions on judgement bias in Japanese quail. Appl Anim Behav Sci 185:121–130. https://doi.org/10.1016/j.applanim.2016.09.007

Howarth ERI, Kemp C, Thatcher HR, et al (2021) Developing and validating attention bias tools for assessing trait and state affect in animals: A worked example with Macaca mulatta. Appl Anim Behav Sci 234:105198. https://doi.org/10.1016/j.applanim.2020.105198

Hubbard AJ, Foster MJ, Daigle CL (2021) Social dominance in beef cattle — A scoping review. Appl Anim Behav Sci 241:. https://doi.org/10.1016/j.applanim.2021.105390

Hucklebridge F, Lambert S, Clow A et al (2000) Modulation of secretory immunoglobulin A in saliva; response to manipulation of mood. Biol. Psychol. 53(1):25-35. https://doi.org/10.1016/s0301-0511(00)00040-5

Hunter LB, Haskell MJ, Langford FM, et al (2021) Heart rate and heart rate variability change with sleep stage in dairy cows. Animals 11:. https://doi.org/10.3390/ani11072095

Huzzey JM, DeVries TJ, Valois P, von Keyserlingk MAG (2006) Stocking Density and Feed Barrier Design

Affect the Feeding and Social Behavior of Dairy Cattle. J Dairy Sci 89:126–133. https://doi.org/10.3168/jds.S0022-0302(06)72075-6

Huzzey JM, Nydam DV, Grant RJ, Overton TR (2012) The effects of overstocking Holstein dairy cattle during the dry period on cortisol secretion and energy metabolism. J Dairy Sci 95:4421–4433. https://doi.org/10.3168/jds.2011-5037

Izard CE, Libero DZ, Putnam P, Haynes OM (1993) Stability of emotion experiences and their relations to traits of personality. J. Pers. Soc. Psychol. 64(5):847. https://doi.org/ 10.1037//0022-3514.64.5.847

Jones S, Neville V, Higgs L, et al (2018) Assessing animal affect: an automated and self-initiated judgement bias task based on natural investigative behaviour. Sci Rep 8:2–13. https://doi.org/10.1038/s41598-018-30571-x

Jones S, Paul ES, Dayan P, et al (2017) Pavlovian influences on learning differ between rats and mice in a counter-balanced Go/NoGo judgement bias task. Behav Brain Res 331:214–224. https://doi.org/10.1016/j.bbr.2017.05.044

Juster R-P, McEwen BS, Lupien SJ (2010) Allostatic load biomarkers of chronic stress and impact on health and cognition. Neurosci. Biobehav. Rev. 35(1):2-16. https://doi.org/ 10.1016/j.neubiorev.2009.10.002

Kalbe C, Puppe B (2010) Long-term cognitive enrichment affects opioid receptor expression in the amygdala of domestic pigs. Genes Brain Behav. 9(1):75-83. https://doi.org/ 10.1111/j.1601-183X.2009.00536.x

Kappel S, Mendl MT, Barrett DC et al (2017) Lateralized behaviour as indicator of affective state in dairy cows. PLoS One 12(9):e0184933. https://doi.org/10.1371/journal.pone.0184933

Karagiannis CI, Burman OHP, Mills DS (2015) Dogs with separation-related problems show a “less pessimistic” cognitive bias during treatment with fluoxetine (Reconcile^{TM}) and a behaviour modification plan. BMC Vet Res 11:80. https://doi.org/10.1186/s12917-015-0373-1

Katayama M, Kubo T, Mogi K et al (2016) Heart rate variability predicts the emotional state in dogs. Behav. Process. 128:108-112. https://doi.org/10.1016/j.beproc.2016.04.015

Keeling LJ, Winckler C, Hintze S, Forkman B (2021) Towards a Positive Welfare Protocol for Cattle: A Critical Review of Indicators and Suggestion of How We Might Proceed. Front Anim Sci 2:1–19. https://doi.org/10.3389/fanim.2021.753080

Kemp AH, Brunoni AR, Santos IS, et al (2014) Effects of Depression, Anxiety, Comorbidity, and Antidepressants on Resting-State Heart Rate and Its Variability: An ELSA-Brasil Cohort Baseline Study. Am J Psychiatry 1328–1334. https://doi.org/10.3779/j.issn.1009-3419.2015.03.04

Kemp AH, Guastella AJ (2011) The role of oxytocin in human affect: a novel hypothesis. Curr. Dir. Psychol. 20(4):222-231. https://doi.org/10.1177/0963721411417547

Kemp AH, Quintana DS, Felmingham KL, et al (2012) Depression, comorbid anxiety disorders, and heart rate variability in physically healthy, unmedicated patients: Implications for cardiovascular risk. PLoS One 7:1–8. https://doi.org/10.1371/journal.pone.0030777

Kindahl H, Kornmatitsuk B, Königsson K, Gustafsson H (2002) Endocrine changes in late bovine pregnancy with special emphasis on fetal well-being. Domest Anim Endocrinol 23:321–328. https://doi.org/10.1016/S0739-7240(02)00167-4

Kiyokawa Y, Hennessy MB (2018) Comparative studies of social buffering: a consideration of approaches, terminology, and pitfalls. Neurosci. Biobehav. Rev. 86:131-141. https://doi.org/ 10.1016/j.neubiorev.2017.12.005

Klaassen T, Riedel W, Deutz N, Van Praag H (2002) Mood congruent memory bias induced by tryptophan depletion. Psychol. Med. 32(1):167-172. https://doi.org/ 10.1017/s003329170100438x

Kluever BM, Howery LD, Breck SW, Bergman DL (2009) Predator and heterospecific stimuli alter behaviour in cattle. 81:85–91. https://doi.org/10.1016/j.beproc.2009.02.004

Knudson D V., Lindsey C (2014) Type I and Type II Errors in Correlations of Various Sample Sizes. Compr Psychol 3:03.CP.3.1. https://doi.org/10.2466/03.cp.3.1

Kok BE, Coffey KA, Cohn MA, et al (2013) How Positive Emotions Build Physical Health: Perceived Positive Social Connections Account for the Upward Spiral Between Positive Emotions and Vagal Tone. Psychol Sci 24:1123–1132. https://doi.org/10.1177/0956797612470827

Konok V, Nagy K, Miklósi Á (2015) How do humans represent the emotions of dogs? The resemblance between the human representation of the canine and the human affective space. Appl. Anim. Behav. Sci. 162:37-46. https://doi.org/10.1016/j.applanim.2014.11.003

Koolhaas JM (2008) Coping style and immunity in animals: Making sense of individual variation. Brain Behav Immun 22:662–667. https://doi.org/10.1016/j.bbi.2007.11.006

Koolhaas JM, De Boer SF, Buwalda B, Van Reenen K (2007) Individual variation in coping with stress: A multidimensional approach of ultimate and proximate mechanisms. Brain Behav Evol 70:218–226. https://doi.org/10.1159/000105485

Koolhaas JM, De Boer SF, Coppens CM, Buwalda B (2010) Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. Front Neuroendocrinol 31:307–321. https://doi.org/10.1016/j.yfrne.2010.04.001

Koolhaas JM, Korte SM, De Boer SF, et al (1999) Coping styles in animals: current in behavior and stress-physiology. Neurosci Biobehav Rev 23:925–935. https://doi.org/10.1016/S0149-7634(99)00026-3

Koolhaas JM, Van Reenen CG (2016) Animal behavior and well-being symposium: Interaction between coping style/personality, stress, and welfare: Relevance for domestic farm animals. J Anim Sci 94:2284–2296. https://doi.org/10.2527/jas2015-0125

Koren L, Mokady O, Karaskov T, et al (2002) A novel method using hair for determining hormonal levels in wildlife. Anim Behav 63:403–406. https://doi.org/10.1006/anbe.2001.1907

Kovács L, Kézér FL, Jurkovich V, et al (2015) Heart rate variability as an indicator of chronic stress caused by lameness in dairy cows. PLoS One 10:1–13. https://doi.org/10.1371/journal.pone.0134792

Kowalik S, Janczarek I, Kędzierski W et al (2017) The effect of relaxing massage on heart rate and heart rate variability in Purebred Arabian racehorses. Anim. Sci. J. 88(4):669-677. https://doi.org/ 10.1111/asj.12671

Kragel PA, LaBar KS (2013) Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. Emotion 13(4):681. https://doi.org/ 10.1037/a0031820

Kremer L, Bus JD, Webb LE, et al (2021a) Housing and personality effects on judgement and attention biases in dairy cows. Sci Rep 1–18. https://doi.org/10.1038/s41598-021-01843-w

Kremer L, Van Reenen CG, Engel B, et al (2021b) Developing a feasible and sensitive judgement bias task in dairy cows. Anim Cogn. https://doi.org/10.1007/s10071-021-01563-8

Kringelbach ML, Berridge KC (2009) Towards a functional neuroanatomy of pleasure and happiness. Trends Cogn. Sci. 13(11):479-487. https://doi.org/10.1016/j.tics.2009.08.006

Kringelbach ML, Berridge KC (2017) The affective core of emotion: linking pleasure, subjective well-being, and optimal metastability in the brain. Emot. Rev. 9(3):191-199. https://doi.org/ 10.1177/1754073916684558

Kubzansky LD, Kawachi I (2000) Going to the heart of the matter: Do negative emotions cause coronary heart disease? J Psychosom Res 48:323–337. https://doi.org/10.1016/S0022-3999(99)00091-4

Kudryavtseva N (2000) Agonistic behavior: a model, experimental studies, and perspectives. Neuroscience and behavioral physiology 30(3):293-305. https://doi.org/10.1007/BF02471782

Kuppens P, Tuerlinckx F, Russell JA, Barrett LF (2013) The relation between valence and arousal in subjective experience. Psychol Bull 139:917–940. https://doi.org/10.1037/a0030811

Kuraoka K, Nakamura K (2011) The use of nasal skin temperature measurements in studying emotion in macaque monkeys. Physiol. Behav. 102(3-4):347-355. https://doi.org/ 10.1016/j.physbeh.2010.11.029

Lacreuse A, Parr L, Chennareddi L, Herndon JG (2018) Age-related decline in cognitive flexibility in female chimpanzees. Neurobiol Aging 72:83–88. https://doi.org/10.1016/j.neurobiolaging.2018.08.018

Lagisz M, Zidar J, Nakagawa S, et al (2020) Optimism, pessimism and judgement bias in animals: a systematic review and meta-analysis. Neurosci Biobehav Rev. https://doi.org/10.1016/j.neubiorev.2020.07.012

Lambert HS, Carder, G (2017) Looking into the eyes of a cow: Can eye whites be used as a measure of emotional state? Appl. Anim. Behav. Sci. 186:1-6

Lane RD, Reiman EM, Ahern GL et al (1997) Neuroanatomical correlates of happiness, sadness, and disgust. Am. J. Psychiatry 154(7):926. https://doi.org/10.1176/ajp.154.7.926

Lange A, Bauer L, Futschik A, et al (2020) Talking to Cows: Reactions to Different Auditory Stimuli During Gentle Human-Animal Interactions. Front Psychol 11:1–14. https://doi.org/10.3389/fpsyg.2020.579346

Langford DJ, Bailey AL, Chanda ML et al (2010) Coding of facial expressions of pain in the laboratory mouse. Nat Meth 7(6):447. https://doi.org/10.1038/nmeth.1455

Lansade L, Nowak R, Lainé A-L et al (2018) Facial expression and oxytocin as possible markers of positive emotions in horses. Sci. Rep. 8(1):14680. https://doi.org/10.1038/s41598-018-32993-z

Larsen JT, McGraw AP, Cacioppo JT (2001) Can people feel happy and sad at the same time? J. Pers. Soc. Psychol. 81(4):684. https://doi.org/10.1037//0022-3514.81.4.684

Lazarus RS (1999) The cognition-emotion debate: A bit of history. Pages 3-19 in Handbook of cognition and emotion. Vol. 5. M. D. Robinson, E. R. Watkins, and E. Harmon-Jones, ed. Guilford Press, New York, USA. https://doi.org/10.1002/0470013494.ch1

Lecorps B, Ludwig BR, von Keyserlingk MAG, Weary DM (2019) Pain-induced pessimism and anhedonia: Evidence from a novel probability-based judgement bias test. Front Behav Neurosci 13:1–6. https://doi.org/10.3389/fnbeh.2019.00054

Lecorps B, Weary DM, Keyserlingk MA (2018) Pessimism and fearfulness in dairy calves. Sci. Rep. 8(1):1421. https://doi.org/10.1038/s41598-017-17214-3

LeDoux J (1996) Emotional networks and motor control: a fearful view. Progress in Brain Research 107:437-446. https://doi.org/10.1016/s0079-6123(08)61880-4

LeDoux JE (1995) Emotion: Clues from the brain. Annu. Rev. Psychol. 46(1):209-235. https://doi.org/10.1146/annurev.ps.46.020195.001233

LeDoux JE (2012) Evolution of human emotion: a view through fear. Progress in Brain Research 195:431-442. https://doi.org/10.1016/B978-0-444-53860-4.00021-0

LeDoux JE, Hofmann SG (2018) The subjective experience of emotion: a fearful view. Curr. Opin. Behav. Sci. 19:67-72. https://doi.org/10.1016/j.cobeha.2017.09.011

Lee C, Cafe LM, Robinson SL et al (2017) Anxiety influences attention bias but not flight speed and crush score in beef cattle. Appl. Anim. Behav. Sci. 205:210-215. https://doi.org/10.1016/j.applanim.2017.11.003

Lee C, Prayaga K, Reed M, Henshall J (2007) Methods of training cattle to avoid a location using electrical cues. Appl Anim Behav Sci 108:229–238. https://doi.org/10.1016/j.applanim.2006.12.003

Lee C, Verbeek E, Doyle R,Bateson M (2016) Attention bias to threat indicates anxiety differences in sheep. Biol. Lett. 12:20150977. https://doi.org/10.1098/rsbl.2015.0977

Leliveld LM, Düpjan S, Tuchscherer A, Puppe B (2017) Vocal correlates of emotional reactivity within and across contexts in domestic pigs (Sus scrofa). Physiol. Behav. 181:117-126. https://doi.org/ 10.1016/j.physbeh.2017.09.010

Leliveld LM, Langbein J, Puppe B (2013) The emergence of emotional lateralization: evidence in non-human vertebrates and implications for farm animals. Appl. Anim. Behav. Sci. 145(1):1-14. https://doi.org/10.1016/j.applanim.2013.02.002

Leliveld LMC, Provolo G (2020) A Review of Welfare Indicators of Indoor-Housed Dairy Cow as a Basis for Integrated Automatic Welfare Assessment Systems. Animals 10:1430. https://doi.org/10.3390/ani10081430

Le Neindre P, Bernard E, Boissy A et al (2017) Animal consciousness. EFSA Supporting Publications 14(4)

Lerner JS, Keltner D (2000) Beyond valence: Toward a model of emotion-specific influences on judgement and choice. Cogn. Emot. 14(4):473-493. https://hal.archives-ouvertes.fr/hal-01594939

Ling J, Xu D, Robbins LB, Meyer JS (2020) Does hair cortisol really reflect perceived stress? Findings from low-income mother-preschooler dyads. Psychoneuroendocrinology 111:104478. https://doi.org/10.1016/j.psyneuen.2019.104478

Lipponen JA, Tarvainen MP (2019) A robust algorithm for heart rate variability time series artefact correction using novel beat classification. J Med Eng Technol 43:173–181. https://doi.org/10.1080/03091902.2019.1640306

Lockwood SA, Kattesh HG, Rhinehart JD, et al (2017) Relationships among temperament, acute and chronic cortisol and testosterone concentrations, and breeding soundness during performance testing of Angus bulls. Theriogenology 89:140–145. https://doi.org/10.1016/j.theriogenology.2016.10.019

Luo L, Reimert I, De Haas EN et al (2019) Effects of early and later life environmental enrichment and personality on attention bias in pigs (Sus scrofa domesticus). Anim Cogn 22:959–972. https://doi.org/10.1007/s10071-019-01287-w

Luo L, Reimert I, Smeets S et al (2018) Effects of (a switch in) enriched vs barren housing on the response to reward loss in pigs in a negative contrast test. Pages 233-233 in Proc. ISAE 2018: 52nd Congress of the International Society for Applied Ethology. Wageningen Academic Publishers. https://www.applied-ethology.org/res/ISAE%202018%20abstract%20book.pdf#page=235

Lv J, Li J, Wang C et al (2018) Positive or negative emotion induced by feeding success or failure can affect behaviors, heart rate and immunity of suckling calves. Physiol. Behav. 196:185-189. https://doi.org/ 10.1016/j.physbeh.2018.09.006

MacLeod C, Mathews A, Tata P (1986) Attentional Bias in Emotional Disorders. J Abnorm Psychol 95:15–20. https://doi.org/10.1037/0021-843X.95.1.15

Maestripieri D, Schino G, Aureli F, Troisi A (1992) A modest proposal: displacement activities as an indicator of emotions in primates. Anim. Behav. 44(5):967-979. https://doi.org/ https://doi.org/10.1016/S0003-3472(05)80592-5

Makowska IJ, Weary DM (2013) Assessing the emotions of laboratory rats. Appl. Anim. Behav. Sci. 148(1):1-12. https://doi.org/10.1016/j.applanim.2013.07.017

Manrique HM, Call J (2015) Age-dependent cognitive inflexibility in great apes. Anim Behav 102:1–6. https://doi.org/10.1016/j.anbehav.2015.01.002

Marin RS, FIrinciogullari S, Biedrzycki RC (1994) Group Differences in the Relationship Between Apathy and Depression. J Nerv Ment Diseaese 182:235–239. https://doi.org/10.1097/00005053-199404000-00008

Mathews A, MacLeod C (1994) Cognitive approaches to emotion and emotional disorders. Annu. Rev. Psychol. 45(1):25-50. https://doi.org/ 10.1146/annurev.ps.45.020194.000325

Mathews A, MacLeod C (2002) Induced processing biases have causal effects on anxiety. Cogn. Emot. 16(3):331-354. https://doi.org/10.1080/02699930143000518

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459(7248):837. https://www.nature.com/articles/nature08028

Mazza V, Eccard JA, Zaccaroni M, et al (2018) The fast and the flexible: cognitive style drives individual variation in cognition in a small mammal. Anim Behav 137:119–132. https://doi.org/10.1016/j.anbehav.2018.01.011

Mazza V, Jacob J, Dammhahn M, et al (2019) Individual variation in cognitive style reflects foraging and anti-predator strategies in a small mammal. Sci Rep 9:1–9. https://doi.org/10.1038/s41598-019-46582-1

McConnachie E, Smid AMC, Thompson AJ, et al (2018) Cows are highly motivated to access a grooming substrate. Biol Lett 14:1–4. https://doi.org/10.1098/rsbl.2018.0303

McCulloch CE, Neuhaus JM. (2014). Generalized linear mixed models. Wiley StatsRef: Statistics Reference Online

McCullough ME, Churchland PS, Mendez AJ (2013) Problems with measuring peripheral oxytocin: can the data on oxytocin and human behavior be trusted? Neurosci. Biobehav. Rev. 37(8):1485-1492. https://doi.org/10.1016/j.neubiorev.2013.04.018

McDonald CL, Beilharz RG, McCutchan JC (1981) Training cattle to control by electric fences. Appl Anim Ethol 7:113–121. https://doi.org/10.1016/0304-3762(81)90089-4

McEwen BS (2003) Mood disorders and allostatic load. Biol Psychiatry 54:200–207. https://doi.org/10.1016/S0006-3223(03)00177-X

McEwen BS (2004) Protection and damage from acute and chronic stress: Allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. Ann N Y Acad Sci 1032:1–7. https://doi.org/10.1196/annals.1314.001

McEwen BS (2006) Sleep deprivation as a neurobiologic and physiologic stressor: allostasis and allostatic load. Metabolism 55:23–26. https://doi.org/10.1016/j.metabol.2006.07.008

McEwen BS (2012) Protective and Damaging Effects of Stress Mediators. N Engl J Med 338:171–179

McEwen BS, Gianaros PJ (2010) Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. Ann. N. Y. Acad. Sci. 1186(1):190-222. https://doi.org/10.31887/DCNS.2006.8.4/bmcewen

McEwen BS, Nasveld P, Palmer M, Anderson R (2012) Allostatic Load: A Review of the Literature | Department of Veterans' Affairs. Aust Gov Dep Veterans' Aff. https://www.dva.gov.au/sites/default/files/allostatic.pdf

McEwen BS, Stellar E (1993) Stress and the Individual. Mechanisms Leading to Disease. Arch Intern Med 153:2093–2101. https://doi.org/10.1001/archinte.1993.00410180039004

McLennan KM, Rebelo CJ, Corke MJ et al (2016) Development of a facial expression scale using footrot and mastitis as models of pain in sheep. Appl. Anim. Behav. Sci. 176:19-26. https://doi.org/10.1016/j.applanim.2016.01.007

Meagher RK, Mason GJ (2012) Environmental Enrichment Reduces Signs of Boredom in Caged Mink. PLoS One 7:. https://doi.org/10.1371/journal.pone.0049180

Melin M, Wiktorsson H, Norell L (2005) Analysis of feeding and drinking patterns of dairy cows in two cow traffic situations in automatic milking systems. J Dairy Sci 88:71–85. https://doi.org/10.3168/jds.S0022-0302(05)72664-3

Mendl M, Brooks J, Basse C et al (2010a) Dogs showing separation-related behaviour exhibit a 'pessimistic'cognitive bias. Curr. Biol. 20(19):R839-R840. https://doi.org/10.1016/j.cub.2010.08.030

Mendl M, Burman OHP, Parker RMA, Paul ES (2009) Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. Appl. Anim. Behav. Sci. 118(3-4):161-181. https://doi.org/10.1016/j.applanim.2009.02.023

Mendl M, Burman OHP, Paul ES (2010b) An integrative and functional framework for the study of animal emotion and mood. Proc. R. Soc. B-Biol. Sci. 277(1696):2895-2904.

Mendonça-De-Souza ACF, Souza GGL, Vieira A, et al (2007) Negative affect as a predisposing factor for cortisol release after an acute stress - The impact of unpleasant priming. Stress 10:362–367. https://doi.org/10.1080/10253890701379999

Meyer JS, Novak MA (2012) Minireview: Hair cortisol: A novel biomarker of hypothalamic-pituitary-adrenocortical activity. Endocrinology 153:4120–4127. https://doi.org/10.1210/en.2012-1226

Meyer PW, Müller LE, Zastrow A, et al (2016) Heart rate variability in patients with post-traumatic stress disorder or borderline personality disorder: relationship to early life maltreatment. J Neural Transm 123:1107–1118. https://doi.org/10.1007/s00702-016-1584-8

Mineka S, Nugent K (1995) Mood-congruent memory biases in anxiety and depression. in Memory Distortions: How Minds, Brains, and Societies Reconstruct the Past. D. L. Schacter and J. T. Coyle, ed. Harvard University Press, London, UK. https://doi.org/ 10.1037//0021-843x.101.3.581

Mineka S, Watson D, Clark LA (1998) Comorbidity of anxiety and unipolar mood disorders. Annu. Rev. Psychol. 49(1):377-412. https://doi.org/ 10.1146/annurev.psych.49.1.377

Mitsui S, Yamamoto M, Nagasawa M et al (2011) Urinary oxytocin as a noninvasive biomarker of positive emotion in dogs. Horm. Behav. 60(3):239-243. https://doi.org/ 10.1016/j.yhbeh.2011.05.012

Mogg K, Bradley BP, Williams R (1995) Attentional bias in anxiety and depression: The role of awareness. Br J Clin Psychol 34:17–36. https://doi.org/10.1111/j.2044-8260.1995.tb01434.x

Monk JE, Belson S, Colditz IG, Lee C (2018a) Attention bias test differentiates anxiety and depression in sheep. Front Behav Neurosci 12:1–14. https://doi.org/10.3389/fnbeh.2018.00246

Monk JE, Belson S, Lee C (2019a) Pharmacologically-induced stress has minimal impact on judgement and attention biases in sheep. Sci Rep 9:1–14. https://doi.org/10.1038/s41598-019-47691-7

Monk JE, Doyle RE, Colditz IG et al (2018b) Towards a more practical attention bias test to assess affective state in sheep. PLoS One 13(1):e0190404. https://doi.org/10.1371/journal.pone.0190404

Monk JE, Lee C, Belson S, et al (2019b) The influence of pharmacologically-induced affective states on attention bias in sheep. PeerJ 7:e7033. https://doi.org/10.7717/peerj.7033

Monk JE, Lee C, Dickson E, Campbell DLM (2020) Attention bias test measures negative but not positive affect in sheep: A replication study. Animals 10:1–18. https://doi.org/10.3390/ani10081314

Moors A (2009) Theories of emotion causation: A review. Cogn. Emot. 23(4):625-662. https://doi.org/10.1080/02699930802645739

Moors A, Ellsworth PC, Scherer KR, Frijda NH (2013) Appraisal theories of emotion: State of the art and future development. Emot. Rev. 5(2):119-124. https://doi.org/10.1177/1754073912468165

Mormède P, Andanson S, Auperin B et al (2007) Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. Physiol. Behav. 92(3):317-339. https://doi.org/10.1016/j.physbeh.2006.12.003

Morris PH, Doe C, Godsell E (2008) Secondary emotions in non-primate species? Behavioural reports and subjective claims by animal owners. Cogn. Emot. 22(1):3-20. https://doi.org/10.1080/02699930701273716

Mück-Šeler D, Pivac N (2011) Serotonin. Period. Biol. 113(1):29-41

Muroy SE, Long KLP, Kaufer D, Kirby ED (2016) Moderate Stress-Induced Social Bonding and Oxytocin Signaling are Disrupted by Predator Odor in Male Rats. Neuropsychopharmacology 41:2160–2170. https://doi.org/10.1038/npp.2016.16

Murphy E, Nordquist RE, van der Staay FJ (2014) A review of behavioural methods to study emotion and mood in pigs, Sus scrofa. Appl. Anim. Behav. Sci. 159:9-28. https://doi.org/10.1016/j.applanim.2014.08.002

Nalon E, Stevenson P (2019) Protection of dairy cattle in the EU: State of play and directions for policymaking from a legal and animal advocacy perspective. Animals 9:1–13. https://doi.org/10.3390/ani9121066

Neave HW, Costa JHC, Weary DM, von Keyserlingk MAG (2020) Long-term consistency of personality traits of cattle. R Soc Open Sci 7:. https://doi.org/10.1098/rsos.191849

Neave HW, Daros RR, Costa JH et al (2013) Pain and pessimism: dairy calves exhibit negative judgement bias following hot-iron disbudding. PLoS One 8(12):e80556. https://doi.org/10.1371/journal.pone.0080556

Neave HW, Weary DM, von Keyserlingk MAG (2018) Review: Individual variability in feeding behaviour of domesticated ruminants. Animal 12:S419–S430. https://doi.org/10.1017/S1751731118001325

Neethirajan S, Reimert I, Kemp B (2021) Measuring farm animal emotions—sensor-based approaches. Sensors (Switzerland) 21:1–23. https://doi.org/10.3390/s21020553

Nettle D, Bateson M (2012) The evolutionary origins of mood and its disorders. Curr. Biol. 22(17):R712-R721. https://doi.org/10.1016/j.cub.2012.06.020

Neumann ID, Landgraf R (2012) Balance of brain oxytocin and vasopressin: Implications for anxiety, depression, and social behaviors. Trends Neurosci. 35(11):649-659. https://doi.org/10.1016/j.tins.2012.08.004

Neumann R, Seibt B, Strack F (2001) The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. Cogn Emot 15:725–747. https://doi.org/10.1080/02699930143000266

Neville V, Nakagawa S, Zidar J, et al (2020) Pharmacological manipulations of judgement bias: A systematic review and meta-analysis. Neurosci Biobehav Rev 108:269–286. https://doi.org/10.1016/j.neubiorev.2019.11.008

Nicol CJ (1995) The social transmission of information and behavior. Appl. Anim. Behav. Sci. 44:79-98. https://doi.org/10.1016/0168-1591(95)00607-T

Nogues E, Lecorps B, Weary DM, von Keyserlingk MAG (2020) Individual variability in response to social stress in dairy heifers. Animals 10:1–10. https://doi.org/10.3390/ani10081440

Novak J, Stojanovski K, Melotti L, et al (2016) Effects of stereotypic behaviour and chronic mild stress on judgement bias in laboratory mice. Appl Anim Behav Sci 174:162–172. https://doi.org/10.1016/j.applanim.2015.10.004

Nummenmaa L,Tuominen L (2017) Opioid system and human emotions. Br. J. Pharmacol. 175:2737-2749. https://doi.org/10.1111/bph.13812

Ogundeji AA, Lakew H, Tesfuhuney W, Lombard W (2021) Influence of heat stress on milk production and its financial implications in semi-arid areas of South Africa. Heliyon 7:e06202 https://doi.org/10.1016/j.heliyon.2021.e06202

Ohl F, Arndt SS, Van Der Staay FJ (2008) Pathological anxiety in animals. The veterinary journal 175(1):18-26. https://doi.org/10.1016/j.tvjl.2006.12.013

Ohl F, Van Der Staay FJ (2012) Animal welfare: At the interface between science and society. Vet J 192:13–19. https://doi.org/10.1016/j.tvjl.2011.05.019

Öhman A, Mineka S (2001) Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. Psychol. Rev. 108(3):483. https://doi.org/ 10.1037/0033-295x.108.3.483

Öhman A, Soares JJ (1994) "Unconscious anxiety": phobic responses to masked stimuli. J. Abnorm. Psychol. 103(2):231. https://doi.org/10.1037//0021-843x.103.2.231

Oka T, Oka K, Hori, T (2001) Mechanisms and mediators of psychological stress-induced rise in core temperature. Psychosom. Med. 63(3):476-486. https://doi.org/ 10.1097/00006842-200105000-00018

Oliveira RF, Faustino AI (2017) Social information use in threat perception: Social buffering, contagion and facilitation of alarm responses. Communicative & integrative biology 10(3):44329. https://doi.org/ 10.1080/19420889.2017.1325049

Packer RMA, Davies AM, Volk HA, et al (2019) What can we learn from the hair of the dog? Complex effects of endogenous and exogenous stressors on canine hair cortisol. PLoS One 14:1–15. https://doi.org/10.1371/journal.pone.0216000

Panksepp J (2003) At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain. Brain Cognition 52(1):4-14. https://doi.org/ 10.1016/s0278-2626(03)00003-4

Panksepp J (2004) Affective neuroscience: The foundations of human and animal emotions. Oxford university press, Oxford, UK

Panksepp J (2005) Affective consciousness: Core emotional feelings in animals and humans. Conscious. Cogn. 14(1):30-80. https://doi.org/10.1016/j.concog.2004.10.004

Panksepp J (2010) Affective neuroscience of the emotional BrainMind: evolutionary perspectives and implications for understanding depression. Dialogues Clin. Neurosci. 12(4):533

Panksepp J (2011) The basic emotional circuits of mammalian brains: do animals have affective lives? Neurosci. Biobehav. Rev. 35(9):1791-1804. https://doi.org/10.31887/DCNS.2010.12.4/jpanksepp

Panksepp J (2016) The cross mammalian neurophenomenology of primal emotional affects: From animal feelings to human therapeutics. J. Comp. Neurol. 524(8):1624-1635. https://doi.org/ 10.1002/cne.23969

Panksepp J, Burgdorf J (2000) 50-kHz chirping (laughter?) in response to conditioned and unconditioned tickle-induced reward in rats: effects of social housing and genetic variables. Behav. Brain Res. 115(1):25-38. https://doi.org/ 10.1016/s0166-4328(00)00238-2

Panksepp J, Lane RD, Solms M, Smith R (2017) Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience. Neurosci. Biobehav. Rev. 76:187-215. https://doi.org/10.1016/j.neubiorev.2016.09.010

Papanicolaou DA, Tsigos C, Oldfield EH, Chrousos GP (1996) Acute glucocorticoid deficiency is associated with plasma elevations of interleukin-6: does the latter participate in the symptomatology of the steroid withdrawal syndrome and adrenal insufficiency? J Clin Endocrinol Metab 81:2303–2306. https://doi.org/10.1210/jcem.81.6.8964868

Papini MR, Penagos-Corzo JC, Pérez-Acosta AM (2019) Avian emotions: Comparative perspectives on fear and frustration. Front Psychol 9:. https://doi.org/10.3389/fpsyg.2018.02707

Parker RMA, Paul ES, Burman OHP, et al (2014) Housing conditions affect rat responses to two types of ambiguity in a reward – reward discrimination cognitive bias task. Behav Brain Res 274:73–83. https://doi.org/10.1016/j.bbr.2014.07.048

Parr L, Waller B (2006) The evolution of human emotion. in Evolution of Nervous Systems: A Comprehensive Reference. J. Kaas, ed. Academic Press Inc., London, UK. https://doi.org/10.1016/B978-0-444-53860-4.00021-0

Parr LA, Waller BM, Burrows AM et al (2010) Brief communication: MaqFACS: a muscle-based facial movement coding system for the rhesus macaque. American journal of physical anthropology 143(4):625-630. https://doi.org/ 10.1002/ajpa.21401

Patron E, Benvenuti SM, Favretto G et al (2014) Depression and reduced heart rate variability after cardiac surgery: the mediating role of emotion regulation. Auton. Neurosci. 180:53-58.

Paul ES, Edgar JL, Caplen G, Nicol CJ (2018) Examining affective structure in chickens: valence, intensity, persistence and generalization measured using a Conditioned Place Preference Test. Appl. Anim. Behav. Sci. 207:39-48. https://doi.org/10.1016/j.autneu.2013.11.004

Paul ES, Harding EJ, Mendl M (2005) Measuring emotional processes in animals: the utility of a cognitive approach. Neurosci. Biobehav. Rev. 29(3):469-491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Paul ES, Mendl MT (2018) Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. Appl. Anim. Behav. Sci. 205:202-209. https://doi.org/10.1016/j.applanim.2018.01.008

Pejman N, Kallas Z, Dalmau A, Velarde A (2019) Should Animal Welfare Regulations Be More Restrictive? A Case Study in Eight European Union Countries. Animals 9. https://doi.org/10.3390/ani9040195

Peng W, Lou W, Huang X, et al (2021) Suffer together, bond together: Brain-to-brain synchronization and mutual affective empathy when sharing painful experiences. Neuroimage 238:118249. https://doi.org/10.1016/j.neuroimage.2021.118249

Perry CJ, Baciadonna L (2017) Studying emotion in invertebrates: what has been done, what can be measured and what they can provide. J. Exp. Biol. 220(21):3856-3868. https://doi.org/10.1242/jeb.151308

Perry CJ, Baciadonna L, Chittka L (2016) Unexpected rewards induce dopamine-dependent positive emotion–like state changes in bumblebees. Science 353(6307):1529-1531. https://doi.org/10.1126/science.aaf4454

Pfister JA, Müller-Schwarze D, Balph DF (1990) Effects of predator fecal odors on feed selection by sheep and cattle. J Chem Ecol 16:573–583. https://doi.org/10.1007/BF01021787

Phillips CJC, Rind MI (2001) The Effects on Production and Behavior of Mixing Uniparous and Multiparous Cows. J Dairy Sci 84:2424–2429. https://doi.org/10.3168/jds.S0022-0302(01)74692-9

Pizzagalli DA, Iosifescu D, Hallett LA, et al (2008) Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. J Psychiatr Res 43:76–87. https://doi.org/10.1016/j.jpsychires.2008.03.001

Pomerantz O, Terkel J (2009) Effects of positive reinforcement training techniques on the psychological welfare of zoo-housed chimpanzees (Pan troglodytes). Am J Primatol 71:687–695. https://doi.org/10.1002/ajp.20703

Poppe M, Mulder HA, Kamphuis C, Veerkamp RF (2021) Between-herd variation in resilience and relations to herd performance. J Dairy Sci 104:616–627. https://doi.org/10.3168/jds.2020-18525

Poppe M, Veerkamp RF, van Pelt ML, Mulder HA (2020) Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. J Dairy Sci 103:1667–1684. https://doi.org/10.3168/jds.2019-17290

Portfors CV (2007) Types and functions of ultrasonic vocalizations in laboratory rats and mice. J. Am. Assoc. Lab. Anim. Sci. 46(1):28-34.

Proctor H, Carder G (2016) Can changes in nasal temperature be used as an indicator of emotional state in cows? Appl. Anim. Behav. Sci. 184:1-6. https://doi.org/10.1016/j.applanim.2016.07.013

Proctor HS, Carder G (2015) Measuring positive emotions in cows: Do visible eye whites tell us anything? Physiol. Behav. 147:1-6. https://doi.org/10.1016/j.physbeh.2015.04.011

Quigley K, Lindquist KA, Barrett LF (2014) Inducing and measuring emotion and affect: Tips, tricks, and secrets. in Handbook of research methods in social and personality psychology. H. T. Reis and C. M. Judd, ed. Cambridge University Press, Cambridge, UK.

Quinones MP, Kaddurah-Daouk R (2009) Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. Neurobiol. Dis. 35(2):165-176. https://doi.org/10.1016/j.nbd.2009.02.019

Rajala-Schultz PJ, Gröhn YT, McCulloch CE, Guard CL (1999) Effects of clinical mastitis on milk yield in dairy cows. J Dairy Sci 82:1213–1220. https://doi.org/10.3168/jds.S0022-0302(99)75344-0

Raoult CMC, Gygax L (2019) Mood induction alters attention toward negative-positive stimulus pairs in sheep. Sci Rep 9:1–11. https://doi.org/10.1038/s41598-019-44330-z

Raoult CMC, Moser J, Gygax L (2017) Mood as cumulative expectation mismatch: A test of theory based on data from non-verbal cognitive bias tests. Front Psychol 8:1–10. https://doi.org/10.3389/fpsyg.2017.02197

Rault J (2016) Effects of positive and negative human contacts and intranasal oxytocin on cerebrospinal fluid oxytocin. Psychoneuroendocrinology 69:60–66. https://doi.org/10.1016/j.psyneuen.2016.03.015

Rault J (2019) Be kind to others : Prosocial behaviours and their implications for animal welfare. Appl Anim Behav Sci 210:113–123. https://doi.org/10.1016/j.applanim.2018.10.015

Rault J-L (2012) Friends with benefits: social support and its relevance for farm animal welfare. Appl. Anim. Behav. Sci. 136(1):1-14. https://doi.org/10.1016/j.applanim.2011.10.002

Rault J-L, Van Den Munkhof M, Buisman-Pijlman FT (2017) Oxytocin as an indicator of psychological and social well-being in domesticated animals: a critical review. Front. Psychol. 8:1521. https://doi.org/10.3389/fpsyg.2017.01521

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Reefmann N, Kaszàs FB, Wechsler B, Gygax L (2009a) Physiological expression of emotional reactions in sheep. Physiol. Behav. 98(1-2):235-241. https://doi.org/ 10.1016/j.physbeh.2009.05.017

Reefmann N, Muehlemann T, Wechsler B, Gygax L (2012) Housing induced mood modulates reactions to emotional stimuli in sheep. Appl. Anim. Behav. Sci. 136(2-4):146-155. https://doi.org/10.1016/j.applanim.2011.12.007

Reefmann N, Wechsler B, Gygax L (2009b) Behavioural and physiological assessment of positive and negative emotion in sheep. Anim. Behav. 78(3):651-659. https://doi.org/10.1016/j.anbehav.2009.06.015

Rehkämper G, Perrey A, Werner CW, et al (2000) Visual perception and stimulus orientation in cattle. Vision Res 40:2489–2497. https://doi.org/10.1016/S0042-6989(00)00113-9

Reijgwart ML, Schoemaker NJ, Pascuzzo R et al (2017) The composition and initial evaluation of a grimace scale in ferrets after surgical implantation of a telemetry probe. PLoS One 12(11):e0187986. https://doi.org/10.1371/journal.pone.0187986

Reimert I, Bolhuis JE, Kemp B, Rodenburg TB (2013) Indicators of positive and negative emotions and emotional contagion in pigs. Physiol. Behav. 109:42-50. https://doi.org/ 10.1016/j.physbeh.2012.11.002

Reimert I, Bolhuis JE, Kemp B, Rodenburg TB (2014) Social support in pigs with different coping styles. Physiol. Behav. 129:221-229. https://doi.org/10.1016/j.physbeh.2014.02.059

Reimert I, Bolhuis JE, Kemp B, Rodenburg TB (2015) Emotions on the loose: emotional contagion and the role of oxytocin in pigs. Anim. Cogn. 18(2):517-532. https://doi.org/10.1007/s10071-014-0820-6

Reimert I, Fong S, Rodenburg TB, Bolhuis JE (2017) Emotional states and emotional contagion in pigs after exposure to a positive and negative treatment. Appl. Anim. Behav. Sci. 193:37-42

Richman LS, Kubzansky L, Maselko J, et al (2005) Positive emotion and health: Going beyond the negative. Heal Psychol 24:422–429. https://doi.org/10.1037/0278-6133.24.4.422

Ridout KK, Ridout SJ, Price LH et al (2016) Depression and telomere length: a meta-analysis. J. Affect. Disord. 191:237-247. https://doi.org/10.1016/j.applanim.2017.03.009

Rief W, Shaw R, Fichter MM (1998) Elevated levels of psychophysiological arousal and cortisol in patients with somatization syndrome. Psychosom Med 60:198–203. https://doi.org/10.1097/00006842-199803000-00016

Rius MM, Cozzi A, Bienboire-Frosini C et al (2018) Providing straw to allow exploratory behaviour in a pig experimental system does not modify putative indicators of positive welfare: peripheral oxytocin and serotonin. Animal:1-9. https://doi.org/ 10.1017/S175173111800006X

Robbins J, Franks B, von Keyserlingk MAG (2018) 'More than a feeling': An empirical investigation of hedonistic accounts of animal welfare. PLoS One 13:1–11. https://doi.org/10.1371/journal.pone.0193864

Robinson SA, Erickson RL, Browne CA, Lucki I (2017) A role for the mu opioid receptor in the antidepressant effects of buprenorphine. Behav. Brain Res. 319:96-103. https://doi.org/ 10.1016/j.bbr.2016.10.050

Rochais C, Henry S, Fureix C, Hausberger M (2016) Investigating attentional processes in depressive-like domestic horses (Equus caballus). Behav. Process. 124:93-96. https://doi.org/ 10.1016/j.beproc.2015.12.010

Roelofs S, Boleij H, Nordquist RE, van der Staay FJ (2016) Making decisions under ambiguity: judgement bias tasks for assessing emotional state in animals. Front. Behav. Neurosci. 10:1–16. https://doi.org/10.3389/fnbeh.2016.00119

Rolls ET (2000) Precis of the brain and emotion. Behav. Brain Sci. 23(2):177-234. https://doi.org/ 10.1017/s0140525x00002429

Rørvang MV, Nawroth C (2021) Understanding the behaviour and improving the welfare of dairy cattle. In: Endres M (ed). Burleigh Dodds Science Publishing

Rosas JM, Callejas-Aguilera JE, Escarabajal MD et al (2007) Successive negative contrast effect in instrumental runway behaviour: a study with Roman high-(RHA) and Roman low-(RLA) avoidance rats. Behav. Brain Res. 185(1):1-8. https://doi.org/10.1016/j.bbr.2007.07.027

Ross M, Garland A, Harlander-matauschek A, Kitchenham L (2019) Welfare-improving enrichments greatly reduce hens ' startle responses , despite little change in judgment bias. Sci Rep 1–14. https://doi.org/10.1038/s41598-019-48351-6

Ross RA, Foster SL, Ionescu DF (2017) The Role of Chronic Stress in Anxious Depression. Chronic Stress 1:. https://doi.org/10.1177/2470547016689472

Roy AK, Ph D, Vasa RA, et al (2009) Attention Bias Toward Threat in Pediatric Anxiety Disorders. J Am Acad Child Adolesc Psychiatry 47:1189–1196. https://doi.org/10.1097/CHI.0b013e3181825ace.

Ruhé HG, Mason NS, Schene AH (2007) Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. Mol. Psychiatry 12(4):331. https://doi.org/10.1038/sj.mp.4001949

Ruis MA, Te Brake JH, Engel B et al (1997) The circadian rhythm of salivary cortisol in growing pigs: effects of age, gender, and stress. Physiol. Behav. 62(3):623-630. https://doi.org/ 10.1016/s0031-9384(97)00177-7

Rushen J, Munksgaard L, Marnet PG, DePassillé AM (2001) Human contact and the effects of acute stress on cows at milking. Appl Anim Behav Sci 73:1–14. https://doi.org/10.1016/S0168-1591(01)00105-8

Russell JA (1980) A circumplex model of affect. J. Pers. Soc. Psychol. 39(6):1161. https://doi.org/10.1037/h0077714

Russell JA (2003) Core affect and the psychological construction of emotion. Psychol. Rev. 110(1):145. https://doi.org/10.1037/0033-295x.110.1.145

Russell JA, Barrett LF (1999) Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. J. Pers. Soc. Psychol. 76(5):805. https://doi.org/10.1037//0022-3514.76.5.805

Rygula R, Pluta H, Popik P (2012) Laughing Rats Are Optimistic. PLoS One 7(12):e51959. https://doi.org/10.1371/journal.pone.0051959

Saito Y, Yuki S, Seki Y et al (2016) Cognitive bias in rats evoked by ultrasonic vocalizations suggests emotional contagion. Behav. Process. 132:5-11. https://doi.org/10.1016/j.beproc.2016.08.005

Sandem A-I, Braastad BO, Bakken M (2006) Behaviour and percentage eye-white in cows waiting to be fed concentrate—A brief report. Appl. Anim. Behav. Sci. 97(2-4):145-151. https://doi.org/10.1016/j.applanim.2005.08.003

Sander D, Grandjean D, Scherer KR (2005) A systems approach to appraisal mechanisms in emotion. Neural Networks 18:317–352. https://doi.org/10.1016/j.neunet.2005.03.001

Sanger ME, Doyle RE, Hinch GN, Lee C (2011) Sheep exhibit a positive judgement bias and stress-induced hyperthermia following shearing. Appl. Anim. Behav. Sci. 131(3-4):94-103. https://doi.org/ 10.1016/j.applanim.2011.02.001

Sankey C, Richard-Yris M-A, Leroy H et al (2010) Positive interactions lead to lasting positive memories in horses, Equus caballus. Anim. Behav. 79(4):869-875. https://doi.org/10.1016/j.anbehav.2009.12.037

Schaafsma FG, Hulsegge G, de Jong MA, et al (2021) The potential of using hair cortisol to measure chronic stress in occupational healthcare; a scoping review. J Occup Health 63:e12189. https://doi.org/10.1002/1348-9585.12189

Schechtman E, Laufer O, Paz R (2010) Negative valence widens generalization of learning. J Neurosci 30:10460–10464. https://doi.org/10.1523/JNEUROSCI.2377-10.2010

Schenk HM, Jeronimus BF, van der Krieke L et al (2018) Associations of positive affect and negative affect with allostatic load: A lifelines cohort study. Psychosom. Med. 80(2):160-166. https://doi.org/10.1097/PSY.0000000000000546

Scherer KR (1999) Appraisal theory. in Handbook of Cognition and Emotion. T. Dalgleish and M. J. Power, ed. Wiley, Hoboken, USA. https://doi.org/10.1002/0470013494.ch30

Scherer KR (2001) Appraisal considered as a process of multilevel sequential checking. in Appraisal Processes in Emotion: Theory, Methods, Research. K. R. Scherer, A. Schorr, and T. Johnstone, ed. Oxford University Press, Oxford, UK.

Scherer KR (2009a) The dynamic architecture of emotion: Evidence for the component process model. Cogn. Emot. 23(7):1307-1351. https://doi.org/10.1080/02699930902928969

Scherer KR (2009b) Emotions are emergent processes: they require a dynamic computational architecture. Philosophical Transactions of the Royal Society of London B: Biological Sciences 364(1535):3459-3474. https://doi.org/10.1098/rstb.2009.0141

Schirmann K, Chapinal N, Weary DM, et al (2011) Short-term effects of regrouping on behavior of prepartum dairy cows. J Dairy Sci 94:2312–2319. https://doi.org/10.3168/jds.2010-3639

Schiweck C, Piette D, Berckmans D et al (2019) Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. A systematic review. Psychol. Med. 49(2):200-211. https://doi.org/10.1017/S0033291718001988

Schrader L, Ladewig J (1999) Temporal differences in the responses of the pituitary adrenocortical axis, the sympathoadrenomedullar axis, heart rate, and behaviour to a daily repeated stressor in domestic pigs. Physiol. Behav. 66(5):775-783. https://doi.org/ 10.1016/s0031-9384(99)00015-3

Schubach KM, Cooke RF, Brandão AP, et al (2017) Impacts of stocking density on development and puberty attainment of replacement beef heifers. Animal 11:2260–2267. https://doi.org/10.1017/S1751731117001070

Schutte NS, Palanisamy SK, McFarlane JR (2016) The relationship between positive psychological characteristics and longer telomeres. Psychology and Health 31(12):1466-1480. https://doi.org/10.1080/08870446.2016.1226308

Schütz KE, Huddart FJ, Sutherland MA, et al (2015) Effects of space allowance on the behavior and physiology of cattle temporarily managed on rubber mats. J Dairy Sci 98:6226–6235. https://doi.org/10.3168/jds.2015-9593

Schwarz N (2000) Emotion, cognition, and decision making. Cogn. Emot. 14(4):433-440. https://doi.org/10.1080/026999300402745

Scollo A, Gottardo F, Contiero B, Edwards SA (2014) Does stocking density modify affective state in pigs as assessed by cognitive bias, behavioural and physiological parameters? Appl Anim Behav Sci 153:26–35. https://doi.org/10.1016/j.applanim.2014.01.006

Sgoifo A, Carnevali L, Pico Alfonso MDLA, Amore M (2015) Autonomic dysfunction and heart rate variability in depression. Stress 18:343–352. https://doi.org/10.3109/10253890.2015.1045868

Shaffer F, Ginsberg JP (2017) An Overview of Heart Rate Variability Metrics and Norms. Front Public Heal 5:1–17. https://doi.org/10.3389/fpubh.2017.00258

Shaffer F, McCraty R, Zerr CL (2014) A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Front Psychol 5:1–19. https://doi.org/10.3389/fpsyg.2014.01040

Sharma A, Umapathy G, Kumar V, Phillips CJC (2019) Hair Cortisol in Sheltered Cows and Its Association. Animals 9:1–27. https://doi.org/10.3390/ani9050248

Shi H, Yang L, Zhao L et al (2017) Differences of heart rate variability between happiness and sadness emotion states: a pilot study. J. Med. Biol. Eng. 37(4):527-539. https://doi.org/ 10.1007/s40846-017-0238-0

Sinclair M, Fryer C, Phillips CJC (2019) The benefits of improving animalwelfare from the perspective of livestock stakeholders across asia. Animals 9:. https://doi.org/10.3390/ani9040123

Sinha R, Lovallo WR, Parsons OA (1992) Cardiovascular differentiation of emotions. Psychosom Med 54:422–435. https://doi.org/10.1097/00006842-199207000-00005

Siniscalchi M, d'Ingeo S, Fornelli S, Quaranta A (2018) Lateralized behavior and cardiac activity of dogs in response to human emotional vocalizations. Sci. Rep. 8(1):77. https://doi.org/ 10.1038/s41598-017-18417-4

Slack MK, Draugalis JR (2001) Establishing the internal and external validity of experimental studies. Am Soc Heal Pharm 58:2173–2181. https://doi.org/10.1093/ajhp/58.22.2173

Smyth N, Bianchin M, Thorn L, et al (2016) Hair cortisol concentrations in relation to ill-being and well-being in healthy young and old females. Int J Psychophysiol 102:12–17. https://doi.org/10.1016/j.ijpsycho.2016.03.001

Soltis J, Leighty KA, Wesolek CM, Savage A (2009) The expression of affect in African elephant (Loxodonta africana) rumble vocalizations. J. Comp. Psychol. 123(2):222. https://doi.org/ 10.1037/a0015223

Southwick SM, Bonanno GA, Masten AS, et al (2014) Resilience definitions, theory, and challenges: interdisciplinary perspectives. Eur J Psychotraumatol 5:25338. https://doi.org/10.1016/B978-1-78548-298-4.50016-6

Southwick SM, Charney DS (2012) The Science of Resilience: Implications for the Prevention and Treatment of Depression. Science (80) 338:79–83. https://doi.org/10.1126/science.1222942

Spain CV, Freund D, Mohan-Gibbons H, et al (2018) Are they buying it? United states consumers' changing attitudes toward more humanely raised meat, eggs, and dairy. Animals 8:1–14. https://doi.org/10.3390/ani8080128

Speer ME, Mauricio R (2017) Reminiscing about positive memories buffers acute stress responses. Nat Hum Behav 1:1–9. https://doi.org/10.1038/s41562-017-0093

Špinka M (2012) Social dimension of emotions and its implication for animal welfare. Appl. Anim. Behav. Sci. 138(3):170-181. https://doi.org/10.1016/j.applanim.2012.02.005

Spoor JR, Kelly JR (2004) The evolutionary significance of affect in groups: Communication and group bonding. Group Process. Intergroup. Relat. 7(4):398-412. https://doi.org/10.1177/1368430204046145

Spruijt BM, van den Bos R, Pijlman FT (2001) A concept of welfare based on reward evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the state of reward systems. Appl. Anim. Behav. Sci. 72(2):145-171. https://doi.org/10.1016/s0168-1591(00)00204-5

Spunt RP, Ellsworth E, Adolphs R (2017) The neural basis of understanding the expression of the emotions in man and animals. Soc. Cogn. Affect. Neurosci. 12(1):95-105. https://doi.org/10.1093/scan/nsw161

Stahl SM (1977) The human platelet: a diagnostic and research tool for the study of biogenic amines in psychiatric and neurologic disorders. Arch. Gen. Psychiatry 34(5):509-516. https://doi.org/10.1001/archpsyc.1977.01770170019001

Stalder T, Steudte-Schmiedgen S, Alexander N, et al (2017) Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. Psychoneuroendocrinology 77:261–274. https://doi.org/10.1016/j.psyneuen.2016.12.017

Standing Committee of Council of Europe (1988) Recommendation concerning cattle. https://www.coe.int/t/e/legal_affairs/legal_co-operation/biological_safety_and_use_of_animals/farming/Rec cattle E.asp#TopOfPage. Accessed 19 Sep 2021

Stephenson E, Haskell MJ (2020) The Use of a "Go/Go" Cognitive Bias Task and Response to a Novel Object to Assess the Effect of Housing Enrichment in Sheep (Ovis aries). J Appl Anim Welf Sci 00:1–13. https://doi.org/10.1080/10888705.2020.1824786

Steptoe A, O'donnell K, Badrick E et al (2007) Neuroendocrine and inflammatory factors associated with positive affect in healthy men and women: the Whitehall II study. Am. J. Epidemiol. 167(1):96-102. https://doi.org/10.1093/aje/kwm252

Steptoe A, Wardle J, Marmot M (2005) Positive affect and health-related neuroendocrine, cardiovascular, and inflammatory processes. Proc. Natl. Acad. Sci. U. S. A. 102(18):6508-6512. https://doi.org/10.1073/pnas.0409174102

Steudte S, Stalder T, Dettenborn L, et al (2011) Decreased hair cortisol concentrations in generalised anxiety disorder. Psychiatry Res 186:310–314. https://doi.org/10.1016/j.psychres.2010.09.002

Stewart JA, Chaiken MF, Wang F, Price CM (2012) Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. Mutat. Res-Fund. Mol. M. 730(1):12-19. https://doi.org/10.1016/j.mrfmmm.2011.08.011

Stone LA, Nielson KA (2001) Intact physiological response to arousal with impaired emotional recognition in alexithymia. Psychother. Psychosom. 70(2):92-102. https://doi.org/10.1159/000056232

Stracke J, Otten W, Tuchscherer A et al (2017a) Serotonin depletion induces pessimistic-like behavior in a cognitive bias paradigm in pigs. Physiol. Behav. 174:18-26. https://doi.org/ 10.1016/j.physbeh.2017.02.036

Stracke J, Otten W, Tuchscherer A et al (2017b) Dietary tryptophan supplementation and affective state in pigs. J. Vet. Behav. 20:82-90. https://doi.org/10.1016/j.jveb.2017.03.009

Stuart SA, Butler P, Munafò MR et al (2013) A translational rodent assay of affective biases in depression and antidepressant therapy. Neuropsychopharmacology 38(9):1625-1635. https://doi.org/10.1038/npp.2013.69

Stuart SA, Butler P, Munafò MR et al (2015) Distinct neuropsychological mechanisms may explain delayed- versus rapid-onset antidepressant efficacy. Neuropsychopharmacology 40(9):2165-2174. https://doi.org/10.1038/npp.2015.59

Takamatsu H, Noda A, Kurumaji A et al (2003) A PET study following treatment with a pharmacological stressor, FG7142, in conscious rhesus monkeys. Brain Res. 980(2):275-280. https://doi.org/10.1016/s0006-8993(03)02987-1

Takatsu-Coleman AL, Patti CL, Zanin KA et al (2013) Short-term social isolation induces depressive-like behaviour and reinstates the retrieval of an aversive task: Mood-congruent memory in male mice? J. Psychiatry Neurosci. 38(4):259-268. https://doi.org/ 10.1503/jpn.120050

Tamioso PR, Rucinque DS, Taconeli CA et al (2017) Behavior and body surface temperature as welfare indicators in selected sheep regularly brushed by a familiar observer. J. Vet. Behav. 19:27-34. https://doi.org/10.1016/j.jveb.2017.01.004

Temple D, Manteca X, Velarde A, Dalmau A (2011) Assessment of animal welfare through behavioural parameters in Iberian pigs in intensive and extensive conditions. Appl. Anim. Behav. Sci. 131(1-2):29-39. https://doi.org/10.1016/j.applanim.2011.01.013

Terlouw CE, Boissy A, Blinet P (1998) Behavioural response of cattle to the odour of blood and urine from the conspecifics and to the odour of faeces from carnivores. Appl Anim Behiour Sci 57:9–21. https://doi.org/10.1016/S0168-1591(97)00122-6

Thayer JF, Hansen AL, Saus-Rose E, Johnsen BH (2009) Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. Ann Behav Med 37:141–153. https://doi.org/10.1007/s12160-009-9101-z

Trimmer PC, Paul ES, Mendl MT et al (2013) On the evolution and optimality of mood states. Behav. Sci. 3(3):501-521. https://doi.org/10.3390/bs3030501

Triscoli C, Croy I, Olausson H, Sailer, U (2017) Touch between romantic partners: being stroked is more pleasant than stroking and decelerates heart rate. Physiol. Behav. 177:169-175. https://doi.org/ 10.1016/j.physbeh.2017.05.006

Trudel-Fitzgerald C, Qureshi F, Appleton AA, Kubzansky LD (2017) A healthy mix of emotions: underlying biological pathways linking emotions to physical health. Curr Opin Behav Sci 15:16–21. https://doi.org/10.1016/j.cobeha.2017.05.003

Tuchscherer M, Puppe B, Tuchscherer A, Kanitz E (1998) Effects of social status after mixing on immune, metabolic, and endocrine responses in pigs. Physiol. Behav. 64(3):353-360. https://doi.org/10.1016/s0031-9384(98)00084-5

Ursinus WW, Bolhuis JE, Zonderland JJ et al (2013) Relations between peripheral and brain serotonin measures and behavioural responses in a novelty test in pigs. Physiol. Behav. 118:88-96. https://doi.org/10.1016/j.physbeh.2013.05.018

Vallortigara G (2000) Comparative neuropsychology of the dual brain: a stroll through animals' left and right perceptual worlds. Brain Lang. 73(2):189-219. https://doi.org/ 10.1006/brln.2000.2303

Van Der Harst J, Spruijt B (2007) Tools to measure and improve animal welfare: reward-related behaviour. Anim. Welf. 16:67-73

Van Der Kooy KG, van Hout HP, van Marwijk HW et al (2006) Differences in heart rate variability between depressed and non-depressed elderly. Int. J. Geriatr. Psychiatry 21(2):147-150. https://doi.org/10.1002/gps.1439

Van Eerdenburg FJCM, Hof T, Doeve B, et al (2021) The relation between hair-cortisol concentration and various welfare assessments of dutch dairy farms. Animals 11:1–17. https://doi.org/10.3390/ani11030821

Van Hooff J (1967) The facial displays of the catarrhine monkeys and apes

Van Reenen CG (2012) Identifying temperament in dairy cows: a longitudinal approach

Van Reenen CG, Engel B, Ruis-Heutinck LFM, et al (2004) Behavioural reactivity of heifer calves in potentially alarming test situations: A multivariate and correlational analysis. Appl Anim Behav Sci 85:11–30. https://doi.org/10.1016/j.applanim.2003.09.007

Van Reenen CG, Hopster H, Van der Werf JTN, et al (2009) The benzodiazepine brotizolam reduces fear in calves exposed to a novel object test. Physiol. Behav. 96:307–314. https://doi.org/ 10.1016/j.physbeh.2008.10.016

Van Reenen CG, O'Connell NE, Van Der Werf JTN, et al (2005) Responses of calves to acute stress: Individual consistency and relations between behavioral and physiological measures. Physiol Behav 85:557–570. https://doi.org/10.1016/j.physbeh.2005.06.015

Van Steenbergen H, de Bruijn ER, van Duijvenvoorde AC, van Harmelen AL (2021) How positive affect buffers stress responses. Curr Opin Behav Sci 39:153–160. https://doi.org/10.1016/j.cobeha.2021.03.014

Veenhoven R (2000) The four qualities of life: ordering concepts and measures of the good life. J. Happiness Stud. 1(1):1-39. https://doi.org/10.1023/A:1010072010360

Veerapa E, Grandgenevre P, El Fayoumi M, et al (2020) Attentional bias towards negative stimuli in healthy individuals and the effects of trait anxiety. Sci Rep 10:1–10. https://doi.org/10.1038/s41598-020-68490-5

Veissier I, Boissy A, Desire L, Greiveldinger L (2009) Animals' emotions: studies in sheep using appraisal theories. Anim. Welf. 18(4):347-354

Verbeek E, Colditz I, Blache D, Lee C (2019) Chronic stress influences attentional and judgement bias and the activity of the HPA axis in sheep. PLoS One 14:1–23. https://doi.org/10.1371/journal.pone.0211363

Verbeek M, Boon A, Drent PJ (1996) Exploration, aggressive behaviour and dominance in pair-wise confrontations of juvenile great tits. Behaviour. https://doi.org/10.1163/156853996X00314

Vick S-J, Waller BM, Parr LA et al (2007) A cross-species comparison of facial morphology and movement in humans and chimpanzees using the facial action coding system (FACS). Journal of nonverbal behavior 31(1):1-20. https://doi.org/10.1007/s10919-006-0017-z

Von Borell E, Ladewig J (1992) Relationship between behaviour and adrenocortical response pattern in domestic pigs. Appl. Anim. Behav. Sci. 34(3):195-206. https://doi.org/10.1016/S0168-1591(05)80115-7

Von Borell E, Langbein J, Després G et al (2007) Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—A review. Physiol. Behav. 92(3):293-316. https://doi.org/10.1016/j.physbeh.2007.01.007

Wallbott HF (1986) Experiencing emotion: A cross-cultural study. Cambridge University Press, Cambridge, UK

Waller BM, Micheletta J (2013) Facial expression in nonhuman animals. Emot. Rev. 5(1):54-59. https://doi.org/10.1177/1754073912451503

Wallis LJ, Virányi Z, Müller CA, et al (2016) Aging effects on discrimination learning, logical reasoning and memory in pet dogs. Age (Omaha) 38:1–18. https://doi.org/10.1007/s11357-015-9866-x

Wang FX, Shao DF, Li SL, et al (2016) Effects of stocking density on behavior, productivity, and comfort indices of lactating dairy cows. J Dairy Sci 99:3709–3717. https://doi.org/10.3168/jds.2015-10098

Wascher CAF (2021) Heart rate as a measure of emotional arousal in evolutionary biology. Philos Trans R Soc B Biol Sci 376:. https://doi.org/10.1098/rstb.2020.0479

Watanabe S (2007) How animal psychology contributes to animal welfare. Appl Anim Behav Sci 106:193–202. https://doi.org/10.1016/j.applanim.2007.01.003

Wathan J, Burrows AM, Waller BM, McComb K (2015) EquiFACS: the equine facial action coding system. PLoS One 10(8):e0131738. https://doi.org/10.1371/journal.pone.0137818

Watson D, Tellegen A (1985) Toward a consensual structure of mood. Psychol. Bull. 98(2):219. https://doi.org/10.1037//0033-2909.98.2.219

Weary DM, Robbins JA (2019) Understanding the multiple conceptions of animal welfare. Anim Welf 28:33–40. https://doi.org/10.7120/09627286.28.1.033

Weary DM, Von Keyserlingk MAG (2017) Public concerns about dairy-cow welfare: How should the industry respond? Anim Prod Sci 57:1201–1209. https://doi.org/10.1071/AN16680

Webb LE, Van Reenen CG, Jensen MB, et al (2015) Does temperament affect learning in calves? Appl Anim Behav Sci 165:33–39. https://doi.org/10.1016/j.applanim.2015.01.013

Webb LE, Veenhoven R, Harfeld JL, Jensen MB (2018) What is animal happiness? Ann. N. Y. Acad. Sci. 1438: 62–76. https://doi.org/10.1111/nyas.13983

Weidman AC, Steckler CM, Tracy JL (2017) The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. Emotion 17(2):267. https://doi.org/10.1037/emo0000226

Welp T, Rushen J, Kramer DL, et al (2004) Vigilance as a measure of fear in dairy cattle. Appl Anim Behav Sci 87:1–13. https://doi.org/10.1016/j.applanim.2003.12.013

Wemelsfelder F (1999) The problem of animal subjectivity and its consequences for the scientific measurement of animal suffering. Attitudes to animals: Views in animal welfare:37-53.

Wemelsfelder F, Hunter TEA, Mendl MT, Lawrence AB (2001) Assessing the 'whole animal': a free choice profiling approach. Anim. Behav. 62:209-220. https://doi.org/10.1006/anbe.2001.1741

Westermann R, Spies K, Stahl G, Hesse FW (1996) Relative effectiveness and validity of mood induction procedures: A meta-analysis. Eur J Soc Psychol 26:557–580. https://doi.org/10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Wilbourn RV, Froy H, McManus M-C et al (2017) Age-dependent associations between telomere length and environmental conditions in roe deer. Biol. Lett. 13(9):4-7. https://doi.org/10.1098/rsbl.2017.0434

Wilcox CS, Schutz MM, Rostagno MR, et al (2013) Repeated mixing and isolation: Measuring chronic, intermittent stress in Holstein calves. J Dairy Sci 96:7223–7233. https://doi.org/10.3168/jds.2013-6944

Wilmink JB. (1987) Adjustement of test-day milk, fat and protein yield for age, season and stage of lactation. Livest Prod Sci 335–348. https://doi.org/10.1016/0301-6226(87)90003-0

Winckler C, Tucker CB, Weary DM (2015) Effects of under- and overstocking freestalls on dairy cattle behaviour. Appl Anim Behav Sci 170:14–19. https://doi.org/10.1016/j.applanim.2015.06.003

Wingfield JC, Moore IT, Goymann W et al (2006) Contexts and ethology of vertebrate aggression: implications for the evolution of hormone-behavior interactions. Biology of aggression 1:179-210. https://doi.org/10.1093/acprof:oso/9780195168761.003.0008

Winkielman P, Berridge KC, Wilbarger JL (2005) Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgements of value. Pers. Soc. Psychol. Bull. 31(1):121-135. https://doi.org/10.1177/0146167204271309

Winkielman P, Berridge KC, Wilbarger JL (2007) Emotion, behaviour, and conscious experience. Once more without feeling. in Emotion and Consciousness. L. F. Barrett, P. M. Niedenthal, and P. Winkielman, ed. Guilford Press, New York, London

Winter KA, Kuiper NA (1997) Individual differences in the experience of emotions. Clin. Psychol. Rev. 17(7):791-821. https://doi.org/10.1016/S0272-7358(97)00057-3

Wittekind DA, Kluge M (2015) Ghrelin in psychiatric disorders - A review. Psychoneuroendocrinology 52:176–194. https://doi.org/10.1016/j.psyneuen.2014.11.013

Wosu AC, Valdimarsdóttir U, Shields AE, et al (2013) Correlates of cortisol in human hair: Implications for epidemiologic studies on health effects of chronic stress. Ann Epidemiol 23:797-811.e2. https://doi.org/10.1016/j.annepidem.2013.09.006

Yeates JW, Main DCJ (2008) Assessment of positive welfare: A review. Vet J 175:293–300. https://doi.org/10.1016/j.tvjl.2007.05.009

Yu X, Qiao S, Wang D, et al (2016) A metabolomics-based approach for ranking the depressive level in a chronic unpredictable mild stress rat model. RSC Adv 6:25751–25765. https://doi.org/10.1039/C6RA00665E

Zajonc RB (1965) Social facilitation. Science 149:269-274. https://doi.org/10.1126/science.149.3681.269

Zebunke M, Puppe B, Langbein J (2013) Effects of cognitive enrichment on behavioural and physiological reactions of pigs. Physiol Behav 118:70–79. https://doi.org/10.1016/j.physbeh.2013.05.005

Ziomkiewicz A, Babiszewska M, Apanasewicz A, et al (2021) Psychosocial stress and cortisol stress reactivity predict breast milk composition. Sci Rep 11:1–14. https://doi.org/10.1038/s41598-021-90980-3

Zheng P, Fang Z, Xu X-J et al (2016) Metabolite signature for diagnosing major depressive disorder in peripheral blood mononuclear cells. J. Affect. Disord. 195:75-81. https://doi.org/10.1016/j.jad.2016.02.008

Acknowledgements

When I started this PhD project, I never *truly* thought that it would end one day. And yet, here we are! During these four years (and a half), I have crossed paths with many people who influenced my PhD trajectory and made my PhD experience unforgettable. Now is finally the time to properly thank them.

My first thanks go to **Laura** for supervising my thesis. No matter the circumstances, you were always there to support me and to guide me whenever I needed it. You gave me many professional opportunities – from writing a review to giving lectures for your course. You are an idea-generator and a project-machine on your own. Despite your busy schedule, you always take the time for your colleagues and your students to celebrate big (or small) achievements and to encourage them during difficult times. I will always remember when you came all the way up North to help me collect milk samples at 5:00 AM just because I needed social contacts. I won't forget neither that you came to my place on several occasions for our weekly meetings when I was no longer able to drive without being sick. I highly doubt that many supervisors would have done the same for their students and, for that and so much more, I am extremely grateful.

With all my heart, I also want to thank **Kees** for supervising my thesis. You always see the bigger picture in every project and you make sure that we can achieve the impossible if we want to – with the resources that we need. You simply are the master of negotiations. I have always felt like my project would be safe and I knew that you would solve any challenge that would pop up along the way if I could not tackle it myself. These four years have been quite a journey for all of us, and not once have I heard you complain about anything (which, from a French perspective, is actually suspicious…). Even worse, you have not stopped once working with great enthusiasm! (Also, thank you for understanding my humour most of the time).

I also want to thank **Joop** for his invaluable help during my experiments. I consider you as my third supervisor and we would have never achieved what we have achieved without you. I have rarely met someone as kind and humble as you, and I feel so privileged to have worked with you. You have taught me so much about setting-up experiments, staying organized (my main strength according to you hahaha) and becoming a better supervisor/colleague. We have always had each other's back, and my favourite days at Dairy Campus were those when you were there – even though that meant that we would spend hours dealing with camera equipment and solving technical issues.

I am also very grateful for **Eddie**'s work as my promotor. Eddie, you have accepted to be my promotor despite my craziness, and it appears that you even got accustomed to it! During these last few months, you have made sure that the end of my PhD would run as smoothly as possible and you helped me prioritize my work. I would like to thank you for arranging all the necessary paperwork for my PhD defence, and – above all – I want to thank you for making sure that I could still work on my PostDoc proposal in spite of a (delightful but unexpected) delay. Thanks also for letting me use your office when you were on holidays, I felt like a boss for a couple of weeks – and I absolutely loved it hahaha! The only thing I regret so far is not having seen you dancing, but I am sure that we can fix it in the near future.

I also would like to thank **Imke** for her support (and the fun that we had at the farm when you visited us at Dairy Campus!). I wish I could have spent more time with you during my PhD. You have always given me excellent advice and I have never felt so well understood by anyone else. You just seem to know who I am, what I want and what I need.

Overall, I wish to thank everyone at Animal Production Systems who contributed to the family-like spirit in our department. Due to the length of my experiments and to Covid, I only spent six months in Wageningen, but these were fantastic six months. What a pleasure to work hard and still have the feeling to be on holiday every day! **Fokje**, thank you for being such a great listener and for your professional and personal advice. Every PhD student needs a Fokje to succeed. **Lia**, many thanks for your constant happiness and our small-talk about what matters the most: our pets! **Cindy**, **Iris** and **Aart**, I was very lucky to end up in your non-PhD room. You were so welcoming and reassuring, I was really fortunate to be able to share my doubts and struggles with you. **Akke**, **Margret**, **Abi**, **Titis**, and **Windi**, many thanks for our numerous laughs, stimulating conversations, and for your unshakeable support. I wish you nothing but the best for the future! May our paths cross again. Of course, I do not forget my colleague **Francesca** from WLR either. Thank you for all your advice and encouragement!

I cannot thank enough the amazing team of statisticians who helped us conduct countless (and various) statistical analyses. **Bas** and **Sabine**, you have helped me tremendously for most part on my PhD. You have always taken the time to answer all of my questions – from the stupidest to the most complex ones – and I have learnt so much thanks to you. I also want to thank you for our small-talk at the beginning of each meeting, they always helped creating an inspiring and stimulating working environment! **Jasper** and **Gerrit**, many thanks for joining the Biomarkers project during my final year.

Jasper, you explain complex statistical concepts like no one else, and it is a real pleasure to listen to you talking about statistics. May we find a positive result together one day (0/694 so far hahaha). **Gerrit**, it has been very fun and enlightening working with you. We have tackled numerous issues together and went to the bottom of each analysis like real detectives sometimes! Thank you so much for your reactivity, your patience and for always accepting my “emergency” meetings.

I will not forget my second family at **Dairy Campus** either. After all, I spent most of my PhD at the farm with you and my sweet little cows. **Fedde**, **Martin** and **Frank** – you have made my life so easy. There is not one problem that we have not managed to fix together within 24h. Your door was always open whenever necessary, and you always did everything possible to make sure that I could conduct my experiments under the best conditions possible. **Eke** and **Tiny**, thanks for your kindness and support with the (numerous) students. **Hans** and **Jan**, I cannot thank you enough for all your help with the blood collection, the faecal collection and the acute-stress tests. You have both been very patient with me and you have taught me everything there is to know about veni-punctures. One day, I will be as good as you at taking blood samples (without panicking!) and I will come back to DC to prove it to you in a blood-sample racing contest! **Edwin** and **Manca**, thanks for being such great roommates and for inviting me to all these crazy parties (while we were still young and irresponsible ;-)). I have not discovered the Friesian Islands but there is not one bar that I do not know in Leeuwarden thanks to you. **Sybren**, many thanks for your kindness, support and interest in my project at all times! **Durk**, thanks for always delivering the feed on time and managing my barn like no one else! You have all made sure that my experiment (and our cows’ wellbeing) was always the number one priority.

I could not have written this PhD thesis without the help and resolute motivation of the students who took part in this PhD project. You have all accepted to work in the evenings or during the weekends when our beloved meisjes had some trouble learning the JBT, and for that, I am extremely grateful. I could honestly write a book about everything that we shared together – from the bird rescue to the sudden manure flood – but I will restrict myself to one sentence per (pair of) student(s) in chronological order. **Jauke**, thanks for the late dinner with the cows. **Zjane**, thanks for staying positive when we had to work outside at -20°C. **Nolwenn**, you are an example of resilience for us all. **Joanne**, you are a true team player, and it was a real pleasure working with you. **Shen**, many thanks for our late philosophical discussions about animal welfare, your love of learning will take you far! **Jacinta**, I am honored to have worked so much with you: it is a privilege to have seen you bloom in your personal life and your professional career. **Sarah**, thanks for sharing your veterinarian knowledge with your beautiful accent (and for teaching us new words!). **Sietske**, thank you for your perseverance despite the numerous challenges. **Marion** and **Sabine**, I hope that we will have the opportunity to work again together – we were a fantastic team. I would redo this experiment by your sides over and over again if it was possible. I am looking forward to seeing you again in Brittany. **Pleunie**, you were always attentive, creative and cheerful – I was delighted to work with you! **Marlies**, your strength and sense of priorities will be essential skills in your future career. **Ilse** and **Suze**, I am eternally grateful for your help with the dog statue: you saved the day and the experiment! **Gustavo**, I appreciated your dedication to knowing each cow as a unique being: your passion for cows was inspiring. **Eléa**, I could not have wished for a better student to finalize this experiment: you are hard-working, smart, supportive and cheerful.

I also wish to thank my friends and family members, who made my experience abroad remarkable. **Corentin** and **Nevena**, you made my journey in the Netherlands unforgettable from the very beginning. Living with you two was simply awesome! **Coco**, thanks for introducing me to your business and gaming world. I will never forget our last night before your departure on the parking lot with Clément! What a fantastic evening for a good-bye party, hahaha! **Nena**, we made our first steps in the streets of Ede together – and you were the best travel buddy possible. I still can't believe how lucky I am to have met you. You are a real role model for me, and I want to thank you for all your advice. Thank you also for our girl nights, and for agreeing to save the ducks in the middle of the road – even though we created a traffic jam and ran late to work!

Sun, many thanks for our philosophical talks in Leeuwarden and elsewhere. You are the sweetest of us all, and I won't forget any of your little attentions! I am looking forward to hearing about the outcomes of your study, and I hope to see you very soon!

Cécile et **Hadrien** (et vos deux crapules poilues !), merci d'avoir toujours été là pour Clément, Zouzou, Clochou et moi. **Cécilou**, I cannot remember the first day that I met you – it feels like you have always been around, at Dairy Campus and in Wageningen. You are the friend that everyone would like to have, and you have all the qualities that a human being should have (just like Clément)! Thanks for not judging me when I uncontrollably laugh in inappropriate situations hahaha, it could have ruined our friendship from the very beginning! **Hadri**, we are all very lucky and happy that you have moved to the Netherlands! I love your jokes and crazy stories. We will be there to support you during your PhD project, just like you were there for us!

My parents-in-law, **Cathy** et **Benoit**, thank you for your kindness and professional advice. Merci aussi de me tolérer même dans les moments les plus difficiles – Noël inclus ! **Jipé** and **Fiona**, illustrating my book by your sides was a real pleasure. I could not have dreamt for a better result, and I am eternally grateful!

Papa et **Cornélia**, many thanks for your free coaching sessions! **Papa**, thanks for helping me become a better supervisor and for advising me whenever I needed it. Many thanks also for spending Christmas in the Netherlands with us: you came to us when we could not come to you! **Cornélia**, many thanks for encouraging me to write and write and write for my PhD thesis. I followed our writing plan every single day!

Maman et **Patrice**, I cannot thank you enough for all your help and support during the last – and rather intense – phase of my PhD. **Mamoune**, you made sure that during these last three months, I only had to take care of my two babies: the professional one and the non-professional one. Thanks to you, I did not have to spend time and energy moving in my new house or visiting it hahaha. I would not have made it without you. **Patou**, thanks for putting us up in Rennes. The conditions were ideal to finalize this thesis in peace!

My siblings, **Charlotte** and **Pierre** – you are the best siblings on earth! **Choune**, you made this COVID time brighter for me. Working with you on Teams, day after day, was an absolute delight and it was very motivating. We worked like dogs together, and we had such a great time at it! I am looking forward to doing it again! Thank you for making sure that I was taking enough breaks. Thank you for allowing us to go on holiday with Clément. Thank you for being here. **Pitt**, you are my favourite twin and the best uncle possible already. I do not have the words to thank you enough for your support, your meals, and your hilarious jokes (including those on my thesis). I was so lucky to spend these final months of my PhD thesis by your side, after 5 years apart! I cherish every memory that I have with you: I would never have thought that I would have enjoyed this last stretch of my thesis so much – and this is thanks to you.

Banzai and **Clochette**, I want to thank you for your hairy, purry and woofy support. You have made all kinds of weird noises during my meetings, but you are forgiven. A pet as a (working) companion is a must in life, and being locked down with you was a blessing.

My final thanks go to my beloved husband, **Clément**. Research has definitely changed both of our lives. We met during my first research project in Bristol and we got married during my (but I should say “our”) PhD thesis. You are the most patient, compassionate and flexible human being on this planet. You agreed to move to the Netherlands within 5 minutes, to move to Friesland within 2 minutes, and to move back near the university within 1 minute. You are my rock, and I would not want to celebrate any achievements or mourn any failures with anyone else. Our love and our family have considerably grown during our journey abroad, and I cannot wait to see what tomorrow will be made of with you.

About the author

Louise Kremer was born on the 12th of October 1993 in Saint-Cloud, France. After two years of preparatory class (BCPST) in Pierre de Fermat, she was admitted into the higher education and international institute of AgroSup Dijon, France – from which she graduated as an Agricultural Engineer in 2016. During the 2nd year of her Engineering School, Louise studied the influence of feed distribution on sticklebacks' behaviour in the University of Bristol, England (2015). During her 3rd year, she assessed the influence of the quality of human-animal relationships on beef cattle behaviours; and she investigated the effects of various light characteristics on piglets' feeding behaviours and sows' welfare and reproductive success (2015-2016).

In 2017, Louise started her PhD at the Animal Production Systems and Livestock Research chair groups of Wageningen University & Research, the Netherlands – where she investigated the potential of different markers as indicators of mood in dairy cows. During her PhD, Louise has developed a Judgment Bias Task specifically for cows, which allows to assess dairy cow positive and negative moods objectively. Researchers can use this task in the future to identify farm practices that cows *truly* perceive as pleasant or aversive. Moreover, Louise investigated the extent to which cows' personality shifts individual perception of the living conditions and questioned the idea of “one-fit-for-all” optimal housing conditions. Finally, she attempted to identify physiological correlates of dairy cow mood in the optic to design a non-invasive tool that would be compatible with on-farm assessments of dairy cow welfare.

Currently, Louise is writing a post-doctoral grant proposal which aims at identifying indicators of resilience in dairy cows. To undertake this work within the Animal Production Systems chair group, Louise was awarded a WIAS fellowship.

Publications

Refereed scientific journals

Herbert-Read JE, Kremer L, Bruintjes R, Radford AN, Ioannou CC (2017) Anthropogenic noise pollution from pile-driving disrupts the structure and dynamics of fish shoals. Proceedings of the Royal Society B: Biological Sciences, 284(1863). https://doi.org/10.1098/rspb.2017.1627

Kremer L, Holkenborg SK, Reimert I, Bolhuis JE, Webb LE (2020) The nuts and bolts of animal emotion. Neuroscience & Biobehavioral Reviews, 113, 273-286. https://doi.org/10.1016/j.neubiorev.2020.01.028

Kremer L, Bus JD, Webb LE, Bokkers EAM, Engel B, van der Werf JTN, Schnabel S, van Reenen CG (2021) Housing and personality effects on judgement and attention biases in dairy cows. Sci Rep 1–18. https://doi.org/10.1038/s41598-021-01843-w

Kremer L, van Reenen CG, Engel B, Bokkers EAM, Schnabel S, van der Werf JTN, Webb LE (2021) Developing a feasible and sensitive judgement bias task in dairy cows. Anim Cogn. https://doi.org/10.1007/s10071-021-01563-8

Abstracts in conference proceedings

Kremer LAM, Webb LE, van Reenen CG, van der Werf JTN, de Boer IJM (2018) *Elaboration of a Judgment Bias Test suitable for dairy heifers*. Poster session presented at WIAS Science Day 2018, Wageningen, Netherlands. https://edepot.wur.nl/445871

Kremer LAM, Webb LE, Bokkers E, & van Reenen CG (2018) *Development of a valid judgment bias test for dairy cattle*. Poster session presented at Really relaxed or deeply depressed? Vienna, Austria.

Kremer LAM, Webb LE, Bokkers E, van der Werf JTN, Engel B, Schnabel SK, & van Reenen CG (2020) Influence of the punisher on the feasibility and sensitivity of a Judgment Bias Task for cattle. In *ISAE Benelux meeting 2020: Elevating animal Lives* (pp. 7-7). International Society for Applied Ethology (ISAE). https://www.applied-ethology.org/res/ISAE%20Benelux%20meeting%202020%20-%20proceedings.pdf#page=7

Kremer LAM, Webb LE, Bokkers E, Engel B, van der Werf JTN, Schnabel SK & van Reenen CG (2021) Is the Judgement Bias Task biased? *WIAS Annual Conference 2021: Reslience.* 26th ed. Wageningen University & Research, p. 37-37 1 p.

Education certificate

Completed training and supervision plan[1]

The Basic Package	**1.8 ECTS**
WIAS Introduction Day	
Course on philosophy of science and/or ethics	
Disciplinary Competences	**12.0 ECTS**
Writing proposal	
Fundamentals of animal emotion	
Bayesian Statistics	
Laboratory Animal Sciences: Design & Ethics *(including specialization on rodents)*	
Laboratory Animal Sciences: Specialisation on Ruminants	
Professional Competences	**8.3 ECTS**
Supervising BSc and MSc thesis students	
Scientific Writing	
Poster and Pitching	
Efficient Writing Strategies	
Project and Time Management	
Presenting with Impact	
Stress Identification and Management	
Writing Propositions for your PhD	
The Final Touch	

---

[1] With the educational activities listed, the PhD candidate has complied with the educational requirements set by the graduate school Wageningen Institute of Animal Sciences (WIAS) of Wageningen University & Research, which comprises a minimum of 30 ECTS (European Credit Transfer and accumulation System). One ECTS equals a study load of 28 hours.

Presentation Skills (max 4 credits)	**4 ECTS**
7th International conference on the WAFL – Ede, 05 to 08/09/2017 – Poster presentation	
WIAS Science Day – Wageningen, 05/02/2018 – Poster presentation	
53rd Congress of the ISAE – Vienna, 19 and 20/09/2018 – Poster presentation	
Benelux ISAE – Virtual meeting, 03/11/2020 – Oral presentation	
WIAS Science Day - Wageningen, 28/04/2021 – Oral presentation	
Teaching Competences (max 6 credits)	**6 ECTS**
Lecturing – Physiological markers of animal affective states in the Fundamentals of Animal Emotion (x2)	
Supervising BSc thesis (x5)	
Supervising students from van Hall Larenstein (x3)	
Supervising MSc thesis from WUR (x6)	
Reviewing Proposals for the Research Master Cluster	
TOTAL (min 30 credits)	**32.1 ECTS**

Colophon

The research described in this thesis was financially supported by the Dutch Ministry of Agriculture Nature and Food Quality, and ZuivelNL (organisation of the Dutch dairy sector), Grant no: TKI-AF-16162

Front cover and illustrations: Jipé Gerardin - Instagram and Twitter: @PunkManouche

Back cover: Your-Comics - https://your-comics.com

Printing: www.proefschriftmaken.nl

ISBN: 978-94-6447-084-0

DOI: https://doi.org/10.18174/562024


54.7