Graduate School for Cellular and Biomedical Sciences

University of Bern

The effects of long-term cognitive training on the behaviour and welfare of goats

PhD Thesis submitted by

Katrina Rosenberger

from **Austria**

for the degree of

PhD in Biomedical Sciences

Supervisor
Prof. Dr. Hanno Würbel
Vetsuisse Faculty of the University of Bern
Co-advisor
Dr. Christian Nawroth
Institute of Behavioural Physiology of the Leibniz-Institute for Farm Animal Biology

Accepted by the Faculty of Medicine, the Faculty of Science and the Vetsuisse Faculty of the University of Bern at the request of the Graduate School for Cellular and Biomedical Sciences

Bern, Dean of the Faculty of Medicine

Bern, Dean of the Faculty of Science

Bern, Dean of the Vetsuisse Faculty Bern

Inhalt

ABSTRACT

Within animal husbandry, the biological needs of animals are rarely met. Captive housing environments are often barren and provide animals only with the basic requirements. The introduction of cognitively challenging tasks has been suggested to put captive animals in a position where they can engage evolved cognitive skills and actively control some aspects of their environment. Current evidence suggests that frequent exposure to cognitive stimulation not only constitutes a need, but also has potential positive effects on the welfare of animals. First results indicate that by increasing behavioural flexibility and reducing stress reactivity, cognitive stimulation could improve an animal's ability to adapt and cope with its environment.

The introduction of automated cognitive enrichment devices in the housing environment of animals has been commonly used to provide cognitive stimulation to animals in captivity. Cognitive tests are increasingly conducted in many zoos as well as research facilities around the world to assess cognitive skills of many species. While the goal of cognitive enrichment devices is the engagement of cognitive skills, cognitive tests primarily aim to assess these. Cognitive tests often require animals to be isolated, handled by a human and/or given food reinforcement by a human. To assess the effects of cognitive testing on welfare, it has to be disentangled from these confounding factors (Chapter I).

In this thesis, it was investigated whether goats want to be cognitively challenged and whether cognitive stimulation, by means of long-term exposure to cognitive tests, might positively affect behaviour and welfare. Goats are a good model species for this research question because lots of literature on their cognitive skills and their reaction to stress exists. For instance, genetic selection for certain traits such as high productivity can lead to reduced motivation to work for feed as well as reduced activity and stress reactivity. To account for this fact and to increase external validity and variability of our results, dairy goats selected for high productivity, and dwarf goats, not selected for production traits were tested.

The first aim of this thesis was to assess whether goats have an intrinsic motivation for cognitive stimulation and thus work for food in a manipulation task that resembles their natural foraging behaviours and requires low effort (Chapter II). We examined whether domestic goats choose to open a sliding door to receive a reward rather than getting the same reward for free. This phenomenon, known as Contrafreeloading (CFL), has been explained by an intrinsic motivation to search for information, novelty, and challenges, amongst other things. Using an Item Response Tree generalized linear mixed model, we found that goats do work for food. Both selection lines of goats are similarly motivated to do so but differ in their performance over several trials.

Farm animals must be able to deal with many stressors such as isolation or handling by humans. To investigate if cognitive stimulation has the potential to improve an animal's ability to cope with these stressors and thus to reduce stress reactivity, we conditioned three treatment groups (Chapter III). Goats from the COG treatment group were tested individually in human-presented object-choice tests. Goats in the POS treatment were isolated individually in the same arena as COG goats but received rewards without being administered the object-choice tests. Goats in the ISO treatment group were isolated individually but neither received a reward nor were administered the tests. Subsequently, we tested all treatment groups in four tests: a Novel Arena test (NA), a Novel Object test (NO), a Novel Human test (NH) and a weighing test (WH) where goats were handled on a weigh scale. To increase external validity, we tested both selection lines (dairy and dwarf goats) at two research sites. We did not find evidence that long-term cognitive testing did have a substantial effect on stress reactivity in any of these tests. However, positive human contact seemed to increase boldness towards a novel object and increased reactivity towards handling in dwarf goats. Furthermore, we found that reactivity towards different stressors is strongly affected by selection line.

As farm animals are exposed to different husbandry systems throughout their life, they need to be able to flexibly adapt to their surroundings. It has been proposed that mastering tasks successfully makes the animal proficient in manipulating its environment, and likely improves behavioural flexibility. Using the same three treatment groups as in Chapter III, we investigated whether cognitive testing improves behavioural flexibility of goats in two conceptually different cognitive tests, namely a spatial A-not-B detour test and an instrumental problem-solving test (Chapter IV). Again, we tested both selection lines (dairy and dwarf goats) at two research sites. We found that cognitive testing *per se* (COG) and exposure to a testing environment via human-given object-choice tests (POS) do not notably affect the performance in subsequent conceptually different cognitive tests in goats.

In summary, we found that two different selection lines of domestic goats are similarly interested in cognitive stimulation and are willing to work for it. Further, we did not find general effects of cognitive testing (COG) *per se* or human-animal-interaction (POS) on responses to different stressors in goats in a novel arena test, a novel object test, a novel human test and during weighing on a scales. Selection lines did differ in some aspects of stress reactivity, but cognitive testing and positive human contact seem to have caused some differences to disappear in the POS and COG dwarf goats. With respect to the goats' detour or problem-solving performance, we found only subtle differences between treatments. Finally, our multi-lab approach in Chapter III and IV allowed us to detect large variances between research sites that should be considered when making claims from data obtained on single sites.

Chapter 1

General introduction

Animal welfare – a multidimensional concept

Animal welfare has become an increasingly important topic over the last decades due to the intensification of livestock production systems, technological innovations, evolving dietary habits, and changes in consumer perception (Alonso, González-Montaña, & Lomillos, 2020; Donald M. Broom, 2019). As animal welfare consists of multiple domains (e.g., behaviour, health, nutrition, environment, mental states) and is perceived by animals subjectively, the definition of welfare is neither definitive nor exhaustive (Hemsworth, Mellor, Cronin, & Tilbrook, 2015; Mellor et al., 2020). In 1986, Broom defined animal welfare as “the animals state as regards its attempts to cope with its environment.”(1986, p. 524). Thus, only if an animal can react to and cope with environmental challenges, its welfare may be termed good (Wechsler, 1995; Wemelsfelder and Birke, 1997).

Initially, the assessment of welfare was focused on negative experiences and mental states that affect biological functioning, health and behaviour (Mellor, 2016). From the early 2000s, more attention was given to the assessment of positive experiences and mental states (e.g., Boissy et al., 2007; Désiré, Boissy, & Veissier, 2002; Mellor, 2016). This was mainly due to increasing scientific literature discussing the concepts of affective states, feelings and emotions in animals (Bekoff, 2000; Boissy et al., 2007; Dawkins, 2000; Désiré et al., 2002; Mellor, 2015) and the recognition that good animal welfare cannot be achieved simply by avoiding negative experiences (Mellor, 2015; Mellor et al., 2020). While the assessment of welfare through biological functions is relatively clear-cut, measuring mental welfare states is more challenging (Brydges & Braithwaite, 2008). Emotions are composed of neurophysiological, behavioural, cognitive and subjective components (Keltner & Lerner, 2010). The subjective component of emotions cannot be assessed directly in animals as they cannot verbally communicate; however, the physiological, behavioural and cognitive components can potentially be used as indicators of welfare in nonhuman animals (Mendl, Burman, & Paul, 2010).

Species-specific behaviour and the role of cognitive capacities

Animals in the wild have to locate mating partners, avoid predators, and/or find and hunt prey in order to survive. They also must learn how to obtain or avoid certain food types; remember food sources and they need to be able to acquire information about their environment via exploration or social learning. Necessarily, these adaptations to a changing physical and social environment, as well as the use of different foraging strategies and food sources led to the development of species-specific behaviour and cognitive capacities (Dunbar, 1998; Shettleworth, 2001). Shettleworth (2001, p. 277) states: “Cognitive processes such as perception, learning, memory and decision making play an important role in mate choice, foraging and many other behaviours”. In captivity, these capacities and cognitive skills are challenged infrequently, or inappropriately (Meehan & Mench, 2007; Spinka & Wemelsfelder, 2017). The inability of an animal to perform species-specific behaviour is a major source of stress (Morgan & Tromborg, 2007) which can impair health and welfare. More recently, the stimulation of evolved cognitive capacities has been linked to the mental well-being of an animal and its lack may result in boredom and anhedonia (Fureix and Meagher, 2015, Meagher and Mason, 2012). Thus, the performance of specific behavioural patterns and activities as well as the stimulation of the cognitive processes involved in these activities are prerequisites for good animal welfare and thus might be called needs.

Behavioural needs and how they were affected by domestication.

A behavioural need has been described as a behavioural pattern that the animal apparently needs to perform, regardless of environmental stimuli (Mench, 1998). The seeking of novelty and exploration of novel environments may be examples for such behavioural needs, as the prevention of performance of these behaviours has been found to cause frustration in animals, especially in intensive housing systems (Studnitz, Jensen, & Pedersen, 2007; Wood-Gush & Vestergaard, 1989). According to the “Information Primacy Theory” by Inglis (1997) in nature, animals gather information about potential food sources where food shortages can occur. The more information an animal has, the better it can adapt to its environment and the higher its chance to survive (Dall, Giraldeau, Olsson, McNamara, & Stephens, 2005). This behaviour may be partially explained by an intrinsically motivated drive to manipulate the environment (Harlow, Harlow, & Meyer, 1950).

Domestication

Behaviour of domesticated animals does not differ much from that of ancestors and thus, domestic animals have similar behavioural needs (E. Price, 1999). One common definition of domestication has been suggested by Prize (1984) and states that “the process of animal domestication involves adaptation, in particular, adaptation to man and the environment he provides”. Through the intervention of man in the lives of animals and their subsequent artificial selection, certain behavioural traits that were originally necessary for survival have been genetically changed (Price, 1999). However, most natural behaviours were not unlearned or only slightly reduced to fit captive conditions and domesticated animals normally still show almost all of their species-specific behaviours (P. Jensen, 2006; Price, 1999). For example, the nest building behaviour in sows has not been lost during the process of domestication and is shown even in pigs kept in modern husbandry systems (Wischner, Kemper, & Krieter, 2009). If an animal cannot perform an ethologically motivated behaviour and thus a behavioural need, its welfare can be impaired and it may be suffering (Jensen & Toates, 1993).

Selection for high productivity

It has been documented that artificial selection for high productivity in farm animals can lead to changes in other traits, such as health or reproduction (as reviewed by Rauw, Kanis, Noordhuizen-Stassen, & Grommers, 1998). For example, the selection for high productivity in many farm animal species has reduced stress reactivity during human-animal interactions (Campler, Jöngren, & Jensen, 2009; Colpoys et al., 2014; Lindqvist & Jensen, 2008; Schütz & Jensen, 2001), and altered the reaction to isolation (Kilgour & Szantar-Coddington, 1995; Romeyer & Bouissou, 1992). These “side-effects” of selecting for high productivity may be explained by a resource allocation towards productivity and away from other energy demanding traits (Beilharz, Luxford, & Wilkinson, 1993). For instance, in chicken bred for high egg productivity, the motivation to forage, explore, and feed from a hidden food source (contrafreeloading) has been reduced in comparison to red jungle fowl, which have not been selected for high productivity (Lindqvist & Jensen, 2009; Schütz & Jensen, 2001). The high artificial selection pressure on specific traits did not only affect behavior but also cognitive abilities, by means of altering brain morphology (Mehlhorn & Caspers, 2021). In laying hens, the selection for high egg productivity affected learning behavior with high productive hens showing enhanced learning strategies compared to less productive hens (Dudde, Krause, Matthews, & Schrader, 2018). Thus, selection apparently shaped adaptive strategies of domesticated animals towards more predictable man-made environments, but the motivation to perform species-specific behaviour such as foraging for food and engage evolved cognitive skills did not disappear.

Cognitive stimulation as a behavioural need

As defined by Shettleworth (2001, p. 278) cognition "includes perception, learning, memory and decision making, or in other words, all ways in which animals take in information about the world through the senses, process, retain and decide to act on it". Cognitive abilities are often categorised in two domains: the domain of physical cognition including e.g. knowledge of space, quantities, and causality, and the domain of social cognition including e.g. communication with humans, attributing mental states to others (Kaminski, Call, & Tomasello, 2008; Nawroth, 2017). Over the last decades, cognitive research in farm animals has become of increasing interest and some of the many cognitive capacities farm animals possess have been unravelled (Gieling, Nordquist, & van der Staay, 2011; Nawroth et al., 2019; Nicol, 1996). For instance, goats can solve visual discrimination and spatial detour tests (Langbein, Siebert, & Nuernberg, 2008; Langbein, Nürnberg, & Manteuffel, 2004; Nawroth, Baciadonna, & McElligott, 2016; Raoult, Osthaus, Hildebrand, McElligott, & Nawroth, 2021). Horses have the ability to discriminate the quantity of different abstract symbols (Gabor & Gerken, 2014). And pigs are able to able to differentiate between different amounts of food (Held, Baumgartner, Kilbride, Byrne, & Mendl, 2005). As the outcomes of cognitive tests can be used as the basis for husbandry recommendations (Croney, Gardner, & Baggot, 2004; Zobel & Nawroth, 2020), investigating cognitive abilities of animals can help to improve animal welfare.

It has been even suggested that animals *want* to be challenged (Clark, 2017; Spinka & Wemelsfelder, 2017). The phenomenon when animals invest effort to access a resource even if the identical resource is available for free has been called contrafreeloading (CFL) and has been described in many captive species (reviewed by Inglis et al., 1997). Neuringer (1969) was one of the first to demonstrate that pigeons will peck at a disc repeatedly to receive food when the same food was also available for free. Not only pigeons have been found to be motivated to work for resources such as food. Similar results have been shown in many farm animals such as pigs, cattle and chickens (de Jonge, Tilly, Baars, & Spruijt, 2008; Hessle, Rutter, & Wallin, 2008; C. Lindqvist & Jensen, 2008; Van Os, Mintline, DeVries, & Tucker, 2018). In a study by Langbein et al. (2009), dwarf goats readily used a touch screen on an automated drinking device to receive water although water was also available for free. Meagher et al. (2020) found that cattle are highly motivated to work to access a discrimination learning task and concluded that learning tasks can offer welfare benefits. As there is evidence that animals will voluntarily engage in opportunities to use their cognitive abilities, the stimulation of cognitive functions might be a behavioural need and important for the welfare of captive animals (Meehan & Mench, 2007).

The functional reason of CFL is still debated, but there are several, not mutually exclusive, theories. According to the “Information Primacy Theory” by Inglis (1997), CFL is a form of acquiring information of potential future food sources and thus adaptive in an unpredictable environment where food shortage can occur. Another theory to explain CFL is White’s Competence Theory (White, 1959) which suggests that animals have a need to control and manipulate their surroundings. Performing a cognitive task may give the animal a feeling of control over its situation (Bassett & Buchanan-Smith, 2007; Singh & Query, 1971). It has also been suggested that solving a cognitive challenge successfully is stimulating in itself (Jensen, 1963). If being involved in a problem-solving experience is itself rewarding and may lead to positive emotions, we can assume that an animal will seek cognitive stimulation even if no extrinsic motivation exists. Indeed, chimpanzees have been observed to seek interaction with a cognitive challenging device even if they are not provided a food reward or other extrinsic positive reinforcement (Clark & Smith, 2013). Thus, the stimulation of cognitive processes may serves different purposes, but nevertheless appears to be important to animals and linked to their welfare.

Cognition and welfare are linked

The association between cognition and welfare is complex and not uni-directional (Franks, 2019). Cognitive processes can influence welfare, but at the same time, cognition can be affected by an animal's welfare state (Mendl, 1999). For example, the understanding that an object exists, even if it is out of sight ('object permanence'), can improve the predictability of an environment for the animal and thus its welfare (Nawroth et al., 2019). Conversely, if an animal's welfare is impaired due to fear or stress, its cognitive functions may be impaired (i.e. Regolin, Vallortigara, & Zanforlin, 1995). For example, Regolin (1995) used detour tasks to assess object permanence in chicks and reported that their performance was impaired when the chicks were tested under stressful conditions such as social isolation.

Based on current findings in animal welfare research, it becomes obvious that animal welfare and cognitive processes are inevitably linked. Most research to date has focused on the link between welfare and emotions (Baciadonna, Duepjan, Briefer, de la Torre, & Nawroth, 2018; Boissy et al., 2007; Paul, Harding, & Mendl, 2005), whereas the need of animals to receive sufficient cognitive stimulation in captive environments to provide good welfare has been rather neglected (Boissy & Lee, 2014; Brydges & Braithwaite, 2008). In the past decade though, the importance of cognitive stimulation for animal welfare has gained interest from researchers (Clark, 2011, 2017; Manteuffel, Langbein, & Puppe, 2009; Meyer, Puppe, & Langbein, 2010) and as a result the environment of captive animals was increasingly enriched with opportunities to engage cognitive skills in addition to more common enrichment forms such as opportunities to search for food or engage with conspecifics (Clark, 2011, 2017).

Environmental enrichment to account for behavioural needs

Captive animals in farms, zoos or laboratories often live in very restricted environments (Newberry, 1995). The performance of species-specific behaviour may be impaired in such barren housing conditions causing stress in animals (Morgan & Tromborg, 2007). According to Newberry (1995), modifications to the environment causing an improvement in the biological functioning of captive animals can be defined as environmental enrichment. The application of environmental enrichment is originating from the zoo setting and has also been defined as "...an animal husbandry principle that seeks to enhance the quality of captive animal care by identifying and providing the environmental stimuli necessary for optimal psychological and physiological well-being" (Shepherdson, Mellen, & Hutchins, 1999). Environmental enrichment does not only aim to provide physical occupation but also appropriate challenges for psychological stimulation, hence different types of enrichment can be classified.

Structural enrichment

Structural enrichment is the most applied type and defined as a temporary or permanent provision of objects or substrates in the housing environment. Enrichment items have to be chosen carefully, i.e., they should be stimulating biologically/ecologically relevant behaviour of a species to make them effective tools to improve wellbeing. For instance, hiding and climbing are ecologically relevant behaviours for goats. These behaviours may even be necessary for physiological and psychological wellbeing of this species. Indeed, the opportunity to climb and feed in elevated spaces has been found to improve welfare by increasing feeding bout duration and decreasing disruption when lying (Aschwanden, Gygax, Wechsler, & Keil, 2009). Enriching the captive environment with structural elements to climb and hide allows the animal to perform these natural behaviours and thus may serve a behavioural need (Zobel & Nawroth, 2020; Zobel, Neave, & Webster, 2019). Depending on the aim of the enrichment, the timing and amount of provision has to be right, in order to avoid a loss of interest (e.g., Jones, Webster, & Vea Salvanes, 2021; Wells, 2004). In order to stimulate paying behaviour in dogs, Wells (2004) provided them with 5 different types of toys on separate days and found that dogs spent relatively little (<8%) of the overall observation time playing with the toys. Their interest in the toys also waned over time. Thus, the authors suggested that regular rotation of such enrichment items is necessary to prevent habituation and increase the welfare benefit.

Sensory enrichment

Sensory enrichment contains acoustic, olfactory and visual stimulation and is most commonly used in zoo animals (de Azevedo, Cipreste, & Young, 2007). The best effects will most likely be achieved by targeting the most dominant senses, but will also depend on species, sex, age, and housing conditions (Wells, 2009). For instance, the introduction of olfactory enrichment (rose and ginger oil, vanilla, orange, mint, and banana extracts) to zoo-housed Rothschild giraffe decreased inactivity and altered exhibit utilization in the short-term (Fay & Miller, 2015). Thus, scent provision encouraged more species-specific behaviour by the introduction of novelty into the zoo environment targeting one of the goals of environmental enrichment. In another study by Powell (1995), African lions were introduced to scents of musk cologne, allspice and extracts of peppermint and almond. The study found that the scents increased the activity level of the lions. In terms of sounds, Kogan et al. (2012) found that kennelled dogs were found to spend more time showing behaviours indicative of relaxation (i.e. decreased barking, increased sleeping) when exposed to classical music. Individuals differ in their preferences e.g. for smells (Fay & Miller, 2015). Thus, the kind of stimuli used for enrichment has to be carefully chosen to avoid unnecessary additional stress. For instance, in felids some odours have been found to have arousing effects (de Resende et al., 2011; Wells & Egli, 2004), which may want to be avoided if a reduction of activity or stress is the aim of the enrichment. Visual stimulation by means of moving televised video images as a form of enrichment have been found to induce positive behavioural changes in a wide variety of species such as primates (e.g. Platt & Novak, 1997), chicken (Clarke & Jones, 2000), and kennelled dogs (Graham, Wells, & Hepper, 2005). Surprisingly, in her review, Wells (2009) mentions computer-based challenges as a form of visual enrichment, while other studies use computer-based tasks to cognitively enrich animals (Langbein et al., 2004).

Social enrichment

Social enrichment may be provided in two ways. Most commonly it describes the introduction of conspecifics into the housing environment of an animal (e.g., De Paula Vieira, von Keyserlingk, & Weary, 2010; Toyoshima, Yamada, Sugita, & Ichitani, 2018; Wells, 2004) allowing it to perform species-specific social behaviour. However, also the interaction with a human may be seen as enriching and can have positive effects on behaviour if the human-animal-relationship is good (Baker, 2004; Cerrone, 2019). The positive effects of keeping animals in social groups on behaviour and cognitive development have been first assessed mainly in laboratory animals (e.g. rhesus monkeys: Sánchez, Hearn, Do, Rilling, & Herndon, 1998; Schapiro, Bloomsmith, Porter, & Suarez, 1996), but are increasingly appreciated and adopted also in farm animal husbandry due to evidence on improved behaviour and performance (De Paula Vieira, de Passillé, & Weary, 2012; De Paula Vieira et al., 2010) as well as improved cognitive functions (Gaillard, Meagher, Von Keyserlingk, & Weary, 2014).

Cognitive enrichment

The concept of cognitive enrichment has initially been applied to zoo animals to improve welfare through rewarded instrumental learning (Carlstead & Shepherdson, 2000). In 2007, cognitive enrichment was the least commonly applied type of enrichment (de Azevedo et al., 2007). Since then, appropriate cognitive challenges for captive animals have been suggested by many more researchers to promote psychological well-being (Meehan & Mench, 2007;

Meyer et al., 2010; Spinka & Wemelsfelder, 2017). However, it is difficult to proof that an animal is mentally enriched.

In contrast to other forms of enrichment, cognitive enrichment is difficult to define. For instance, while it is intuitive that providing climbing racks for animals that have the behavioural need to climb fall into the category of structural enrichment, it is hard to decide if a push gate to access food is experienced as cognitive stimulation or is rather seen as physical work to access a resource for an animal. According to Clark (2011, p. 6), cognitive enrichment "(1) engages evolved cognitive skills by providing opportunities to solve problems and control some aspect of the environment, and (2) is correlated to one or more validated measures of well-being". The idea of using instrumental learning as enrichment is based on the assumption that welfare of animals in captivity is compromised by conditions that are neither resolvable nor controllable. This also applies to farm animals which are living in often barren environments with no possibility to execute control over their environment. Broom (Broom, 2008, p. 81) states “Poor welfare is often associated with lack of control over interactions with the environment of the individual, i.e. with difficulty in adapting”. This theory is linked to the concepts of competence and agency (White, 1959). It postulates the need of animals to control and modify their surroundings. In addition, the successful learning of the link between the performance of a specific behaviour and its functional outcome has been also suggested to reduce uncertainty and improve predictability and controllability for an animal (Bassett & Buchanan-Smith, 2007). Zebunke et al. (2013, p. 70) stated that “Through cognitive interaction with the environment, the animals regain a certain control over their environment”. Cognitive stimulation may be a promising way to flexibly provide farm animals with challenging situations and with opportunities to gain some control over their environment and has been found to have a broad range of welfare impacts (Jan Langbein, Nürnberg, Puppe, & Manteuffel, 2006; Jan Langbein, Siebert, Nuernberg, & Manteuffel, 2007; Oesterwind, Nürnberg, Puppe, & Langbein, 2016; Puppe, Ernst, Schön, & Manteuffel, 2007; Zebunke et al., 2013).

It should be kept in mind that only suitable challenges that can be solved by the individual have the potential to improve welfare. Boredom and anhedonia may be caused in case of under-stimulation (Fureix & Meagher, 2015), while an unsolvable problem can lead to anxiety and distress, and hence to physiological and behavioural problems such as learned helplessness (Maier & Seligman, 1976). According to Meehan and Mench (2007) appropriate challenges for captive animals may elicit initial frustration but they should be solvable via the application of cognitive and behavioural skills. In a study by Langbein et al. (2004), goats that were objected to a novel cognitive task initially showed an increase in heart rate indicative of stress. After they had learned to master the task successfully, heart rate decreased again (Jan Langbein et al., 2004) which may be a sign of positive stress (=eustress, Selye, 1975). Therefore, it is recommended that the cognitive capacities need to be carefully considered at the species level, but also at the individual level when implementing tasks to cognitively enrich animals. Challenges for the purpose of cognitive enrichment should not be forced upon the animal and adapted to the individual's skill level to ensure positive effects on welfare.

Cognitive enrichment and its impact on welfare-related parameters

Effects on reactivity towards different stressors

Captive animals are frequently exposed to stressors related to housing and husbandry routines that include the confrontation with novel stimuli to various extents (e.g., during health checks, temporal separation or relocation and cleaning of facilities), separation from the group or handling by humans. Current evidence suggests that frequent exposure to cognitive stimulation/training may alter behavioural responses to future challenges and thus can have positive effects on several aspects of welfare such as human-animal interactions or the reactivity towards novelty and towards isolation (Oesterwind et al., 2016; Puppe et al., 2007; Zebunke et al., 2013).

Novelty

Novelty is often a potent trigger of fear reactions in animals. Nevertheless, it is unavoidable that animals will be confronted with many novel stimuli during their life. The novel object test is often used to assess stress reactivity towards a novel stimulus and to investigate the motivation to explore (Finkemeier, Oesterwind, Nürnberg, Puppe, & Langbein, 2019; Sneddon, Braithwaite, & Gentle, 2003). There is empirical evidence for the positive relationship between welfare and exploratory tendency whereby the absence of curiosity and exploration can be indicative of stress (Forkman, Boissy, Meunier-Salaün, Canali, & Jones, 2007). Environmental enrichment increases the animals exposure to novel stimuli and has been successfully applied in farm animal husbandry where it decreased fear responses to novel objects (Beattie, O'Connell, Kilpatrick, & Moss, 2000; Hillmann, Von Hollen, Bünger, Todt, & Schrader, 2003; Oesterwind et al., 2016). Similar effects have been documented for the introduction of cognitive enrichment with cognitively enriched animals being more explorative in a novel object test than non-enriched ones (Oesterwind et al., 2016; Puppe et al., 2007; Zebunke et al., 2013). For example, in a study by Oesterwind et al. (2016), the impact of structural and cognitive enrichment on learning as well as on the behavioural and physiological responses of goats was investigated. While structural enrichment increased activity of the animals in a novel arena test, cognitive enrichment increased curiosity towards and duration of contact with an unknown object in a novel object test. Similarly, a study investigating the effect of a call-feeding station as cognitive enrichment found that it increases exploration and reduces fear in stressful situations such as a in novel arena (Zebunke et al., 2013). This may be attributed to an improvement in behavioural competence in challenging situations. The previously mentioned novel arena and novel object tests do not only assess stress caused by novelty, but likely also stress caused by separation from the group, as both are commonly testing animals in isolation.

Separation from the group

Captive animals are frequently separated temporarily for veterinary treatments, management, or experimental procedures. Separation from the group is stressful for most social animals and has been found to increase vocalisations, heart rate and cortisol levels (Aschwanden, Gygax, Wechsler, & Keil, 2008a; Boissy & Le Neindre, 1997; Da Costa, Leigh, Man, & Kendrick, 2004; Patt et al., 2013; Siebert, Langbein, Schön, Tuchscherer, & Puppe, 2011). Even though habituation can help to reduce stress responses towards many novel stimuli, repeated confrontation to isolation in goats does not appear to lead to habituation or sensitisation (Siebert et al., 2011). Further, Siebert et al. (2011) stated that isolation should be avoided or only applied if absolutely necessary (e.g., essential veterinary treatments), as it does impair welfare. The open field or novel arena test has been used to assess behavioural reactions towards isolation and a novel environment in many species (Graunke, Nürnberg, Repsilber, Puppe, & Langbein, 2013a; Neave, Costa, Weary, & von Keyserlingk, 2018; Oesterwind et al., 2016; Prut & Belzung, 2003). Cognitive stimulation has been found to help increase the ability to cope with stress in this type of test. For example, in pigs, the introduction of cognitive stimulation in the form of an automated call-feeding station altered behaviour in an open field test towards a reduction in activity as well as reduced vocalisation, lower escape rates and excretion behaviour compared to controls (Puppe et al., 2007). The reaction to separation from conspecifics may differ depending on the presence or absence of a familiar human in the test pen (Tallet et al., 2014).

Human-animal interaction

The relationship between human and animal can strongly influence an animals' welfare. Positive experiences in early life stages are reported to have durable positive effects on future interactions (e.g. goats: Lyons, 1989). In a study by Lesink et al. (2001), calves that received additional human contact interacted longer with the unfamiliar person, were easier to load for transport and had lower heart rates than calves which received only minimal contact by humans. Similarly, previous positive handling or human contact has been shown to increase the readiness of animals to approach people (Hemsworth, Verge, & Coleman, 1996; Jago, Krohn, & Matthews, 1999; Waiblinger, Menke, & Fölsch, 2003), and to affect behaviour in the presence and absence of a human in unfamiliar environments (Boivin, Tournadret, & Le Neindre, 2000; Tallet et al., 2014; Tallet, Veissier, & Boivin, 2009). In contrast, negative experiences with humans have been found to increase fear and impair welfare in farm animals (Breuer, Hemsworth, & Coleman, 2003). The perception of and reactivity towards a human is not only affected by past experiences, but also by genetics (Le Neindre, Poindron, Trillat, & Orgeur, 1993; Murphey, Duarte, & Torres Penedo, 1981). For instance, selection for high productivity has been shown to reduce stress reactivity during human-animal interactions (Boissy et al., 2005; Campler et al., 2009; Colpoys et al., 2014; Schütz & Jensen, 2001).

Whether cognitive stimulation can have positive effects on the human-animal relationship or vice versa has not been investigated yet. Most studies to date that assessed the effects of cognitive enrichment used computer-controlled learning devices where human-animal contact is not required (e.g., Jacobson, Kwiatt, Ross, & Cronin, 2019a; Jan Langbein et al., 2004; Manteuffel et al., 2009; Puppe et al., 2007; Whitehouse, Micheletta, Powell, Bordier, & Waller, 2013; Yamanashi & Hayashi, 2011; Zebunke et al., 2013). For the purpose of scientific experiments however, animals are often required to be trained and handled and/or given food reinforcement by a human (e.g., Morton, Lee, & Buchanan-Smith, 2013; Nawroth, Brett, & McElligott, 2016; Nawroth, Von Borell, & Langbein, 2014). The favoured technique nowadays is called positively reinforced training and supposed to improve the human-animal relationship by building trust and providing rewards to the trained subjects (Brando, 2012; McKinley, Buchanan-Smith, Bassett, & Morris, 2003). As a result, positively reinforced animals may associate the presence of a familiar human with food which may reduce fear responses towards novel humans and ease handling in future situations (Boissy & Bouissou, 1988; Boivin, Neindre, & Chupin, 1992).

Effects on behavioural flexibility

Behavioural flexibility is described as the ability of an individual to change its behaviour by developing new responses when faced with novel challenges or altering already hard wired responses to familiar stimuli (Leal & Powell, 2012). The more behavioural flexible an animal, the better it will cope with its husbandry setting. Thus, in addition to the level of stress an animal experiences, also the level of behavioural flexibility and problem-solving ability of animals under our care has implications for their welfare in several ways. The ability to behave flexibly and to adapt to changing environments is especially important for farm animals as animals may be regrouped as they age, housed in several different pens or locations over their life cycle, and may have to learn to use different feeders, drinkers or gates. Memory of previously learned information can help to facilitate subsequent learning in a conceptually similar subsequent task, an effect called learning-to-learn (Jan Langbein, Siebert, Nürnberg, & Manteuffel, 2007). Previous exposure to cognitive challenges requiring behavioural flexibility thus may have the potential to also improve this flexibility in future situations by the process of 'learning-to-learn'. Therefore, the introduction of cognitive stimulation into the husbandry systems of farm animals may ease adaptation to novel environments and situations by increasing the animals' behavioural flexibility. Facilitation of the adaptation to housing conditions and husbandry procedures may in turn reduce stress for the animal and thus improve its welfare.

The ability to adapt to and flexibly react to environmental challenges is regarded as essential for survival in the wild, but also significant for farm animal welfare (Reader & Laland, 2002; Wechsler & Lea, 2007). High behavioural flexibility is usually associated with species that live in complex social structures and that possess a rich repertoire of foraging strategies (R. I. M. Dunbar & Shultz, 2007; Reader & Laland, 2003). Behavioural flexibility is dependent on a set of skills including good inhibitory control which is defined as an individual's ability to inhibit an impulsive and possibly harmful response to allow for a more appropriate response (Bray, MacLean, & Hare, 2014). Inhibitory control is commonly measured using detour tests such as the AnotB test (Nawroth, von Borell, & Langbein, 2015; Britta Osthaus, Proops, Hocking, & Burden, 2013), reversal-learning tasks (Tapp et al., 2003; Bond, Kamil & Balda, 2007) and detour-reaching tasks (Kabadayi, Bobrowicz, & Osvath, 2018; Jan Langbein, 2018). Behavioural flexibility is also required to solve complex problems such as reversing a previously learned colour association or finding a novel solution to access a resource (Leal & Powell, 2012). The ability to solve problems is a pre-requisite for animals in the wild to survive, but it is also an important tool to have for captive animals. A greater ability to solve cognitive challenges can help animals to come up with novel ways to manipulate their environment and access food sources using strategies different from their natural foraging behaviour (e.g., tool use in cod: Millot et al., 2014). Thus, enhanced behavioural flexibility and problem-solving skills may facilitate adaptation to and coping with stressful situations, while their lack could cause frustration and stress in animals. Captive-bred animals generally have reduced behavioural diversity and less behavioural flexibility likely caused by barren non-stimulating rearing environments (Kihslinger, Lema, & Nevitt, 2006; E. Price, 1999; Salvanes & Braithwaite, 2005). While the effect of cognitive stimulation *per se* on behavioural flexibility, has not been investigated yet, several studies found positive effects of environmental enrichment on reversal learning and spatial abilities (Leggio et al., 2005; e.g., Schrijver, Pallier, Brown, & Würbel, 2004). For instance, in rodents, enriching the environment with objects such as Plexiglas tunnels, ladders, ceramic pots, toy balls, wooden planks, and running wheels rather than standard Plexiglas cages has been shown to have a positive effect on the ability to learn and remember new tasks (Ickes et al., 2000). It has to be mentioned however, that the positive effects might be caused by social rather than structural enrichment, as rats in the enriched group were housed in groups of 6 rats, while the non-enriched group was housed individually. In a study by Salvanes et al. (2013), rearing Atlantic salmon in structurally enriched tanks containing pebbles, cobbles and vertically floating plastic structures led to neural changes and improved learning ability assessed in a spatial task compared to salmon reared in plain hatchery tanks. In pigs, environmental enrichment provided in the form of bigger pens, a rooting area, and toys slightly improved the cognitive performance in a spatial hole-board task (Grimberg-Henrici, Vermaak, Elizabeth Bolhuis, Nordquist, & van der Staay, 2016). Although these studies found positive effects of enrichment on measures of behavioural flexibility and problem-solving, it is hardly possible to disentangle which effects were a result of cognitive stimulation *per se* and which were caused by other enrichment aspects such as engagement in social activities or the use of toys or rooting material. Thus, the question remains wether it is the “thinking” or the “doing” that is causing the positive effects on welfare.

Cognitive testing as a form of cognitive enrichment

Most research to date assessed the welfare effects of cognitive test protocols specifically developed with the aim to improve welfare parameters. In contrast, the main aim of animal cognition research is the assessment of cognitive capacities by confronting animals with cognitive tests. Thus, the question arises if cognitive testing for the purpose of scientific knowledge gain is also enriching for the test subjects.

Comparative cognitive research involving cognitive tests has been increasing over the last years and is targeting animals in research facilities, including laboratory, zoo, and farm animals (MacDonald & Ritvo, 2016). As cognitive tests are usually conducted with the aim to assess cognitive capacities of animals, each animal needs to be subjected to the same cognitive test under the same conditions. However, most studies that provided evidence for positive effects of cognitive stimulation on farm animals provided enrichment via automated enrichment devices that were incorporated in the housing environment (Katrin Ernst, Puppe, Schön, & Manteuffel, 2005; Jan Langbein et al., 2006; Jan Langbein, Siebert, Nuernberg, et al., 2007; Puppe et al., 2007; Zebunke et al., 2013). The automatisation of the training process by means of self-controlled learning is useful for the farm animal setting, e.g. to avoid additional stress due to separation from the group, or to increase control for the animal (Jan Langbein et al., 2006). Nevertheless, this process of self-controlled learning is usually not suitable for more controlled experimental settings where, e.g., a strict standardization of the test procedure is necessary or in case the tests require the presence of human experimenters (e.g., research on human-animal interactions). In addition, many cognitive tests require animals to be isolated (e.g., Ruby & Buchanan-Smith, 2015), handled, and/or given food reinforcement by a human (Morton et al., 2013; Nawroth, Brett, et al., 2016; Nawroth et al., 2014). Very little research has investigated the effects of stressors inherent to these tests such as isolation on test subjects (Ruby & Buchanan-Smith, 2015) or training by humans (Whitehouse et al., 2013). This is making it difficult to disentangle actual effects of cognitive stimulation from confounding factors such as the human contact or the habituation to new environments and isolation.

Over the course of cognitive tests, test subjects will become habituated to separation from peers, the interaction with humans or stimuli associated with the test environment. This habituation may help them to be less stressed in future stressful situations and thus perform better in cognitive tests. For instance, Hölscher (1999) compared handled vs. unhandled rats performance in a spatial water maze task and found that stress-induced learning deficits impaired performance in this task. However, pre-training of animals improved performance in learning spatial tasks, probably by reducing the amount of stress caused by the task and habituation to the test setup. Similarly, in pigs, Brajon et al. (2016) found that learning performance is linked to the level of fear towards a handler. Thus, previous experiences with cognitive tests might alter the motivation to participate and the performance in future cognitive tests by reducing neophobia and/or stress levels and a change in the perception of the animal. It remains to be determined if cognitive tests are enriching to test subjects and have similar positive effects on welfare as cognitive enrichment devices.

Hypothesis and aims

Although research in the field of animal cognition has increased immensely, the impact and consequences of cognitive tests on behaviour and welfare of test animals have not been investigated sufficiently. There are indicators that cognitive stimulation is a behavioural need of animals and that the exposure to cognitive challenges does increase the ability to cope with different stressors and to solve problems more flexibly. However, previous studies did not control for the many confounding factors attached to scientific tests, e.g., isolation during testing or human-animal interactions during tests.

The aim of the present thesis was to test if goats work for food in a cognitive test and to assess the effect of cognitive testing on measures of welfare in goats. More specifically, the objectives of this thesis were:

I. to assess whether goats show CFL in an instrumental task (Chapter II).

II. to investigate the effect of participation in long-term cognitive tests on stress reactivity towards isolation, novelty, and humans (Chapter III).

III. to assess the effect of participation in long-term cognitive tests on behavioural flexibility and problem-solving ability (Chapter IV).

IV. to assess the biological variance found in CFL, stress reactivity and behavioural flexibility caused by genetics, i.e., selection line (Chapter II-IV) and the environment, i.e., research site (Chapter III-IV).

Chapter II

Based on

Goats work for food in a contrafreeloading task.

Rosenberger K.[1,2], Simmler M.[3], Nawroth C.[4], Langbein J.[4], and Keil N.[1]

[1]Swiss Federal Veterinary Office, Centre for Proper Housing of Ruminants and Pigs, Agroscope, 8356 Ettenhausen, Switzerland

[2]Graduate School for Cellular and Biomedical Sciences, University of Bern, 3012 Bern, Switzerland

[3]Digital Production Group, Agroscope, 8356 Ettenhausen, Switzerland

[4]Leibniz-Institute for Farm Animal Biology, Institute of Behavioural Physiology, 18196 Dummerstorf, Germany

Published in Scientific Reports (2020), 10:22336, https://doi.org/10.1038/s41598-020-78931-w

---

Author contributions

K.R and C.N. collected the data. K.R. and M. S. analysed the data. K.R. drafted the manuscript, wrote the main parts of the manuscript, made all tables and Fig. 1 and Appendix I: Fig. S1. M.S. provided statistical support, adapted the IRT GLMM method to analyse and interpret the data, and made all corresponding statistical plots. C.N., J.L., and N.K. conceptualised the study and provided supervision. All authors interpreted the data and provided critical feedback on the manuscript.

Abstract

Contrafreeloading (CFL) is the phenomenon when animals work for a resource although an identical resource is available for free. Possible explanations for CFL are that animals seek context for species-specific behaviours or to control their environments. We investigated whether goats show CFL and whether breeding for productivity traits has altered its occurrence. In a manipulation task, we compared two selection lines: 27 Nigerian dwarf goats, not bred for productivity traits, and 30 dairy goats, bred for high milk yield. Over 10 trials, each goat could perform one of three behaviours: not participating in the trial, feeding for free from an open door, or opening a sliding door for a feed of similar value. The results were analysed using an Item Response Tree (IRTree) generalized linear mixed model (GLMM). The fitted probabilities to participate were >0.87 over all trials in both selection lines. For dwarf goats, the probability of choosing the closed door, and thereby demonstrating CFL, increased from 0.30 in Trial 1 to 0.53 in Trial 10. For dairy goats, this probability was constant at approximately 0.43. Unlike dwarf goats, dairy goats were faster to approach the closed compared to the open door. Overall, our results suggest that both selection lines were similarly interested in CFL.

Introduction

Contrafreeloading (CFL) describes the phenomenon when animals, given the choice, work for a resource even though an identical resource is simultaneously available for free (I. Inglis et al., 1997; G. D. Jensen, 1963; Osborne, 1977). This phenomenon contradicts optimal foraging theory (e.g. Stephens, D. W., Krebs, 1986) which suggests that an animal will maximize the net energy gain by choosing the food source providing most energy for the lowest cost (Pyke, Pulliam, Charnov, & Charnov, 1977). Although CFL is documented in captive wildlife and domestic pigs, cattle, goats, and chicken(de Jonge et al., 2008; Hessle et al., 2008; C. Lindqvist & Jensen, 2008; Van Os et al., 2018), it has not been reported in animals living in the wild (I. R. Inglis & Ferguson, 1986). One prominent theory to explain the occurrence of CFL is the Information Primacy Model (Ian R. Inglis, Langton, Forkman, & Lazarus, 2001). It assumes that CFL is driven by the urge to gather information about optimal food sources in a natural environment where food shortages can occur. As a result, if food deprivation increases, optimal foraging strategies will increase and the preference to gather information decreases. The need to explore the environment might, therefore, be an important adaptive mechanism (I. R. Inglis & Ferguson, 1986) and possibly explain why CFL occurs in animals in captivity (I. Inglis et al., 1997) where food is abundant.

Another theory to explain CFL is White's Competence Theory (White, 1959) which is not mutually exclusive from the Information Primacy Model. White's theory postulates the need of animals to control and modify their surroundings and assumes that the successful performance of a task reinforces itself (de Jonge et al., 2008; B. O. Hughes & Duncan, 1988; G. D. Jensen, 1963) by increasing the perceived control over the environment (Bassett & Buchanan-Smith, 2007; I. Inglis et al., 1997; Singh & Query, 1971). Several studies found that mastering a task can induce positive emotions in farm animals (Hagen & Broom, 2004; Jan Langbein et al., 2004; Manteuffel et al., 2009; Meehan & Mench, 2007; Puppe et al., 2007) that often live in barren environments with little stimulation or possibilities to control their surroundings. For example, in a study by De Jonge et al. (2008), pigs preferred searching for food rewards in straw rather than receiving the identical rewards freely available from a trough. The authors concluded that the display of CFL could be best explained by the rewarding effect of the anticipation of food while foraging. Therefore, if the need to control aspects of the environment and/or the need to perform species-specific behaviours are the motivators behind CFL, providing tasks to satisfy CFL motivation within a farm setting might enhance animal welfare. To effectively enrich housing conditions, all individuals of a species should frequently take part in the CFL task.

The type of training used, and the characteristics of the task offered to measure CFL may affect the proportion of individuals participating in the particular CFL experiment. Meagher et al. (2020)assessed the motivation to learn in cattle. They trained 30 heifers to perform an operant response (nose touch) to access a compartment providing a discrimination task. Although they used positive reinforcement (clicker training), the authors had to omit 10 animals mainly due to poor engagement in the task during the training phase. Participation proportions in the discrimination learning task varied between individuals and ranged from 0–100% of the offered sessions. Also, in studies where all of the individuals participated, huge individual differences in the extent of CFL display were found (e.g., Bean, Mason, & Bateson, 1999; de Jonge et al., 2008; Ogura, 2011; Sasson-Yenor & Powell, 2019). For example, in a study with starlings, the percentage of choosing to work for food ranged from 0–100% across individuals (Bean et al., 1999), whereas, in a study on pigs, the relative fraction of 'earned' food from total food consumed ranged from 0.4–12% (Young & Lawrence, 2003). Similar variations in CFL levels were reported in goats by Langbein et al. (Jan Langbein et al., 2009). In this study, 12 dwarf goats were trained to operate an automated learning device that posed a discrimination task to receive drinking water. They found that 10 out of 12 goats chose to direct, on average, one-third of their total daily button presses towards the device instead of pressing the normal drinker where no prior discrimination task had to be solved. Some goats gained more than 80% of their daily water at the learning device, while other goats gained very little water at the device.

It has been postulated that tasks that resemble the species' foraging behaviours or need low effort to receive the reward will promote CFL (de Jonge et al., 2008; I. R. Inglis & Ferguson, 1986). Goats are considered an intermediate ruminant type in the line of browsing - grazing species (Hofmann, 1989). They feed on a mixture of shrubs, trees, and grasses, often switching seasonally (Dicko & Sikena, 1992), and browse for the most nutritive fractions in their food (Lu, 1988; Morand-Fehr, Hervieu, & Sauvant, 1980). In contrast to sheep and cattle, which are mainly grazers, the time spent browsing can make up to 73–93% of their feeding time, depending on season (Dicko & Sikena, 1992). A suitable CFL task for goats might resemble this natural browsing behaviour, allowing for oral manipulation of the test apparatus. Additionally, a low-effort and easily executable manipulation task is expected to increase the number of individuals frequently choosing to work for the reward (I. Inglis et al., 1997).

Domestication in general, and the selection for high productivity in particular, were found to have altered not only stress reactivity (Campler et al., 2009; Colpoys et al., 2014; Dodd, Pitchford, Hocking Edwards, & Hazel, 2012), but also foraging (Hessle et al., 2008; C. E. S. Lindqvist, Schütz, & Jensen, 2002; C. Lindqvist & Jensen, 2008; Schütz & Jensen, 2001) and exploration behaviour (Colpoys et al., 2014) in farm animals. Compared to less productive breeds, pigs selected for high feed efficiency showed less behavioural reactivity towards fear-eliciting stimuli and displayed increased latencies when approaching a novel object and an unfamiliar human (Colpoys et al., 2014). Such alterations in behaviour may reflect in the motivation to show CFL. Schütz and Jensen (Schütz & Jensen, 2001) compared White leghorn chicken, selected for high egg productivity, to red jungle fowl, the ancestor of the domestic chicken, and to Swedish bantam, a domestic breed not strongly selected for production traits. They found that White leghorn chicken obtained a lower proportion of their food through CFL than both junglefowl and bantam chicken. Selection for high egg productivity in White leghorn chicken might thus have either directly reduced their motivation for CFL or indirectly decreased traits such as curiosity and risk-taking that are likely to affect the preferences of animals to perform CFL. Whether selection for high milk yield had a similar effect on the motivation to show CFL is yet unknown. To address this question, goats may represent a suitable species as selection lines differ strongly in milk production performance. Goats specifically selected for the dairy industry, such as Saanen goats, can produce up to 2–3 kg of milk per day (Vacca et al., 2018). On the other hand, common pet goats such as Nigerian Dwarf goats were not selected for productivity traits and their milk yield is much lower (Akinsoyinu, Mba, & Olubajo, 1977).

We examined whether domestic goats show CFL and repeatedly do so over several trials in a low-effort manipulation task that resembles their natural foraging behaviours. We provided goats with the choice between receiving a desired food item at an open door or opening a sliding door to access an identical food item. As some dwarf goats readily worked for more than 80% of their daily water intake in a previous CFL study (Jan Langbein et al., 2009), we expected that goats in our experiment would also show CFL to a certain proportion, i.e., push the closed sliding door open to receive the food reward instead of choosing the free reward. If goats are more motivated to work for a reward instead of receiving it for free, we would expect the approach time towards the closed door to be shorter than towards the open door (Rebecca K. Meagher et al., 2020). Following up on the findings by Schütz and Jensen (Schütz & Jensen, 2001), who showed lower motivation for CFL in the high-productivity chicken line, we assessed whether selection for high milk yield had a similar effect on CFL in goats. We compared dairy goats with a pedigree for high milk production to Nigerian dwarf goats that have not been selected for productivity traits. To increase genetic variability in our sample, we used individuals of two dairy breeds and of their crossbred (Saanen, Chamois Coloured, Saanen x Coloured). Using an Item Response Tree (IRTree) generalized linear mixed model (De Boeck & Partchev, 2012), the experiment was statistically modelled as a sequence of binary decisions between mutually exclusive behaviours (participate or not, choose open or closed door, approach fast or slow). Additionally, a linear mixed model, with approach time as a continuous response, was used to compare approach times between the open and the closed door.

Material, Animals and Methods

Location, animals, and housing conditions

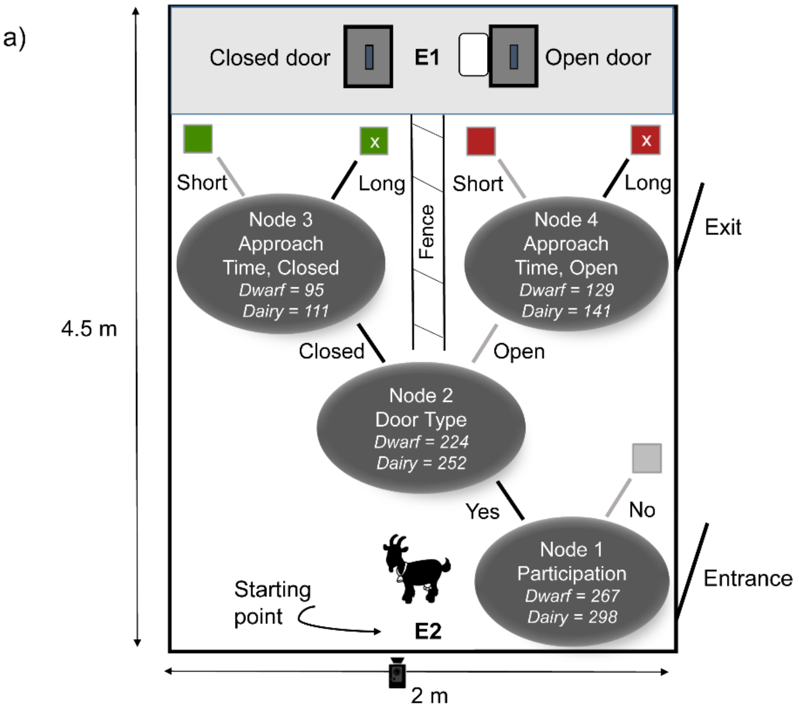
The study was carried out in August 2018 at the Agroscope Research Station in Ettenhausen, Switzerland. In total, we housed 60 domestic goats from two different selection lines: 30 Nigerian dwarf goats and 30 dairy goats. The Nigerian dwarf goat is commonly kept as pet and zoo animal in Europe and not selected for productivity traits. We used dwarf goats bred at the Leibnitz Institute for Farm Animal Biology in Dummerstorf, Germany. The only selection aim in this population was to avoid inbreeding, and the potential milk yield of dwarf goats does likely not exceed 0.3 kg per day (Akinsoyinu et al., 1977). All dwarf goats were born in January/February 2017. As it was common practice at the institute in Dummerstorf, dwarf goat kids stayed with their dams for six weeks before they were separated. They were moved to Ettenhausen in June 2017. To investigate the effect of selection for high productivity on CFL, we compared dwarf goats to dairy goats with a pedigree for high milk production. To increase genetic variability in our sample (Voelkl, Vogt, Sena, & Würbel, 2018), we used two of the most common high-producing dairy breeds in Switzerland and their crossbred, namely Saanen (n = 15), Chamois coloured (n = 12), and Saanen x Chamois (n = 3). These breeds have a potential milk yield of up to 3 kg per day (Vacca et al., 2018). All dairy goats were born on Swiss farms in February to April 2017. In accordance with common practice in the dairy goat industry, the dairy goat kids from Swiss farms were separated from their dam shortly after birth and artificially raised. They were moved to Ettenhausen in June/July 2017.

In Ettenhausen, dwarf and dairy goats were initially housed in one group pen each. At the age of 7–8 months, all goats were moved to pens of 10 goats each: three groups of dairy goats and three groups of dwarf goats. The total area of each dwarf goat pen was 14 m^2 (approximately 3.6 m x 3.9 m), consisting of a deep-bedded straw area of 11 m^2 (approximately 2.8 m x 3.9 m) and a 0.5 m elevated feeding place (3.0m^2). The total area of each dairy goat pen was 17.5 m^2 (approximately 3.9 m x 4.5 m), consisting of a deep-bedded straw area of 13.2 m^2 (approximately 4.5 m x 2.9 m) and a 0.6 m elevated feeding place (4.3 m^2). Hay was provided *ad libitum* behind a feeding fence and replenished twice a day at approximately 8 am and 4 pm. Each pen had one drinker and a mineral supply. Additional structures in the straw-bedded area included a wooden bench (dairy: 2.4m long, 0.6m high, 0.6m wide; dwarf: 2.3m long, 0.5m high, 0.5m wide) along the wall of the pen and a round wooden table (dairy: 0.8 m high, Ø 1.1 m; dwarf: 0.6 m high, Ø 1 m) in the centre of the pen. The goats were between 15 and 18 months old at the start of the study (mean ± SD, dairy goats: 529 ± 18.7 d, dwarf goats: 578 ± 4.7 d).

All procedures involving animal handling and treatment were approved by the Swiss Cantonal Veterinary Office Thurgau (Approval No. TG04/17 – 29343) and were performed in accordance with all relevant Swiss legislative and regulatory requirements and the ASAB/ABS guidelines for the use of animals in research (ASAB & ABS, 2018).

Test apparatus and test arena

The goats were individually tested in a test arena (4.5 m x 2 m). The test apparatus was installed as part of a wall on the shorter side of the test arena and consisted of two identical wooden sliding doors. A metal grid fence (1.35 m high x 1.27 m long) was installed between the two sliding doors to prevent goats from switching sides and feeding from both openings right after each other (Appendix I, Fig. S1). A human experimenter (E1) was positioned behind the wall of the test apparatus. Whenever a goat opened the closed door (= CFL) or stuck its muzzle through the open door, it instantly received a piece of uncooked pasta in a plastic dish from E1 as a reward.



b)

Behavioural categories		Symbol	Node 1	Node 2	Node 3	Node 4
No participation		⬜	0	Na	Na	Na
Closed door	Short approach time	🟩	1	1	1	Na
	Long approach time	🟩 x	1	1	0	Na
Open door	Short approach time	🟥	1	0	Na	1
	Long approach time	🟥 x	1	0	Na	0

Figure 1: a) Schematic drawing of the test arena with left door closed and right door open including the binary response tree with four nodes representing the sequential choices leading to one of five behavioural categories (= five square symbols). A number of observations with non-missing information at respective nodes are given in the tree. . Positions of the experimenters are indicated with E1 and E2, and the position of the video camera is marked with a camera symbol. b) The five behavioural categories with their symbols corresponding to the tree in a) as well as the encoding of the node for the IRTree model.

Habituation phase

The goats were familiar with the arena and with opening a sliding door from a previous experiment on social learning, and additional training in case it was necessary. In the previous experiment the apparatus consisted of one sliding door only. All goats had gone through a 3-day habituation phase to get used to feed from the open door: on day 1, goats had been habituated in pairs for 3 min and 10 pieces of pasta per pair of goats. On day 2 and 3, goats had been habituated individually over 3 min and 10 pieces of pasta per goat. In the subsequent test sessions, goats first observed a human demonstrator opening the closed sliding door and were then given the opportunity to open the door themselves. All goats had received a total of 10 sessions over 10 consecutive days. At the end of the experiment, all but four dwarf goats had learned to open the closed sliding door.

In the current experiment, we presented two identical sliding doors simultaneously. We provided no additional habituation to this novel setup as the previous experiment finished only two days prior to the start of the current experiment. The four dwarf goats that had not learnt to open the sliding door in the previous experiment received additional individual training on the day before the current experiment by leaving a slowly shrinking gap to facilitate door opening until the sliding door was fully closed. However, only one of the four goats successfully learned to open the door and, thus, was included in the study. The other three were excluded from the experiment. Consequently, we included 57 goats in our study, 30 dairy and 27 dwarf goats.

Test procedure

Each goat received 10 consecutive trials on a single day with each trial lasting 30 seconds. In total, goats were tested within four days. In each trial, one of the two sliding doors was kept open, allowing free access to a food reward. The other door was presented closed, requiring manipulation to slide the door to the right or left to access the food, i.e., necessitating work for the reward (Fig. 1a, Appendix I Fig. S1). The closed door was administered on the left and the right sides of the fence in a pseudorandom order, but each goat was constrained to a total of five trials with the closed door on the left and five trials with the closed door on the right side.

Goats were individually led into the test arena by a second experimenter (E2) and released near the centre of the room approximately one meter away from the start of the fence (= starting point, Fig. 1a). In each trial, the goat could choose to walk to the closed or the open door. After each trial, the individual was led back to the starting point by E2, and the next trial started. All trials were videotaped with a camcorder (Sony HDR-CX240E) mounted on top of the wall of the arena above the starting point (Fig. 1a). Due to technical failures, five trials were not videotaped and therefore excluded from the analysis. The videos were analysed with the Observer XT software (Version 12, Noldus Information Technology, The Netherlands). We recorded whether the goat participated or not, and, if it participated, which door type it chose (closed = CFL or open = no CFL), as well as the time it took to approach the door (= nose within less than approximately 5 cm to the door) from the start of the fence (approach time, in sec). Participation was defined as walking towards one of the doors, opening the door, if closed, and feeding through the door from the plastic dish. If a goat did not participate within 30 seconds after it was released, the trial was recorded as ‘no-participation’, and the goat was led back to the starting point to begin the next trial. Opening the closed door without feeding from the dish was never observed. To assess the reliability of the approach times determined from videos, we compared them to times recorded during the experiment with a stopwatch (Pearson correlation coefficient, r_p = 0.85). Participation and the choice of door type were unambiguous.

Data analysis and statistics

All statistical analyses were performed in R v4.0.2 (R Core Team, 2014). We employed an Item Response Tree (IRTree) model of the GLMM family to analyse the multivariate behavioural response in our experiment. Such IRTree models allow representing a multivariate behavioural response as a tree of sequential binary responses, enable incorporating hierarchical sampling, and can account for correlated responses as well as repeated testing of the same individuals (De Boeck & Partchev, 2012). They are well suited for analysing categorical data in behavioural studies, as shown for example for escalating courtship behaviours, antipredator behaviours, and social interactions (De Bona, Valkonen, López-Sepulcre, & Mappes, 2015; López-Sepulcre, De Bona, Valkonen, Umbers, & Mappes, 2015).

For encoding the data as a binary tree and specification of the model, we followed recommendations by López-Sepulcre et al. (2015). Figure 1a shows the binary response tree of our IRTree model. The nodes correspond to the sequential choices (responses) leading to one of five behavioural categories (= five symbols). Node 1 encodes the participation in a given trial (‘yes’, 1; ‘no’, 0). Node 2 encodes the choice of either walking to feed at the open (0) or the closed (1) door. Finally, Nodes 3 and 4 encode the time taken to walk from the start of the fence to the closed or the open door, respectively. This approach time was short (0) or long (1) corresponding to below or above the median over all observations of the respective selection line (dwarf = 2.66 s, dairy = 1.81 s; Appendix I: Fig. S2). Goats that chose not to participate in a trial made only one choice while participating goats made three sequential choices, as Node 3 and 4 are mutually exclusive (see Fig. 1a, b). The IRTree model was estimated as GLMM with a binary response and logit link using the *glmer* function from the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015). The model formula in *lme4* syntax was as follows:

Value ~ 0 + Node:SelectionLine + Node:SelectionLine:I(Trial - 1) + (1 | Obs) + (0 + Node | Pen/Individual/ClosedSide)

The nodes were qualitatively different from each other, and we suspected the selection lines to differ in their behaviour as well as to adapt their responses with repeated trial. As fixed effects, we therefore, included for each node an individual intercept for the two selection lines (0 + Node:SelectionLine) and an individual slope for the trial number for the two selection lines [Node:SelectionLine:I(Trial - 1)]. The trial number was included as Trial - 1 to render the intercept to correspond to Trial 1 instead of the non-meaningful Trial 0. As a random effect, we included a random intercept for observation (1|Obs) to ensure that the sequential binary responses corresponding to a single observation shared the same variance and were not treated as independent observations. Furthermore, we specified for each node a random intercept for the location of the closed door, nested within individual, itself nested within pen (0 + Node | Pen/Individual/ClosedSide). This was done to account for potential side bias, repeated testing of the same individual, and potential effects of pen affiliation. Despite the randomization of the location of open and closed door, a side bias was evident; the choice for the closed door was apparently more likely when the closed door was on the left versus the right side (Wald test in repeated measure logistic regression, p = 0.03).

To investigate the difference in behaviour between the selection lines, we tested selection line contrasts for the fixed effects using the *glht* function from the R package multcomp (Hothorn, Bretz, Westfall, & Heiberger, 2008). Both the p-values for fixed effect estimates (*glmer* function) and for the contrasts (*glht* function) were obtained using Wald tests. Fitted probabilities and bootstrap confidence bands that were only conditioned on the fixed effects were obtained using the *predict.MerMod* function (parameter re.form = ~ 0; lme4 package) in conjunction with the *bootMe*r function (lme4) for parametric bootstrapping (10,000 bootstraps).

To compare approach times towards the open versus the closed door, we also analysed the approach time as a continuous response in a linear mixed model using the *lmer* function (lme4 package) and the following model formula:

log2(ApproachTime) ~ 0 + SelectionLine + SelectionLine:DoorType +
SelectionLine:DoorType:I(Trial - 1) + (1 | Pen/Individual/ClosedSide)

The approach time was right-skewed and, therefore, log2-transformed to approximate normal distribution (Appendix I: Fig. S2). Individually for the two selection lines, we included an intercept (0 + SelectionLine), an effect of the choice of door type (SelectionLine:DoorType), and a slope for trial separately for the closed and open doors [SelectionLine:DoorType:I(Trial-1)]. Besides these fixed effects, a random intercept for the location of the closed door, nested within individual, itself nested within pen (1 | Pen/Individual/ClosedDoor) was included to account for potential side bias, repeated testing of the same individual, and potential effect of pen affiliation. P-values for the fixed effect estimates were obtained using Z-tests through the *glht* function (multcomp package).

Results

We found that all but four goats chose to participate in at least half of the 10 trials of the experimental task. Three individuals participated in less than 5 of 10 trials (B2, Y1, Z8) and only one individual chose in all trials to not participate (X4, Fig. 2). In total, 53 of 57 goats chose to feed from the closed door in at least 1 of 10 trials, but inter-individual variation was substantial with values ranging from 1 to 7 of the 10 trials (Fig. 2).

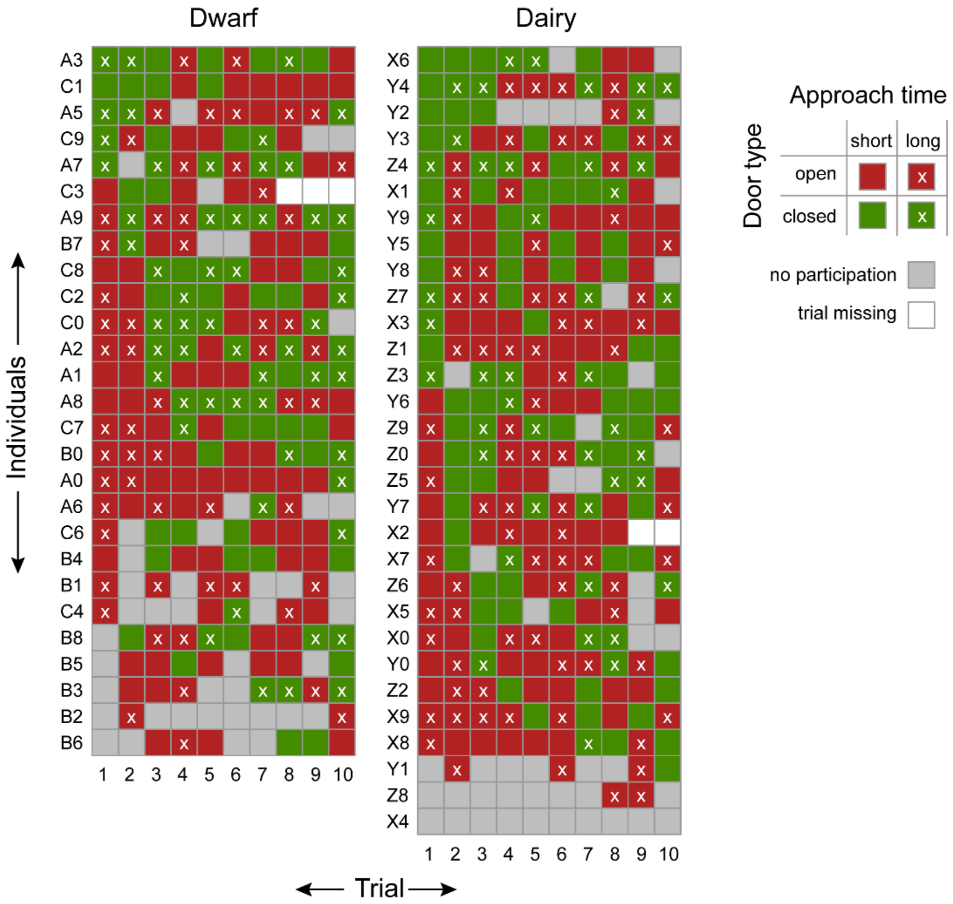


Figure 2: Individual distribution of choices regarding door type and approach time from trials of dwarf and dairy goats. Names of individuals are composed of a letter for the pen affiliation (dwarf: A, B, C; dairy: X, Y, Z) and a number for the individual within each pen (0-9).

Table 1: IRTree GLMM of behavioural responses. Intercepts correspond to Trial 1 (see Methods). Results with p-value ≤ 0.05 are given in bold. An extended version of this table, including standard error and Z-values, is provided in Appendix I: Table S1. Additionally, Appendix I, Table S2 lists the random effect variance components and correlations.

Fixed Effects	**Node 1: Participation** *p* participation				**Node 2: Door type** *p* closed door				**Node 3: Approach time closed** *p* long approach time				**Node 4: Approach time open** *p* long approach time			
	Dwarf		**Dairy**		**Dwarf**		**Dairy**		**Dwarf**		**Dairy**		**Dwarf**		**Dairy**	
	est.	**p**	**est.**	**p**	**est.**	**p**	**est.**	**p**	**est.**	**p**	**est.**	**p**	**est.**	**p**	**est.**	**p**
Intercept	**2.43**	**<0.001**	**3.32**	**<0.001**	**-0.85**	**0.003**	-0.27	0.29	0.25	0.75	-0.72	0.27	0.55	0.25	**0.99**	**0.04**
(Trial – 1)	0.03	0.69	**-0.15**	**0.04**	**0.11**	**0.03**	0.002	0.97	0.07	0.48	0.04	0.57	**-0.21**	**0.01**	-0.07	0.34
	Dwarf - Dairy				**Dwarf - Dairy**				**Dwarf - Dairy**				**Dwarf - Dairy**			
Contrast in…	**est.**		**p**		**est.**		**p**		**est.**		**p**		**est.**		**p**	
Intercept	0.90		0.32		0.58		0.13		-0.98		0.33		0.45		0.50	
(Trial – 1)	-0.18		0.08		-0.11		0.12		-0.03		0.84		0.15		0.17	

The results of the IRTree GLMM are presented in Table 1 (and in Appendix I: Table S1 and S2). Fitted probabilities of behavioural choices represented by the nodes as well as observed proportions are shown in Figure 3. The fitted probability to participate in the task was > 0.87 over all trials for both selection lines (Fig. 3 – Node 1). In dairy goats, but not in dwarf goats, the probability of participating decreased as the trial numbers increased — from 0.97 in Trial 1 to 0.88 in Trial 10 (p = 0.04, Table 1, Fig. 3 – Node 1).

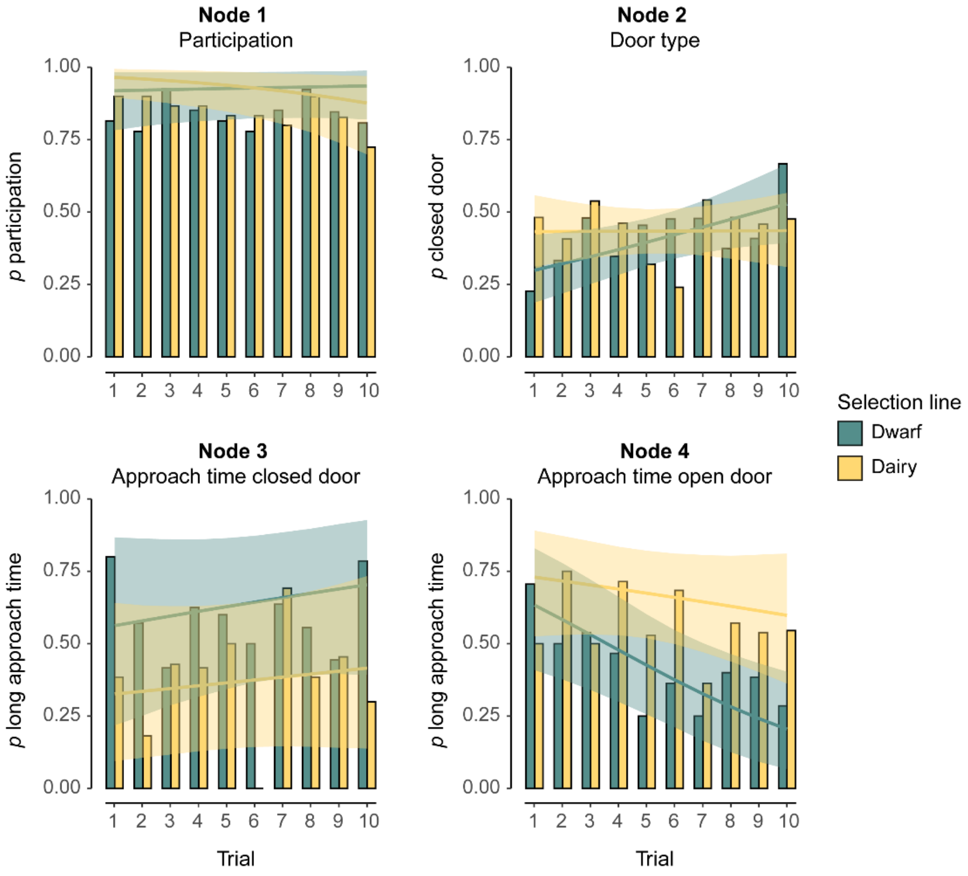


Figure 3: Fitted probabilities (lines) of the IRTree GLMM at the four nodes and observed proportions (bars). The shaded areas represent 95% confidence bands for the fitted values considering the fixed effect uncertainty.

In dwarf goats, the probability of choosing the closed door was 0.30 in Trial 1, but it increased with increasing trial number (p = 0.03) and reached a probability of 0.53 in Trial 10 (Fig. 3 – Node 2). In dairy goats, the probability of choosing the closed door was approximately constant at around 0.43 throughout all trials (Fig. 3 – Node 2).

The probability for a long approach time (= above the median over all trials of the respective selection line) towards the closed door ranged from 0.56 to 0.70 in dwarf goats and from 0.33 to 0.42 in dairy goats (Fig. 3 – Node 3). As the uncertainty in these probabilities was high (Fig. 3 – Node 3, Table 1), a difference between selection lines is statistically not supported (p = 0.33 for intercept contrast in Trial 1).

In dwarf goats, the probability for a long approach time towards the open door decreased with increasing trial number (p = 0.01) — from 0.63 in Trial 1 to 0.20 in Trial 10 (Fig. 3 – Node 4). In dairy goats, the probability for a long approach time towards the open door was 0.73 in Trial 1 and 0.60 in Trial 10 (Fig. 3 – Node 4), with a high uncertainty in these probabilities throughout the various trials (p = 0.34).

Figure 4 shows the probabilities of the five behavioural categories, representing the possible outcomes of the sequential choices. In dwarf goats, the probability for 'open door, long approach time' decreased with increasing trial number in favour of the probabilities for 'open door, short approach time' and for 'closed door, long approach time'. In dairy goats, the probability for 'open door, long approach time' decreased in favour of the probability for 'no participation'.

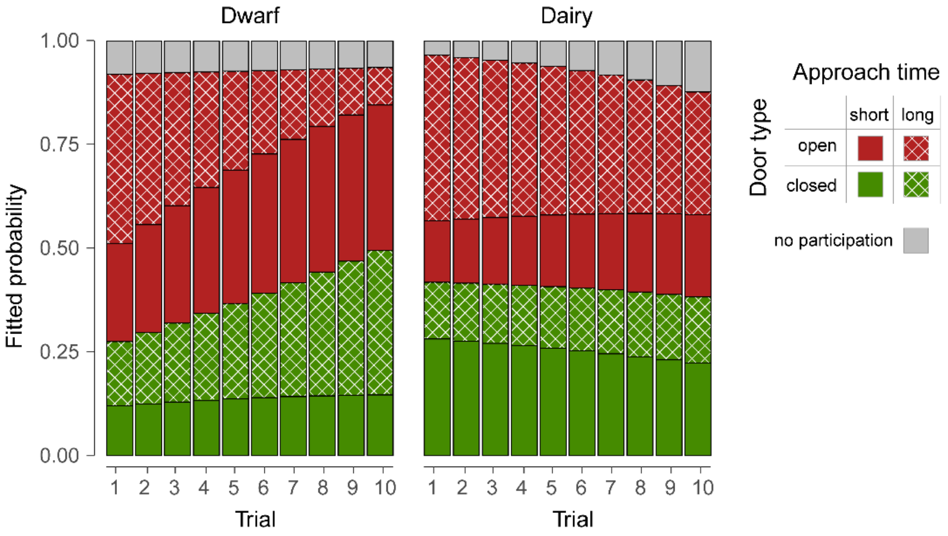


Figure 4: Fitted probabilities of the IRTree GLMM for the five behavioural categories. These probabilities were calculated by multiplying the probabilities of the corresponding sequential choices (Fig. 1a, b). For example, the probability for the behavioural category 'open door, long approach time' (Fig. 1a, b and Fig. 3) was calculated as the probability to participate (Node 1), times the probability to choose the open door (Node 2), times the probability to show a long approach time (Node 4).

Table 2: Estimates of variance components and fixed effects of the LMM with continuous approach time as response. Intercepts correspond to Trial 1 (see Methods). Results with p-value ≤ 0.05 are given in bold. Appendix I: Table S3 lists the random effect variance components.

Fixed Effects	**Dwarf**				**Dairy**			
	est.	**s.e.**	**z**	**P**	**est.**	**s.e.**	**z**	**p**
Intercept	1.73	0.17	10.05	<0.001	1.21	0.17	7.05	<0.001
Door Type Closed	0.12	0.20	0.60	0.55	-0.54	0.18	-3.04	0.002
(Trial – 1): Door Open	-0.07	0.02	-3.14	0.002	-0.02	0.02	-1.01	0.31
(Trial – 1): Door Closed	-0.01	0.03	-0.34	0.73	0.02	0.02	0.81	0.42

The results of the linear mixed model, with approach time as continuous response, are presented in Table 2 (and Appendix I: Table S3). Fitted approach times are shown in Figure 5. Unlike the IRTree model above, this model allows a direct comparison of approach times towards the open versus the closed door. The fitted approach time for the dwarf goats towards the closed door was approximately constant at around 3.5 s (Fig. 5). Towards the open door, the fitted approach time was 3.3 s in Trial 1, decreased over time (p = 0.002), and was 2.1 s in Trial 10 (Fig. 5).

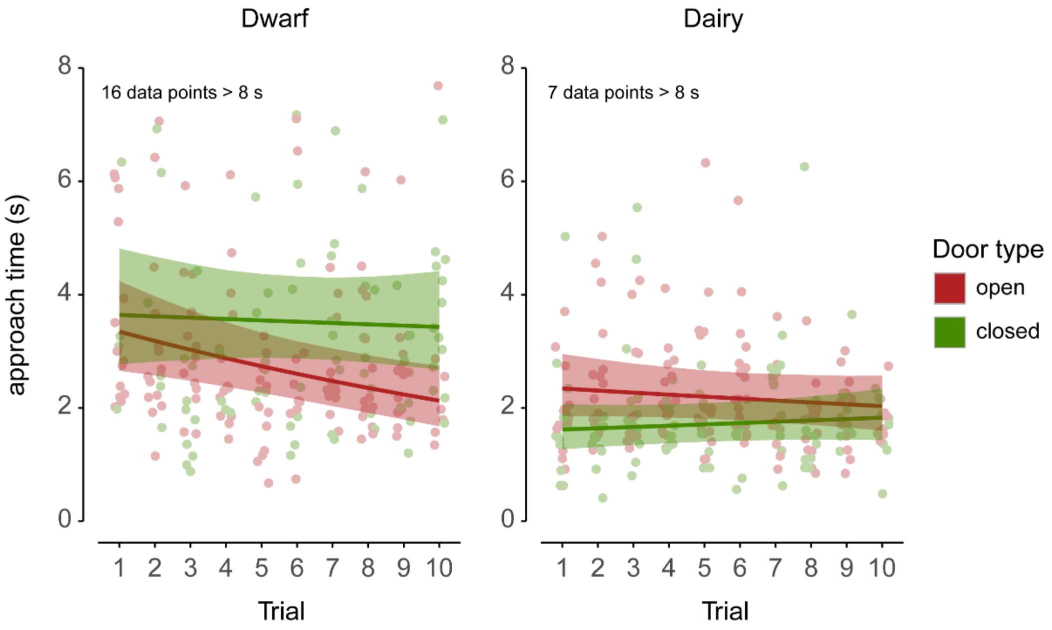


Figure 5: Fitted approach times of the linear mixed model (back-transformed from log2 to linear scale). The shaded areas represent 95% confidence bands for the fitted values considering the fixed effect uncertainty. Dots represent observed approach times. For each trial number, the dots were horizontally jittered for visual clarity. Data points > 8 s are not shown (numbers given in panels).

For dairy goats, the fitted approach time towards the closed and the open door was approximately constant over the trials and ranged from 1.6 (Trial 1) to 1.8 s (Trial 10) and 2.3 (Trial 1) to 2.0 (Trial 10), respectively. Thus, dairy goats approached the closed door faster than the open door ($p = 0.002$).

Discussion

In line with our hypothesis, goats chose to participate in the experiment and were willing to work for a reward in the presence of an identical, free reward, and thus chose to perform CFL. Contrary to our expectations, the selection line was not related to the overall probability of working for food but rather to changes in goats' responses with increasing trial numbers. Our results suggest that both selection lines were motivated to work for food, while this motivation was suppressed in early trials in dwarf goats, presumably due to higher stress reactivity.

The high proportion of goats that frequently chose to participate in the experiment suggests that we used a suitable experimental setting with a highly desired reward and enough previous habituation to the test arena and the sliding door in order to make participation rewarding for goats. To avoid spatial learning, we randomized the location of the open and the closed door in all trials. However, we found indications that the closed door was more likely to be chosen when located on the left side. As the goats' pen mates where located on the right side, this bias seems not socially induced. In our statistical models we considered the potential side bias as a random effect nested in individual, thus allowing an individual side bias for each goat.

In accordance with our expectations, all but four goats exhibited CFL in at least 1 of 10 trials. Overall, the probabilities to choose the closed door were not different between the dwarf and dairy goats. This is in contrast to what we expected and to previous studies on different selection lines of chicken (C. E. S. Lindqvist et al., 2002; C. Lindqvist & Jensen, 2008; Schütz & Jensen, 2001) and cattle (Hessle et al., 2008) which found that animals selected for high productivity were choosing more energy conserving strategies and thus were less likely to show CFL than animals not selected for production traits. In our experiment, the probability in dwarf goats to choose the closed door increased with increasing trial number, while it remained approximately constant for dairy goats (Fig. 3 – Node 2). As all goats were familiar with receiving food out of both doors, open and closed, the initial reluctance to choose CFL in dwarf goats is unlikely to be explained by neophobia towards one of the doors (Mitchell, Scott, & Williams, 1973). Although all goats had been habituated to the single sliding door in a previous experiment, no additional habituation to the novel setting with two sliding doors was performed in the current experiment. This new setting may have induced more stress in dwarf goats than in dairy goats and resulted in dwarf goats initially choosing the option that appears to be less risky (i.e. the open door). Increasing habituation and positive reinforcement from opening the door might then have increased motivation to choose CFL in subsequent trials. This is in line with the notion that stress reactivity has been reduced in animals selected for high productivity (C. Lindqvist & Jensen, 2009; Schütz & Jensen, 2001), which would suggest reduced stress reactivity in dairy goats compared to dwarf goats.

However, not only a genetic disposition but also differences in rearing may have caused different stress responses between the selection lines in our experiment. Whereas the dairy goats were artificially raised without their mothers, the dwarf goats stayed with their dams for six weeks. Previous studies (Lyons, 1989; Lyons, Price, & Moberg, 1988) found indications for higher fearfulness of dam-reared goats in goat-human encounters as compared to human-reared goats — dam-reared goats exhibited greater behavioural responsiveness in novel situations, as well as longer latencies to approach an unfamiliar human.

Regarding the effect of door type on approach times, we hypothesised that, if goats are motivated to work for a reward instead of receiving it for free, they would approach the closed door faster than the open door (Rebecca K. Meagher et al., 2020). Our results only support this hypothesis in dairy goats, which approached the closed door faster. In contrast to dairy goats and not in line with our hypothesis, dwarf goats showed similar approach times towards both doors in the first few trials and tended to approach the open door faster in later trials (Fig. 5). An explanation for these observations may again be differences in stress reactivity. Maybe dwarf goats would have required more time to adapt to the test situation to react similarly to dairy goats regarding their approach time towards the closed and open door. Recent research on farm animal personality highlights the need to consider the animals' individual stress levels for their habituation to experimental tasks (Finkemeier, Langbein, & Puppe, 2018), regardless of whether it is genetically based or developed during the ontogenesis.

Over all trials, the probability of choosing CFL was slightly below 50%, raising the question whether our results could partially be explained by goats randomly choosing the open or closed doors. However, we found the probability of choosing CFL in dwarf goats to increase over trials and the approach time to be affected by door type. This indicates that goats deliberately chose the door type rather than choosing at random. An explanation for the occurrence of CFL over several trials might be attributed to intrinsic rewarding properties of the performance of the task itself (de Jonge et al., 2008; B. O. Hughes & Duncan, 1988; G. D. Jensen, 1963). Positive emotions as a result of mastering a task and being in control of the situation have been reported in cattle, pigs, and goats (Hagen & Broom, 2004; Jan Langbein et al., 2004; Manteuffel et al., 2009; Meehan & Mench, 2007; Puppe et al., 2007). This is also in line with White's Competence theory (White, 1959), which postulates that animals are motivated to manipulate and control their environment to attain competence. Hence, it is possible that goats were choosing CFL due to positive feedback from executing the manipulation task.

It has been suggested that animals need appropriate cognitive challenges to express control over their environment (Hagen & Broom, 2004; Jan Langbein et al., 2004; Manteuffel et al., 2009; Meehan & Mench, 2007; Puppe et al., 2007) and that animals value effort (Franks, 2019) such that the incorporation of such challenges into a farm setting can have welfare benefits for the animals (K. Ernst, Tuchscherer, Kanitz, Puppe, & Manteuffel, 2006; Jan Langbein et al., 2004; Puppe et al., 2007). For a successful implementation of CFL tasks in the husbandry system, it would be necessary to evaluate if the motivation to display CFL is stable in various conditions and also persists over a longer period of time. The high inter- and intra-individual variations in the extent of CFL, which are in accordance with other studies on CFL in dwarf goats (Jan Langbein et al., 2009) and in other species (Bean et al., 1999; McGowan, Robbins, Alldredge, & Newberry, 2010; Sasson-Yenor & Powell, 2019), require further research regarding the motivational background to show CFL in animals.

Conclusion

Overall, high CFL proportions in both selection lines, increasing interest in approaching the closed door in dwarf goats, and shorter approach times towards the closed door compared to the open door in dairy goats indicate that both dairy and dwarf goats were motivated to work for a resource in the presence of the same resource for free. The two selection lines of goats differed in the changes of the probabilities to choose CFL with increasing trial number and regarding the comparison of approach times towards the open versus towards to closed door. These results might reflect differences in stress reactivity towards the CFL task, potentially related to selection for productivity or differences during ontogeny. Our findings suggest that goats seem to be motivated to solve a CFL task, stressing the need for the provision of cognitive challenges to improve the welfare of farm animals.

Acknowledgements

We thank all the staff of the Agroscope research station in Ettenhausen who helped in this experiment and who cared for the animals. We additionally want to thank Michael Seehaus for coding our videos. This work was supported by a SNF grant (310030E-170537) to Nina Keil and a DFG grant (LA 1187/6-1) to Jan Langbein.

Chapter III

Based on

Reactivity of domesticated goats towards different stressors following long-term cognitive test exposure

ROSENBERGER, K.[1,2], SIMMLER, M. [3], LANGBEIN, J.[4], NAWROTH, C.[4], KEIL, N.[1]

[1]Swiss Federal Veterinary Office, Centre for Proper Housing of Ruminants and Pigs, Agroscope, 8355 Ettenhausen, Switzerland

[2]Graduate School for Cellular and Biomedical Sciences, University of Bern, 3012 Bern, Switzerland

[3]Digital Production, Agroscope, 8355 Ettenhausen, Switzerland

[4]Leibniz Institute for Farm Animal Biology, Institute of Behavioural Physiology, 18196 Dummerstorf, Germany

Manuscript in preparation for PLOS ONE (June/July 2021)

Author contributions

K.R. and C.N. collected the data. K.R. and M.S. analysed the data. K.R. drafted the manuscript, wrote the main parts of the manuscript, and made tables and figures (except of Fig. 8 and 9 with help from M.S.). M.S. provided statistical support and made all corresponding statistical plots (Fig. 8 and 9 together with K.R.). N.K., C.N., and J.L. conceptualised the study and provided supervision. All authors interpreted the data and provided critical feedback on the manuscript.

Abstract

Current evidence suggests that frequent cognitive stimulation of captive animals may have positive effects on stress reactivity. In this study we used standardised cognitive tests (object choice tests) to assess the effect of long-term cognitive stimulation on behavioural and cardiac responses of domestic goats in subsequent stress tests. To disentangled whether cognitive stimulation *per se* or the reward associated human-animal interaction required for testing was affecting the stress reactivity we conditioned three treatment groups: goats that were isolated for participation in human-presented object-choice tests and rewarded with food (COG treatment), goats isolated as for the test exposure and rewarded with food by the experimenter without being administered the object-choice tests (POS treatment), and goats that were isolated in the same test room but received neither a reward nor were administered the tests (ISO treatment). All treatment groups were subsequently tested in four stress tests: a novel arena test, a novel object test, a novel human test, and a weighing test, in which goats had to enter and exit a scales. To increase variability of our sample we tested two selection lines, dwarf goats, not selected for production traits, and dairy goats selected for high productivity, each at the same two research sites. Analysing the data with principal component analysis and linear mixed-effects models, we did not find evidence that standardized cognitive testing *per se* (COG–POS contrast) reduces stress reactivity of goats in subsequent stress tests. However, for dwarf goats but not for dairy goats we found support for an effect of reward-associated human-animal interactions (POS-ISO contrast) at least for some stress test measures. This highlights the need to consider phenotypic variation when assessing stress reactivity and when interacting with goats.

Introduction

Current evidence suggests that frequent cognitive stimulation of captive animals may have positive effects on stress reactivity. Many animals in zoos, laboratories or research facilities are frequently confronted with cognitive tests for a scientific aim, but little research has investigated the effects of these tests on the welfare of test subjects. First results suggest that cognitive stimulation via enrichment devices has positive effects on activity budgets and social interactions in primates (Jacobson, Kwiatt, Ross, & Cronin, 2019b; Whitehouse et al., 2013; Yamanashi & Hayashi, 2011) and the potential to increase exploration and reduce fear in stressful situations in pigs (Zebunke et al., 2013). In contrast, research with computer-controlled learning devices suggested that cognitive challenge may also cause stress and frustration if training events are unpredictable (Doyle et al., 2011; Galhardo, Vital, & Oliveira, 2011) or during initial stages of learning when failure often occurs (Jan Langbein et al., 2004; Yamanashi, 2010). Whether standardised cognitive testing itself affects stress reactivity has not been investigated.

There are several aspects of cognitive stimulation tasks that may produce positive effects on welfare. Results of studies in farm animals assessing the potential of cognitive stimulation as a form of enrichment indicate that engagement in a solvable task induces positive emotions (Hagen & Broom, 2004; Jan Langbein et al., 2004; Manteuffel et al., 2009; Meehan & Mench, 2007; Puppe et al., 2007). It is argued that the successful execution of a task is not only reinforcing in itself (B. O. Hughes & Duncan, 1988; G. D. Jensen, 1963), but also increasing the perceived control over the environment (Jan Langbein et al., 2004; Meehan & Mench, 2007). Zebunke et al. (2013) used a call-feeding station incorporated in the home pen as cognitive enrichment for pigs and found that the animals were less stressed by isolation and more explorative towards a novel object than pigs without this cognitive enrichment. They concluded that the introduction of cognitive stimulation in the housing environment has the potential to reduce stress reactivity in future situations (Puppe et al., 2007; Zebunke et al., 2013).

Captive animals are frequently exposed to stressors related to housing and husbandry routines that include isolation and handling by humans and the confrontation with novel stimuli to various extents (e.g. during health checks, temporal separation or relocation, and cleaning of facilities). Several stress tests have been applied in animals to study the behavioural responses towards different stressors (Forkman et al., 2007). The so called ‘open-field’ or ‘novel arena test’ has been used to assess behavioural reactions towards isolation and a novel environment in many species (Graunke, Nürnberg, Repsilber, Puppe, & Langbein, 2013b; Neave et al., 2018; Oesterwind et al., 2016; Prut & Belzung, 2003). The ‘novel object test’ has been applied to assess behavioural reactivity towards a novel stimuli and to investigate the motivation to explore (Finkemeier et al., 2018; Sneddon et al., 2003). The reactivity towards a human has been measured using different animal-human-encounter tests with a standing or walking human (Lyons et al., 1988) and with a sitting human (Romeyer & Bouissou, 1992). In addition to reactivity assessment using established stress tests, stress responses can also be assessed in routine handling procedures such as weighing. The weighing situation includes several potentially stressful components such as human-animal interaction, novelty, confinement, and separation from conspecifics (Forkman et al., 2007; Hemsworth, 2003).

Stress responses towards the same stimuli can vary greatly between individuals. This inter-individual variation is caused by an interplay of environmental and genetic factors (Dantzer and Mormède 1983). Breeding for specific traits can intentionally or unintentionally lead to changes in behavioural and physiological stress responses of the selection line (Rauw et al., 1998). Selection for high productivity has been shown to reduce stress reactivity during human-animal interactions (Campler et al., 2009; Colpoys et al., 2014; C. Lindqvist & Jensen, 2008; Schütz & Jensen, 2001), and to affect the reactivity towards isolation (Kilgour & Szantar-Coddington, 1995; Romeyer & Bouissou, 1992). These differences could be relevant when exposing animals to handling procedures or novel environments. In order to achieve greater general validity of research results, it is therefore advantageous to consider different selection lines of a species. Goats are a suitable model species to study the effect of selection for productivity on stress reactivity because not all selection lines were bred with the aim to increase productivity. Dwarf goats for instance evolved under extensive husbandry conditions in West Africa and are commonly not bred for productivity. They are usually kept as pets in zoos or by hobby breeders in Europe. In contrast, dairy goats have been and are still selected for high milk yield and ease of handling during milking. Goats are also suitable to study the effects of cognitive testing on stress reactivity because common stressors such as isolation from the group (Aschwanden, Gygax, Wechsler, & Keil, 2008b; Carbonaro, Friend, Dellmeier, & Nuti, 1992; E. O. Price & Thos, 1980; Siebert et al., 2011), novel environments and objects (Forkman et al., 2007) as well as the presence of a human and handling procedures (Forkman et al., 2007; Lyons, 1989; Lyons et al., 1988) have already been investigated. Additionally, a lot of literature exists on the highly developed cognitive capacities of goats (J. Langbein et al., 2008; Nawroth, 2017) and their motivation to engage in cognitive tasks has been demonstrated in previous studies (Jan Langbein et al., 2009; Rosenberger, Simmler, Nawroth, Langbein, & Keil, 2020).

Standardised cognitive tests often require animals to be isolated (e.g. Ruby & Buchanan-Smith, 2015), handled by a human, and/or given food reinforcement by a human (e.g. Morton et al., 2013; Nawroth, Brett, et al., 2016; Nawroth et al., 2014). However, there is only little research on how the animal is affected by these test conditions, e.g. by isolation (Ruby & Buchanan-Smith, 2015) and interaction with humans (Whitehouse et al., 2013). The presence of a familiar human and the positive association with food may contribute to behavioural changes and may reduce fear responses towards novel humans and ease handling in future situations (Boissy & Bouissou, 1988; Boivin et al., 1992). In addition, separation from the group is stressful for most social animals and has been found to increase vocalisations, heart rate and cortisol levels (Aschwanden et al., 2008a; Boissy & Le Neindre, 1997; Da Costa et al., 2004; Patt et al., 2013; Siebert et al., 2011) and to alter behavioural responses towards novel environments and towards handling (Grignard, Boissy, Boivin, Garel, & Le Neindre, 2000; Veissier & le Neindre, 1992). Habituation to frequent isolation in a novel environment and positive experiences with humans during standardised cognitive tests may also prepare animals to cope better with stress in future similar situations. To assess whether standardised cognitive testing per se is associated with changes in stress reactivity, it is necessary to disentangle the effect of testing from effects of confounding factors such as the human contact and the isolation during testing.

In this study, we investigated whether long-term cognitive stimulation by means of standardised object-choice tasks, including discrimination and reversal learning tests and a cognitive test battery, affects goats' stress reactivity in subsequent potentially stressful situations. We conditioned three treatment groups: goats individually exposed to human-presented and reward-associated object-choice tests (COG treatment), goats that received rewards from the experimenter without being administered the object-choice tests and thus could form a positive association between human and food (POS treatment), and goats that were isolated but received neither a reward nor were administered the tests (ISO treatment). All treatment groups were subsequently tested in four standardized stress tests: a novel arena test (NA), a novel object test (NO), a novel human test (NH), and a weighing test (WH). We hypothesize that if long-term cognitive test exposure itself had a positive impact on behaviour and cardiac activity in subsequent stress tests, COG goats would show less responses indicative of stress than POS goats. Furthermore, we hypothesize that if long-term experience with reward-associated human-animal interaction had a positive impact on behaviour and cardiac activity in subsequent stress tests, POS goats would show less responses indicative of stress than ISO. To increase variability of our sample, we tested two selection lines of goats (dwarf and dairy goats) at two sites under comparable conditions (Voelkl et al., 2020, 2018; Würbel, 2017).

Material, Animals, and Methods

Location, animals, and housing conditions

The study was carried out at two locations, at the Centre for Proper Housing of Ruminants and Pigs at Agroscope in Ettenhausen (ET), Switzerland, and at the Leibniz Institute for Farm Animal Biology in

Dummerstorf (DU), Germany. We used two selection lines of goats, dwarf goats, not selected for productivity traits, and dairy goats, selected for high milk yield. In total, we investigated 61 non-lactating female Nigerian dwarf goats and 59 non-lactating female dairy goats. Apart from few exceptions, we used dwarf goats bred in DU. The only selection aim in this population was to avoid inbreeding. The potential milk yield of Nigerian dwarf goats does likely not exceed 0.3 kg per day (Akinsoyinu et al., 1977). As it was common practice in DU, dwarf goat kids stayed with their dams for six weeks before they were weaned. Regarding dairy goats, we used three of the most common high-producing dairy breeds in Switzerland and Germany and their crossbreds, namely Saanen (n = 15), Chamois Coloured (n = 12), Saanen x Chamois (n = 3), and Deutsche Edelziegen (n = 30). These breeds have potential milk yields of up to 3 kg per day (Vacca et al., 2018). In accordance with common practice in the dairy goat industry, the dairy goat kids were separated from their dam shortly after birth and artificially raised.

At ET, we housed 30 dwarf goats and 30 dairy goats (15 Saanen, 12 Chamois Coloured, 3 Saanen x Chamois crossbreds). The dwarf goats were born in DU and moved to ET in June 2017. The dairy goats were born on different Swiss farms and were moved to ET in June/July 2017. At DU, we housed 31 dwarf goats and 29 dairy goats (Deutsche Edelziege). The dwarf goats were born in DU, except of eight animals which we bought from two German farms due to shortage of female animals in the facility's own breeding stock. The dairy goats were born on a German farm and were moved to DU in July 2018. At the age of 7–8 months, all goats were moved to pens of 10–11 goats each, corresponding to three groups of dairy goats and three groups of dwarf goats at both locations. The total area of each dwarf goat pen was 14 m^2 (approximately 3.6 m x 3.9 m), including of a deep-bedded straw area of 11 m^2 (approximately 2.8 m x 3.9 m) and a 0.5 m elevated feeding place (1.4 m^2). The total area of each dairy goat pen was 17.7 m^2 (approximately 3.9 m x 4.55 m) including of a deep-bedded straw area of 13.4 m^2 (approximately 4.55 m x 2.95 m) and a 0.65 m elevated feeding place (1.82 m^2). Hay was provided behind a feeding fence at the feeding place twice a day at around 8 am and 4 pm in ET and at around 7 am and 1 pm in DU. Each pen had one drinker and a lick block for mineral supply. Additional structures in the straw-bedded area included a wooden bench (for dairy: 2.4m long, 0.6m high, 0.62 m wide; for dwarf: 2.3 m long, 0.5 m high, 0.5 m wide) along the wall of the pen and a round wooden table (0.8 m high, Ø 1.1 m) in the centre of the pen.

All animal care and experimental procedures were performed in accordance with the relevant legislative and regulatory requirements of the corresponding country and the ASAB/ABS Guidelines for the Use of Animals in Research (ASAB & ABS, 2018). All procedures involving animal handling and treatment were approved by the Cantonal Veterinary Office, Thurgau, Switzerland (Approval No.

TG04/17 – 29343) and the Committee for Animal Use and Care of the Ministry of Agriculture, Environment, and Consumer Protection of the federal state of Mecklenburg-Vorpommern, Germany (Approval No. 7221.3-1.1-062/17).

Treatment groups

Three goats from each of the twelve pens (9–11 goats each) were pseudo-randomly assigned to one of the three treatment groups: COG (n = 36), POS (n = 36), and ISO (n = 36). Except one pen, all pens housed 1–2 extra goats not assigned to a treatment group to replace others in case of e.g. disease or injury. Over a period of 4–5 months, goats from the COG group were exposed to standardised cognitive tests, in the form of object-choice tasks to assess their discrimination and reversal learning skills as well as their ability to use physical cues and human gestures to locate a hidden reward in a cognitive test battery (see Appendix II: Text and Table S1 for details regarding habituation and cognitive tests). During these tests, COG goats received food rewards from the experimenter for correct responses. The POS group was not exposed to cognitive tests, but received a similar amount of rewards as individuals in the COG group (= median number of rewards received by COG group in the previous test session), provided by the experimenter in the test arena at random times, but over a similar amount of time as the COG group (= median time taken by COG group to finish all trials in the previous test session). Contrasting COG versus POS allows investigating the effect of the cognitive testing itself, disentangled from the effects of the positive association with the human and the isolation from the group during testing. Individuals administered the ISO treatment neither participated in cognitive tests nor did they receive rewards by the experimenters in the test arena. Instead, they were isolated over a similar amount of time as the COG and the POS group in the same arena (= median time taken by COG group to finish all trials in the previous test session). Contrasting POS versus ISO allows investigating the effect of the positive association of the human with food, disentangled from the effect of isolation from the group during testing. To control for caloric intake, ISO animals received the same amount of food rewards as POS and COG goats, scattered over the floor of the waiting room (to avoid positive association with the human) before they were isolated.

Stress tests

After the cognitive test phase (as described in Appendix II: Text and Table S1), we measured reactivity of all goats towards different stressors in four tests: a novel arena test (NA), a novel object test (NO), a novel human test (NH), and a weighing test (WH) in a mobile scales. In ET, all goats were between 15 and 17 months old when the first stress test (NA) started (mean ± SD: dwarf goats 509 ± 0.9 d, dairy goats 468 ± 3.5 d). In DU, goats were around 18 months old at test start (mean ± SD: dwarf goats 557 ± 3.4 d, dairy goats 540 ± 0 d). All tests were completed within 6 weeks at both locations. While in DU, the NA, NO, and NH took place in a room in the same building where goats were housed, in ET goats had to be moved to a different building for these tests. The WH took place right in front of the goats' home pens at both locations. During all tests, acoustic and olfactory contact between the test subject and its peers was possible.

Novel arena test (NA)

The NA was used to assess reactivity towards isolation in a novel environment. We measured goats' reactivity in an arena (3 m x 5 m) with opaque walls (2 m high), a grid drawn on the floor (with 12 segments), and a start box (1 x 1 x 1 m) attached on the outside but connected to the arena (Fig. 1). Each subject was placed in the start box for 20 s to standardize the beginning of each test. After 20 s, the start box was opened and the animal allowed to enter the arena. Right after the goat had entered the arena the start box was closed and the test subject stayed in the arena for 5 min. Each goat was tested once in the NA on one of two consecutive days on which, for a given selection line and site, all NA were performed.

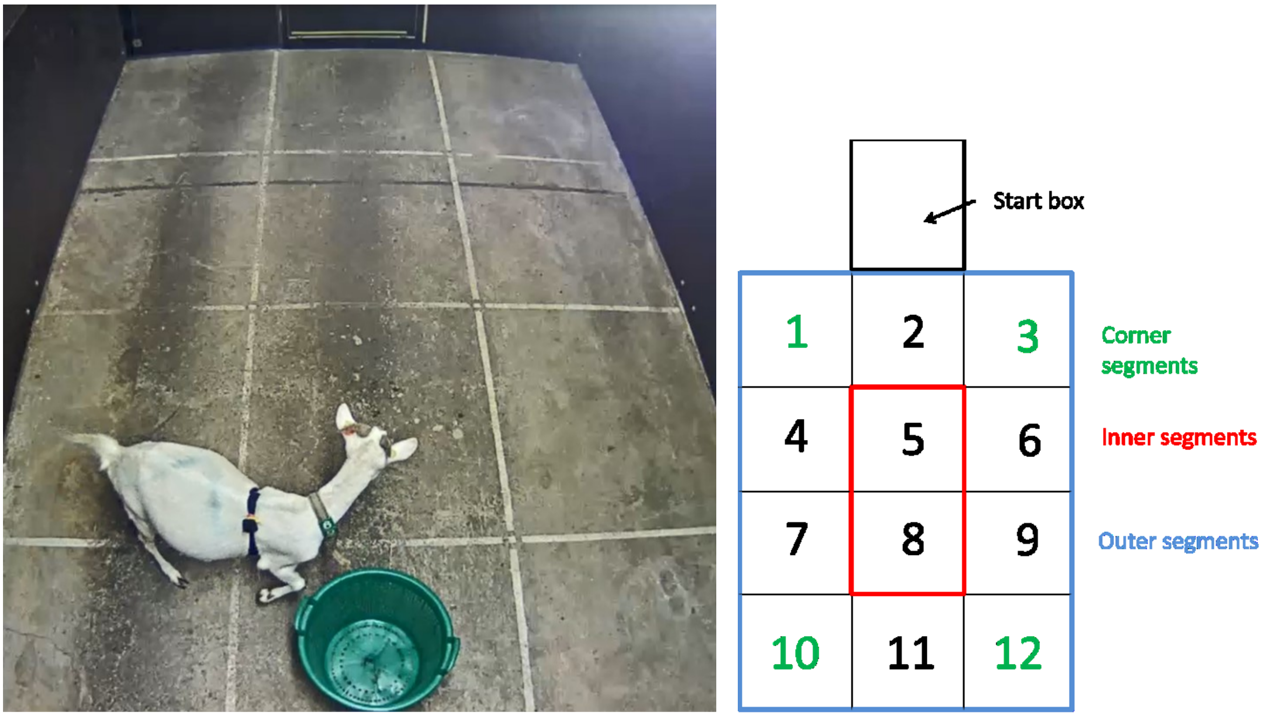


Figure 6: Picture of the arena (3 x 5 m) used for the novel arena (NA), novel object (NO), and novel human (NH) tests. The arena was divided in 12 segments which were drawn on the floor. The numbering used to identify the segments when encoding the animal behaviour are indicated in the sketch on the right.

Novel object test (NO)

With the NO, we assessed the reactivity of animals towards a novel object. In the arena described above (Fig. 6), the same procedure as in the NA was applied, however a novel object (DU: green bucket, 30 cm high, Ø 50 cm, ET: brown bucket, 40 cm high, Ø 30 cm) was placed in segment 11 of the arena. Each goat was tested once in the NO on one of two consecutive days on which, for a given selection line and site, all NO were performed.

Novel human test (NH)

The reactivity towards a human was measured using a NH. In this test, the goat was confronted with a stationary novel human (always wearing a lab coat) standing in the back (segment 11) of the arena (Fig. 6) and looking at the wall above the entrance gate not making eye contact with the goat. All else followed the same procedure as in the NA. Each goat was tested once in the NH on one of two consecutive days on which, for a given selection line and site, all NH were performed.

Weighing test (WH)

Finally, we scored goats' responses to handling during weighing in a mobile scales (FX 21A, Agro Sigmer, Fig. 7). The test subject was first led onto the scales where it stayed for 3 min, starting as soon as the gate of the scales was closed. If the goat refused to walk into the scales, the experimenter gently pushed it. After this period, the experimenter opened the gate on the opposite side of the scales and the goat was allowed to walk out. If the goat refused to walk out, the experimenter gently encouraged it to leave the scales. Two experimenters were simultaneously scoring how easily the goat entered the scales, how it behaved in the scales and how easily it exited the scales using a scoring system adapted from D`Eath et al. (D'Eath et al., 2009; Table 3). During weighing, experimenters were positioned on either side of the scales within 1 m distance. Testing of all goats took place on a single day (one trial per goat).

Figure 7: Picture of a dairy goat on the scales wearing the harness used to record cardiac measures.

Table 3: Scoring system used for scoring responses to handling during weighing (adapted from D'Eath et al., 2009).

Score	Description
Entering the scales	
1	Goat is very difficult to move and tries to escape. Hard pushing or lifting of legs by experimenter necessary.
2	Goat is difficult to move into the scales and tries to resist. Some pushing by experimenter needed.
3	Goat walks into the scales with little encouragement by experimenter.
4	Goat walks/runs forward into the scales voluntarily
Weighing on the scales	
1	Goat moves around a lot during weighing, many escape attempts, rearing, and vocalizing
2	Goat is mostly moving, showing some escape attempts, vocalizes
3	Goat moves around a bit during weighing, max. one escape attempt, vocalizes
4	Goat stands mostly still during weighing or is trying to lie down, no escape attempts, little vocalization
Exiting the scales	
1	Goat resists and is very difficult to push out of the scales
2	Goat moves out of the scales after some pushing by experimenter
3	Goat (slowly) leaves of its own accord once the door is opened
4	Goat quickly runs out of scales, no hesitation

Behavioural and cardiac measures

The behavioural and acoustic responses were videotaped with a camcorder (ET: Sony HDR-CX240E; DU: Panasonic HDC-SD60) and an external microphone (Table 4). Additionally, cardiac measures of goats were recorded in each of the four test situations: in the NA, NO, and NH for 5 min starting when the goat entered the arena, in the WH for 3 min starting as soon as the goat entered the scales. Goats were equipped with an electrocardiogram (ECG) acquisition harness (BioHarness® system, MLE120X BioHarness Telemetry System, Zephyr Technology Corporation, Annapolis, MD, U.S.A.) which was fitted tightly around the chest behind the front legs of the animals (Fig. 7). ECG gel was applied on the parts of the belt containing the electrodes before each use. Electrodes were positioned on the left side of the chest, with one electrode placed close to the sternum and the other electrode over the right scapula. ECG was recorded at 250 Hz by a logger integrated in the BioHarness® system and transmitted live to a laptop with the software AcqKnowledge (v.4.4, BIOPAC System Inc). To minimize stress linked to the novel device, the first test, the NA, was preceded by one day of habituation to wearing the harness.

For each individual, a baseline recording of cardiac activity was taken, either in its social group (DU and ET: NA, NO, NH, DU: WH) or individually in front of its home pen (ET: WH), shortly before it was taken to the test room. Cardiac activity for the baseline was measured for 10 min for the NA, NO, and NH and for 5 min for the weighing test. Right after the baseline measurement was taken, functionality of the harness was assessed and, if needed, readjusted, and the goat was led to the starting box of the test room for the NA, NO, and NH or in front of the scales in case of the WH.

Behaviour coding and ECG processing

The videos (with audio) were analysed with the Observer XT software (v.13, Noldus Information Technology, The Netherlands) to determine frequency and duration of behavioural responses (Table 4). The ECG was processed using a custom-made software called "EasieRR" (Rasmussen, Rosenberger, & Langbein, 2020, see Appendix IV for more details). R peaks were automatically detected using the software's peak prominence algorithm (peak prominence set at 0.05–0.07 depending on the ECG trace) and were reviewed visually for ectopic cardiac beats, missed beats, and outliers by three investigators. As there was too much noise in the data to analyse the whole 3 (WH) to 5 min (NA, NO, NH) recordings of the stress tests, we decided to only analyse selected time intervals of 20 s within certain time windows. For the baseline recordings, this time window started after 60 s to avoid measuring the effect of putting the harness on. For the recordings of the stress tests, the test time was split in half and we analysed 1–3 time intervals of 20 s in the first half and 1–3 time intervals of 20 s in the second half of each test per animal. We selected time intervals in which the heart beats on the ECG trace were clearly visible and the signal-to-noise ratio was adequate. If artefacts in heart rate data could not be avoided by deleting a max of 5% of artefacts (max. of 3 artefacts in a row), the time range was discarded (Jan Langbein et al., 2004; Mohr, Langbein, & Nürnberg, 2002). However, likely due to the high activity in the stress tests and presumably bad electrode contact, data quality was too low to use any parameters of heart rate variability. In the statistical analysis we therefore only included the baseline-subtracted heart rate which is the average of the 20 s time intervals from the recordings of the stress test minus the averaged baseline values.

Statistical analysis

All statistical analyses were performed in R v4.0.3 (R Core Team, 2020). We excluded 7 dwarf goats and 1 dairy goat that did not participate because they were too stressed to be tested in some of the standardized cognitive tests (object-choice tests) administered as part of the COG treatment. Additionally, in each of the four stress tests, several goats had to be excluded from the analysis because of missing values due to technical failure of the video camera, the microphone, or the ECG device (NA: n = 13 goats, NO: n = 11 goats, NH: n = 11 goats, WH: n = 6 goats).

For each test, we performed a Principal Component Analysis (PCA) to reduce the large set of behavioural and cardiac parameters to a smaller set of components. To improve normality before the PCA analysis, all parameters were transformed applying Yeo-Johnson transformation with the R package bestNormalize (Peterson & Cavanaugh, 2019). We used the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) and the Bartlett test to assess the adequacy of our data for PCA (functions *KMO* and *cortest.bartlett* from psych R package (Revelle, 2020). According to Budaev (2010) an overall MSA of less than 0.5 would be unacceptable. If the overall MSA was below 0.5, the MSA of each contributing variables was calculated and variables with MSA < 0.4 were dropped before the test was re-run. This procedure was continued until the overall MSA was >= 0.5. The Bartlett test was finally applied to test whether the correlation matrix is factorable (i.e., the correlations differ from 0). This procedure resulted in a total of nine different behavioural measures and one cardiac measure to be used in the different PCAs (5 measures in NA, 6 in NO, 6 in NH, 4 in WH, see Table 4).

Table 4: Definitions of behavioural and cardiac measures that were included in the different Principal Component Analysis.

Test	**Measure**	**Type**	**Definition**
NA, NO, NH	Total time inactive	sec	Animal is standing still, legs are not moving.
NA, NO, NH	Vocalizing	frequency	Animal is vocalising with open or closed mouth.
NO	Object contact	frequency	Animals' snout touches or is within 5 cm of the object.
NH	Human contact	frequency	Animals' snout touches or is within 5 cm of the human.
NA, NO, NH	Change of segment	frequency	Animal moves to another segment with at least both front legs (see Fig. 6).
NA, NO, NH	Staying in inner segments	sec	Segments number 5 and 8 (see Fig. 6).
NA, NO, NH, WH	Heart Rate	bpm	Baseline-subtracted heart rate during test
WH	Weighing Score	Score 1-4*	Mean score (Table 3) given by experimenters during weighing on scales, from high (=1) to low (=4) stress.
WH	Exiting Score	Score 1-4*	Mean score (Table 3) given by experimenters during exiting of scales, from high (=1) to low (=4) stress.
WH	Entering Score	Score 1-4*	Mean score (Table 3) given by experimenters during entering of scales, from high (=1) to low (=4) stress.

* scoring system according to D'Eath (2009)

The PCA was conducted with the R function *principal* (psych R package) using varimax rotation. To choose the final number of extracted PCs to retain, we applied *nScree* analysis with the *plotnScree* function from the R package nFactors (Raîche, Walls, Magis, Riopel, & Blais, 2013), which uses four methods: the optimal coordinates, the acceleration factor, the parallel analysis, and the Kaiser-Guttman rule. We decided to retain two components in the NA, NH, and NO for the final PCA calculation, since three of the previously mentioned methods suggested so. For the WH, only the first component was retained. As on the second component only the behavioural measure Entering Score loaded considerably (> |0.5|), we used directly the Yeo-Johnson-transformed Entering Score and not the rotated component from PCA for further analysis.

To analyse the effects of the treatments (COG, POS, ISO) on the rotated PCA component scores and on the Entering Score, we employed linear mixed-effects models using the *lmer* function from the R package lme4 (Bates, Maechler, et al., 2015).

For all models, the formula in lme4 syntax was the following:

response ~ 0 + Treatment:SelectionLine + (1|Site/Pen)

We considered a treatment effect individually for each selection line through a corresponding interaction term as fixed effect [0 + Treatment:SelectionLine]. Besides this fixed effect, a random intercept for pen nested within site (1|Site/Pen) was included to account for potential effects of the affiliation to the home pen (A–F, U–Z) and site (ET, DU). To investigate differences in stress responses between the treatments and between the selection lines, we tested contrasts for the fixed effect using the *glht* function from R package multcomp (Hothorn et al., 2008). P-values for fixed effect estimates and for the contrasts were obtained using Wald Z-tests (*summary.ghlt* function, multcomp package). Confidence bands for fixed effect estimates were obtained using the *predict.MerMod* function (lme4 package) in conjunction with the *bootMer* function (lme4) for parametric bootstrapping (10^4 bootstraps). Only the uncertainty in the fixed effects was taken into account (parameter re.form = ~ 0 in *predict.MerMod*).

Results

Principal component analyses

Table 5 and Figure 8 present the results of the four PCAs of the stress tests. In the PCA of the NA, we retained two rotated components (RCs, Overall MSA = 0.52). The first RC (NA_RC1) explained 30 % of the total variance. It contained a high positive loading (> 0.5) for the frequency of segment changes and a high negative loading (< -0.5) for the duration being inactive. Goats that loaded highly on this component were therefore labeled ‘active in NA’ (Table 4). The second RC (NA_RC2) explained 27% of the total variance. It contained high positive loadings for heart rate, frequency of vocalisations, and duration in inner segments. As these behaviours are indicative of a high reactivity to isolation, goats that loaded highly on this component were termed ‘reactive in NA’. In the PCA of the NO (Overall MSA = 0.56) the first component (NO_RC1) explained 40% of total variance. It contained a high negative loading for the duration being inactive and a high positive loading for the frequency of segments changes. Goats loading highly on this component were thus termed ‘active in NO’. The second component (NO_RC2) explained 31% of total variance and contained high loadings for the frequency of object contacts as well for heart rate. As the frequency of object contacts is indicative for explorative-like behaviours, goats that loaded highly on this component were termed ‘explorative in NO’. In the PCA of the NH (Overall MSA = 0.57), the first component (NH_RC1) explained 34% of the total variance. It contained high positive loadings for frequency of vocalisations as well as frequency of human contacts and a high negative loading for heart rate. Goats that loaded highly on this component were therefore termed ‘sociable in NH’. The second component in the NH (NH_RC2) explained 33% of the variation in the data. We negated its loadings (multiplication by -1) to ease interpretation. The negated component contained a high negative loading for duration being inactive and a high positive loading for frequency of segment changes. Goats that loaded highly on this (negated) component were termed ‘active in NH’. In the PCA of the WH (Overall MSA = 0.59), the single component retained described behaviour indicative of reactivity towards handling and weighing (WH_RC1) and explained 51% of total variance. It contained high positive loadings for exiting score and heart rate and a high negative loading for the weighing score. Goats loading highly on this component were termed ‘reactive in WH’.

Table 5: Principal Component Analysis (PCA) results of each of the four stress tests (NA, NO, NH, WH) with eigenvalues, percentage of the total variance, and loadings of the rotated components (RC1 and RC2), along with communalities (= proportion of variance in the variable explained by the components). Loadings above 0.5 and below -0.5 are shown in bold.

novel arena test (NA)	'active in NA' (RC1)	'reactive in NA' (RC2)	Communalities
baseline-substracted heart rate	0.0	**0.7**	0.5
duration being inactive	**-0.8**	-0.1	0.6
frequency of vocalisations	0.3	**0.6**	0.5
frequency of segment changes	**0.8**	-0.1	0.7
duration in inner segments	-0.3	**0.7**	0.6
eigenvalues	1.5	1.3	
% of variance	30.0	27.0	
novel object test (NO)	'active in NO' (RC1)	'explorative in NO' (RC2)	
baseline-substracted heart rate	-0.04	**0.9**	0.8
duration being inactive	**-0.8**	-0.3	0.8
frequency of object contacts	0.3	**0.7**	0.5
frequency of segment changes	**0.9**	-0.01	0.8
eigenvalues	1.6	1.3	
% of variance	40.0	31.0	
novel human test (NH)	'sociable in NH' (RC1)	'active in NH' (-RC2*)	
baseline-substracted heart rate	**-0.6**	-0.2	0.5
duration being inactive	0.0	**-0.9**	0.8
frequency of vocalisations	**0.7**	0.2	0.5
frequency of human contacts	**0.8**	-0.2	0.7
frequency of segment changes	0.2	**0.9**	0.8
eigen values	1.7	-1.7	
% of variance	34.0	-33.0	
weighing test (WH)	'reactive in WH' (RC1)		
baseline-substracted heart rate	**0.6**	**-**	0.4
weighing score	**-0.8**	**-**	0.6
exiting score	**0.8**	**-**	0.6
eigenvalues	1.5	-	
% of variance	51.0	-	

* Loadings for this component have been negated (multiplied by -1) for ease of interpretation.

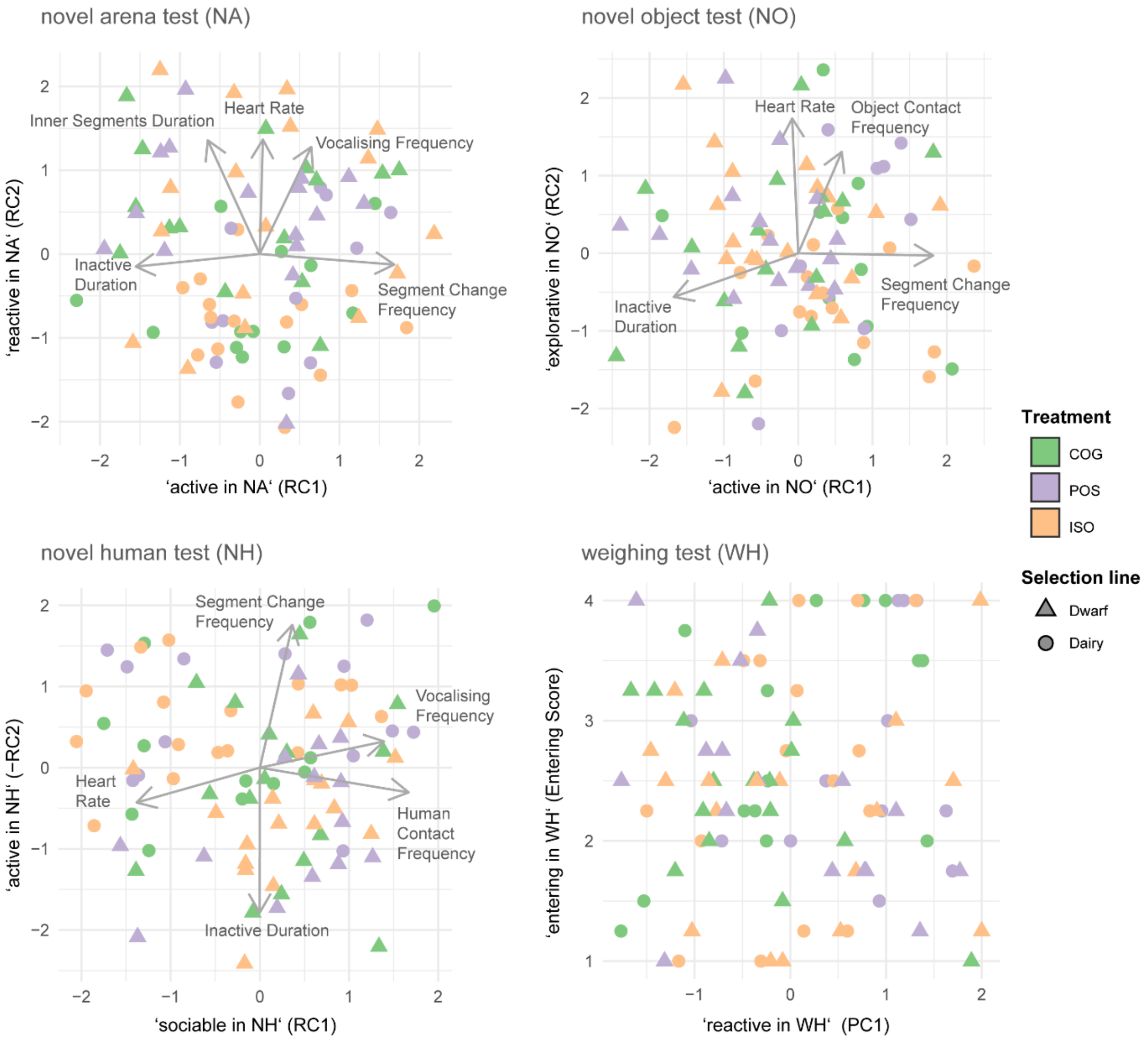


Figure 8: PCA biplot for each of the four stress tests: novel arena (NA), novel object (NO), novel human (NH), and weighing test (WH).

Linear mixed-effects models (LMM)

We analyzed the effect of treatment and selection line on the above described rotated PCA components and the Entering score in WH. Fixed effect estimates are shown in Fig. 9 and fixed effect contrasts with respect to treatment and selection line are listed in Table 6 and Table 7, respectively. In the NA, we found no indication for treatment differences in activity (NA_RC1, Table 6). With weak statistical support POS dwarf goats were more reactive than ISO dwarf goats (NA_RC2, $p = 0.08$). When comparing selection lines, dwarf goats were less reactive to isolation than dairy goats (all $p < 0.03$, Table 7).

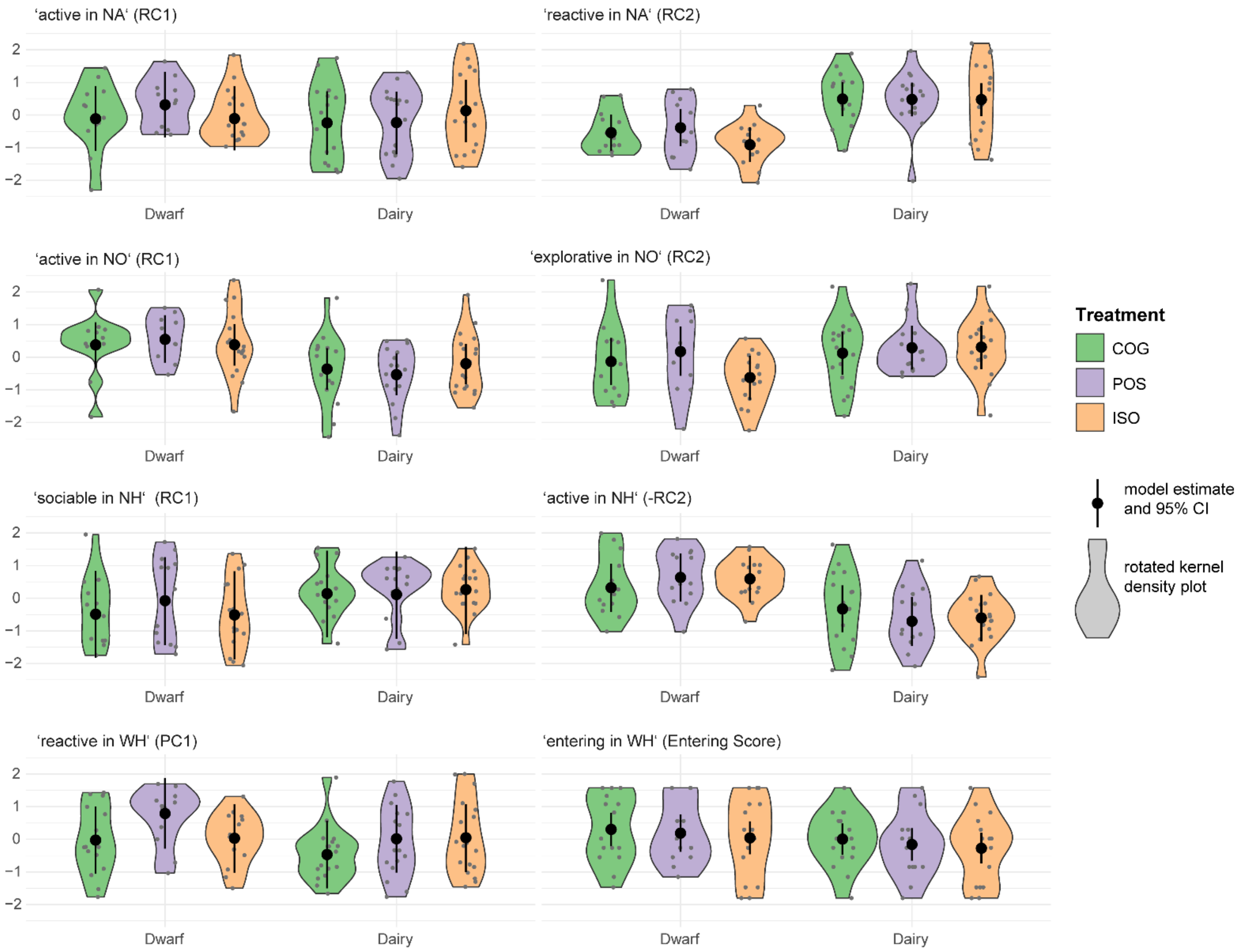


Figure 9: Treatment (COG, POS, ISO) fixed effect estimates with confidence intervals from LMMs with PCA component scores as responses. Distributions of component scores (grey dots) are summarized as rotated kernel density plots (violin plots).

In the NO, we found no indication for treatment differences in activity (NO_RC1, Table 6). Dwarf goats were generally more active than dairy goats, but with varying statistical certainty within the different treatment groups (all $p < 0.08$; Table 7). Dwarf goats from the POS treatment were more explorative (NO_RC2) than dwarf goats from the ISO treatment (NO_RC2, $p = 0.03$). When comparing between selection lines, we found that ISO dwarf goats were less explorative than ISO dairy goats ($p < 0.01$; Table 7).

In the NH, we found, with limited statistical support, that POS dwarf goats were more sociable than ISO dwarf goats (NH_RC1, $p = 0.09$). Sociability of selection lines differed within the COG and ISO treatment, with COG and ISO dwarf goats being less sociable towards a novel human than COG and ISO dairy goats, respectively (both $p < 0.02$; Table 7). No indication for treatment differences were found for activity (NH_-RC2). Dwarf goats were generally more active compared to dairy goats (all $p < 0.03$; Table 7).

In the WH, COG goats were less reactive towards handling (WH_RC; Table 6) than POS goats with varying statistical certainty for dwarf (p = 0.01) and dairy (p = 0.08) goats. POS dwarf goats were furthermore more reactive towards handing than ISO dwarf goats (p = 0.01). Comparing selection lines, we found for the POS treatment that dwarf goats were more reactive towards handling than dairy goats (p = 0.02; Table 7). This was not the case within the COG treatment and the ISO treatment. The LMM with the 'Entering in WH' as response variable did not indicate differences between treatments or selection lines (Table 6 and 7).

Table 6: Summary of treatment contrasts from linear mixed-effects models of all stress tests. The respective results for fixed and random effects are in Appendix II: Table S2-S9.

'active' in the novel arena test (NA_RC1)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.43	0.35	0.23	0.01	0.31	0.98
ISO - POS	-0.42	0.34	0.22	0.36	0.30	0.22
ISO - COG	0.00	0.32	0.99	0.37	0.30	0.22
'reactive to isolation' in the novel arena test (NA_RC2)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.15	0.31	0.62	-0.01	0.27	0.96
ISO - POS	-0.52	0.30	0.08	0.00	0.26	0.99
ISO - COG	-0.37	0.28	0.19	-0.02	0.26	0.95
'active' in the novel object test (NO_RC1)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.17	0.38	0.66	-0.17	0.31	0.58
ISO - POS	-0.16	0.36	0.64	0.34	0.30	0.27
ISO - COG	0.00	0.34	0.99	0.16	0.30	0.58
'explorative' in the novel object test (NO_RC2)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.31	0.40	0.44	0.17	0.32	0.61
ISO - POS	-0.80	0.38	0.03	0.02	0.32	0.96
ISO - COG	-0.49	0.36	0.17	0.18	0.32	0.56
'sociable' in the Novel human test (NH_RC1)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.41	0.27	0.13	-0.03	0.25	0.92
ISO - POS	-0.44	0.26	0.09	0.15	0.25	0.55

ISO - COG	-0.02	0.26	0.93	0.12	0.24	0.61
‘active’ in the novel human test (NH_-RC2)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.31	0.31	0.32	-0.38	0.29	0.20
ISO - POS	-0.04	0.30	0.88	0.10	0.29	0.72
ISO - COG	0.27	0.30	0.37	-0.27	0.28	0.32
‘reactive towards handling’ in the weighing test (WH_RC)						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.81	0.31	0.01	0.48	0.28	0.08
ISO - POS	-0.76	0.31	0.01	0.03	0.27	0.91
ISO - COG	0.06	0.29	0.85	0.52	0.27	0.06
Entering Score* in the weighing test						
	Dwarf			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	-0.12	0.39	0.77	-0.17	0.35	0.64
ISO - POS	-0.14	0.38	0.71	-0.12	0.35	0.74
ISO - COG	-0.26	0.36	0.47	-0.28	0.34	0.41

* Yeo-Johnson transformed variable (no PCA component)

Table 7: Summary of selection line contrasts from linear mixed-effects models of all stress tests. The respective results for fixed and random effects are in Appendix II: Table S2-S9.

‘active’ in the novel arena test (NA_RC1)									
Treatment	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.13	0.39	0.75	-0.55	0.39	0.17	0.24	0.37	0.52
‘reactive to isolation’ in the novel arena test (NA_RC2)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	1.03	0.39	0.01	0.87	0.39	0.03	1.38	0.37	<0.001
‘active’ in the novel object test (NO_RC1)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.74	0.36	0.04	-1.08	0.38	0.00	-0.58	0.33	0.08
‘explorative’ in the novel object test (NO_RC2)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	0.26	0.35	0.46	0.12	0.38	0.75	0.93	0.32	<0.01

'sociable towards a novel human' in the novel human test (NH_RC1)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	0.63	0.27	0.02	0.19	0.28	0.49	0.78	0.25	<0.01
'active' in the novel human test (NH_-RC2)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.65	0.31	0.03	-1.34	0.32	0.00	-1.20	0.29	<0.001
'reactive towards handling' in the weighing test (WH_RC)									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.44	0.32	0.17	-0.77	0.34	0.02	0.02	0.32	0.96
Entering Score* in the weighing test									
	COG			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.30	0.36	0.41	-0.35	0.38	0.37	-0.32	0.35	0.36

*Yeo-Johnson transformed variable (no PCA component)

Discussion

We hypothesized that if cognitive testing itself has a positive impact on behaviour and cardiac activity in subsequent stress tests, goats with long-term experience with standardized cognitive tests (COG) will show less responses indicative of stress than POS goats, which were not exposed to cognitive tests, but received a similar amount of reward-associated human-animal interaction. Furthermore, we hypothesized that if reward-associated human-animal interaction has a positive impact on behaviour and cardiac activity in subsequent stress tests, POS goats will show less responses indicative of stress than ISO goats, which neither participated in cognitive tests nor received reward-associated human-animal interaction. Overall, our results from the four stress tests do not support the two hypotheses.

Effect of cognitive testing (COG vs. POS)

We had developed the hypothesis of COG vs. POS based on studies which found that cognitively challenged goats and pigs display more exploration and less fearful behaviour and activity in an NA and NO compared to animals that did not receive this cognitive enrichment (Oesterwind et al., 2016; Puppe et al., 2007; Zebunke et al., 2013). Except for the reactivity to weighing in WH we did not find support for effects of cognitive testing per se (COG-POS contrast) on the responses in the four stress tests. One possible explanation for the general lack of a treatment effect is that all goats experienced a lot of environmental stimulation already in their home pens (e.g. climbing and hiding opportunities, ad lib hay and straw). Various studies on farm animals have demonstrated that environmental enrichment, such as straw or climbing racks, can make animals less fearful and more exploratory towards an unknown object (Beattie et al., 2000; Hillmann et al., 2003; Oesterwind et al., 2016). Furthermore, all goats were confronted with additional environmental stimuli in the process of the experiment. They were repeatedly taken out of their pens and exposed to the test environment (that included novel visual, acoustic, and olfactory cues). They were also regularly handled to be equipped with loggers to measure activity in the home pen (data not shown) and with harnesses to measure cardiac activity. Taken together, all goats may have received such a high level of environmental stimulation that the additionally administered cognitive stimulation through the COG treatment was rendered negligible. Also, in other studies, intermittent exposure to mildly stressful situations in early life has been found to reduce reactivity and to improve resistance to subsequent stressors in mice and monkeys (Brockhurst, Cheleuitte-Nieves, Buckmaster, Schatzberg, & Lyons, 2015; Parker, Buckmaster, Schatzberg, & Lyons, 2004). A study on horses reported that animals which spend more time outside of their stalls and that are used to be ridden by two or more riders have a less pronounced adrenal response than box stalled horses or horses with only one rider (Sauer et al., 2019).

Another explanation for not finding an effect of the COG treatment could be that the cognitive testing per se was not perceived as enriching by the goats. It is even possible that certain aspects of the testing procedure (i.e. uncontrollability or isolation) were experienced as negative and thus did not lead to the expected decrease of stress reactivity in the COG goats. For example, unpredictable training events (Doyle et al., 2011; Galhardo et al., 2011) or the initial frustration caused by failure in a novel task at early stages of learning may cause stress levels to increase (Jan Langbein et al., 2004). Particularly for social animals such as goats, isolation is a major stressor (Aschwanden et al., 2008a; Patt et al., 2013; Siebert et al., 2011). Although we made sure all goats were sufficiently habituated to isolation and the GOG goats willingly participated in the tests, we cannot exclude that testing was still perceived as stressful by some individuals. Also, studies that found cognitive enrichment effects on stress reactivity had the cognitive enrichment device incorporated in the home pen and the animals actively decided to interact with the device (Jan Langbein et al., 2004; Manteuffel et al., 2009; Puppe et al., 2007). Such voluntary interaction and exploration of an enrichment device may allow the animal to experience agency over its environment which is not only self-rewarding, but also enhances the animal's competence to deal with future challenges (Spinka & Wemelsfelder, 2017). In the current study, the experimenters but not the individual goat decided when the cognitive tests were administered. This may have taken the effect of agency out of the equation and could explain why we did not find reduced stress reactivity in COG versus POS animals. The discrepancy between our study with the administered cognitive stimulation and the studies with the enrichment devices therefore suggests that further research is needed to identify the aspects associated with cognitive enrichment (and the conditions under which it takes place) relevant for the reduced stress reactivity.

Effect of reward-associated human-animal interaction (POS vs. ISO)

We also hypothesised that, if goats are sustainably positively affected by the reward associated interaction with the experimenter, POS goats would be less stress reactive than ISO goats. Overall, we did not find consistent support for this hypothesis. Only in dwarf goats and with varying statistical certainty we found some differences in stress test responses between the POS and the ISO treatment group. Dwarf goats from the POS versus from the ISO treatment were more sociable (NH) and more explorative (NO), presumably both indicating reduced stress and therefore supporting the hypothesis above. However, POS versus ISO dwarf goats were also more reactive towards isolation (NA) as well as more reactive towards weighing (WH). Assuming high levels in the latter two are indicative for stress, these observations would contradict the hypothesis that reward-associated human-animal interaction reduces stress reactivity. However, the assignment of the PCA components and their underlying measures to higher or lower stress is not unambiguous. For example a higher heart rate was associated both with higher reactivity toward isolation (indicative of more stress) as well as with increased exploration (indicative of less stress; Fig. 8). Also, reactivity in the WH consisted of a high negative loading for the weighing score and a high positive loading for the exiting score. This would mean, goats that were more stressed during weighing, tended to be less stressed during exiting. But, the stress scores we adapted from D'Eath et. al. (2009) might not be able to adequately capture stress reactivity in goats depending on the type of their reactions. While an active stress response might result in low exiting score and low weighing score, a passive stress response (e.g. freezing) would result in a low exiting score and a high weighing score. Therefore, single components or stress tests have to be interpreted very cautiously if — like in our case — consistent patterns are not apparent.

However, it is worth mentioning that only for dwarf goats but not for dairy goats we found support for POS versus ISO differences at least for some stress test measures. Dairy goats seemed to be generally less stressed by human presence compared to dwarf goats and therefore less affected by the administered human-animal interaction (POS treatment). These differences between selection lines could be explained by genetic predisposition as selection for high productivity has been shown to increase sociability towards humans (C. Lindqvist & Jensen, 2008; Schütz & Jensen, 2001). They could also be explained by differences in early rearing. Whereas the dwarf goats were reared with their dams until they were 6 weeks old, the dairy goats were separated immediately after birth. It has been shown that goats raised by their dams are more reluctant to get in contact with humans (Boivin & Braastad, 1996). Early human-animal interactions have been found to have lasting effects on temperament and behaviour (Lyons, 1989; Lyons & Price, 1987; Lyons et al., 1988). For example, goat kids gently handled at one week of age remained closer to a human observer and vocalized less when isolated than kids first handled at six months of age or not handled at all (Boivin & Braastad, 1996). Therefore, phenotypic variation in stress reactivity must be considered when assessing stress reactivity and when interacting with goats.

Conclusion

We found that long-term experience with standardized cognitive testing *per se* does not reduce stress reactivity of goats in subsequent stress tests. Further research is needed to identify the aspects associated with cognitive enrichment relevant for the reduced stress reactivity found in other studies. For dwarf goats but not for dairy goats we found support for an effect of reward-associated human-animal interactions at least for some stress test measures. This highlights the need to consider phenotypic variation when assessing stress reactivity or interacting with goats.

Acknowledgements

We thank the staff of the Agroscope Research Station in Ettenhausen, Switzerland, and of the Leibniz Institute for the Biology of Farm Animals in Dummerstorf, Germany, for assisting in this experiment and for caring for the animals. We additionally want to thank Katrin Siebert for her help with the software *The Observe*r, Jonas Barkhau and Jana Benner for video analysis, Charlotte Hursey and Franziska Kleindienst for help with ECG analysis, and Heather Neave for support regarding PCA analysis. This work was supported by an SNF grant (310030E-170537) to Nina Keil and a DFG grant (LA 1187/6-1) to Jan Langbein.

Chapter IV

Based on

Performance of goats in a detour and a problem-solving test following long-term cognitive test exposure

Rosenberger, K.[1,2]*, Simmler, M.[3], Langbein, J.[4], Keil, N[1], Nawroth, C.[4]*

[1]Swiss Federal Veterinary Office, Centre for Proper Housing of Ruminants and Pigs, Agroscope, 8355 Ettenhausen, Switzerland

[2]Graduate School for Cellular and Biomedical Sciences, University of Bern, 3012 Bern, Switzerland

[3]Digital Production Group, Agroscope, 8355 Ettenhausen, Switzerland

[4]Leibniz-Institute for Farm Animal Biology, Institute of Behavioural Physiology, 18196 Dummerstorf, Germany

Manuscript under review in Royal Society Open Science (since April 2021)

Author contributions

K.R. and C.N. collected the data. M.S. and K.R. analysed the data. K.R. drafted the manuscript, wrote the main parts of the manuscript, made all tables and figures except of (Fig. 12-15). M.S. provided statistical support and made all corresponding statistical plots (Fig. 12-15). C.N., N.K. and J.L. conceptualised the study and provided supervision. All authors interpreted the data and provided critical feedback on the manuscript.

Abstract

Cognitive research in long-lived species commonly involves using the same animals in different experiments. It is unclear whether the participation in cognitive tests can noticeably alter the performance of individuals in subsequent conceptually different tests, so we investigated whether exposure to cognitive tests affects future test performance of goats. We conditioned three treatment groups: goats with long-term exposure to human-presented object-choice tests (for visual discrimination and reversal learning tests + cognitive test battery), goats that were isolated as for the test exposure but received a reward from the experimenter without being administered the object-choice tests, and goats that were isolated but neither received a reward nor were administered the tests. All treatment groups were subsequently tested in two conceptually different cognitive tests, namely a spatial A-not-B detour test and an instrumental problem-solving test. We tested dairy goats, selected for high productivity, and dwarf goats, not selected for production traits, each at the same two research sites. We found only subtle differences between treatments with respect to the goats' detour or problem-solving performance. However, high variation was observed between the research sites, the selection lines, and among individuals, highlighting potential pitfalls of making accurate comparisons of cognitive test performances.

Introduction

Comparative approaches that identify key socio-ecological drivers of certain cognitive traits are important to understand the evolutionary origins of human cognition. Comparative research involving cognitive tests in non-human species is targeting animals in research facilities, including laboratory, zoo and farm animals (MacDonald & Ritvo, 2016). Depending on the species, this line of research commonly uses the same animals repeatedly over several experiments, which is often a necessity due to restricted financial resources and/or constraints in the available number of animals at the facility. Although it is known that animals can pick up learned contingencies over similar tests (Harlow, 1949; Jan Langbein, Siebert, Nürnberg, et al., 2007), it is still unclear whether the participation in cognitive tests alters the performance of individuals in subsequent conceptually different cognitive tests. This raises the question whether potential long-term behavioural change due to repeated cognitive testing hampers comparability of data and replicability of study findings that are obtained from subjects with different histories of test exposure.

To answer this question, a variety of potentially confounding factors that can affect test performance in subsequent cognitive tests need to be disentangled. First, the cognitive testing *per se* may have an effect on future test performance. The frequent operation of similar cognitive tests enhances the ability to learn the test-inherent contingencies and thus alter test performances when compared with naïve subjects: the recall of previously learned information can help to facilitate subsequent learning in a conceptually similar following test (Harlow, 1949; Jan Langbein, Siebert, Nürnberg, et al., 2007). For instance, prior experience in detour tests with transparent obstacles also improved performance of pheasants in a subsequent novel detour test (Van Horik et al., 2018). Although this effect of 'learning to learn' was shown for similar tests, little is known on how test exposure affects future performance in cognitive tests of a different type.

Research suggests that experience acquired via one specific test is likely domain specific and hardly transferable to different tests (Souders et al., 2017; Tia, Viaro, & Fadiga, 2018; van Horik & Madden, 2016). For instance, tool-use training in monkeys did improve performance in the physical cognition domain, but not their ability to generalise this tool-use knowledge to a novel tool (Tia et al., 2018). Similar results were found in humans where participants slightly improved their performance in the cognitive test most similar to the test learned, but general transfer from one test to a conceptually different test was poor (Souders et al., 2017). However, it is possible that the experience of short testing periods over 1–2 months in the previously mentioned studies (Souders et al., 2017; Tia et al., 2018; van Horik & Madden, 2016) have not been sufficient to induce biologically meaningful differences in test performance in subsequent tests. For example, highly trained dogs, such as competitive-level sport dogs or certified working dogs, were shown to be more persistent and successful in solving a problem than non-trained dogs (Marshall-Pescini, Frazzi, & Valsecchi, 2016; Marshall-Pescini, Valsecchi, Petak, Accorsi, & Previde, 2008; B. Osthaus, Lea, & Slater, 2003). However, other factors associated with the dog's environment such as rearing history might have caused changes beyond test experience.

Furthermore, cognitive testing is linked to more general changes in behaviour and physiology that could indirectly affect the performance in subsequent cognitive tests. Langbein et al. (Jan Langbein et al., 2004) found that offering cognitive challenges via a computer-based learning device to goats induced changes in vagal activity of the heart during successful learning, suggesting that operating the learning device was experienced as positive stress by those animals. The use of a 'call feeding station' to cognitively challenge pigs over a 12-week period not only led to an increase in locomotor behaviour and a decrease in belly nosing in the pigs' home environment, but also altered behaviour in subsequent open-field and novel-object tests. The cognitively challenged pigs showed reduced activity and excitement in these tests compared with control pigs that were not administered the cognitive challenge (Puppe et al., 2007). The effect was more pronounced after 12 weeks of access to the call feeding station than after 6 weeks. Previous experiences with cognitive tests might therefore alter the motivation to participate and the performance in future cognitive tests by reducing neophobia and/or stress levels.

In addition to previous experiences with cognitive testing itself, habituation to humans and to isolation in a test environment are important non-cognitive factors that can potentially cause relevant differences in test performance between habituated and naïve individuals, because they also affect the motivation to participate in tests and the stress level in the test situation. In contrast to studies using an automated reward delivery (Jacobson et al., 2019b; Jan Langbein et al., 2004; Puppe et al., 2007), other cognitive test paradigms, such as object-choice tests, often require human–animal interaction, for example for positive reinforcement by a human experimenter. This interaction may be stressful for animals if they have not been habituated or if they had negative experiences with humans in the past. For example, high emotional reactivity towards the experimenter was found to affect learning performance of pigs (Brajon et al., 2016). In addition, separation from the group is stressful for most social animals and has been found to increase vocalisations, heart rate and cortisol levels (Aschwanden et al., 2008a; Boissy & Le Neindre, 1997; Da Costa et al., 2004; Patt et al., 2013; Siebert et al., 2011). A lack of sufficient habituation to humans and isolation can thus hamper a correct and reliable assessment of cognitive performance by a decrease in attention and participation (Regolin et al., 1995; Washburn & Rumbaugh, 1991).

Goats are a promising model species for cognitive research. By means of object-choice tests, they have been found to follow human pointing gestures to locate hidden food rewards and anticipate the trajectory of hidden objects (Nawroth et al., 2015, 2014). They have been shown to be good problem-solvers in visual discrimination and spatial detour tests (J. Langbein et al., 2008; Jan Langbein et al., 2004; Nawroth, Baciadonna, et al., 2016; Raoult et al., 2021), but also in instrumental problem-solving tests that involved opening a box or container (Briefer, Haque, Baciadonna, & McElligott, 2014; Finkemeier et al., 2018; Nawroth, Brett, et al., 2016). In addition, domestic goats have been shown to interact with humans in complex ways (Nawroth, 2017). Thus, object-choice tests as well as detour and problem-solving tests have been shown to be suitable tests to assess cognitive capacities in goats.

In the current study, we investigated whether long-term exposure to object-choice tests affects the performance of goats in subsequent conceptually different cognitive tests. To control for habituation to humans and to isolation in a test environment, we conditioned three treatment groups: goats with long-term exposure to human-presented object-choice tests (COG treatment), goats that were isolated as for the test exposure but received a reward from the experimenter without being administered the object-choice tests (POS treatment), and goats that were isolated but neither received a reward nor were administered the tests (ISO treatment). All treatment groups were subsequently tested in two conceptually different cognitive tests, namely a spatial A-not-B detour test (ABT) and an instrumental problem-solving test (PST). The ABT requires animals to learn to detour around a spatial barrier before (= A trials) and after (= B trials) the position of the barrier is altered (Britta Osthaus, Marlow, & Ducat, 2010; Britta Osthaus et al., 2013). The PST in our study is an instrumental manipulation test that requires the animal to open a familiar food container covered with a lid novel to the animal. To increase the heterogeneity of our sample and thus the external validity of our findings, we tested each of two different selection lines of goats (dairy goats and dwarf goats) at the same two research sites (Voelkl et al., 2018; Würbel, 2017).

We hypothesised that previous cognitive test exposure improves performance in an ABT and a PST, which measure behavioural flexibility and problem-solving abilities, respectively. We thus expected the COG group to perform better than the POS group in the tests. We furthermore hypothesised that positive human–animal interaction improves performance in these tests. Correspondingly, we expected the POS group to outperform the ISO group.

Material and Methods

Location, animals and housing conditions

To increase external validity (Voelkl et al., 2020, 2018; Würbel, 2017), the study was carried out at two locations, at the Centre for Proper Housing of Ruminants and Pigs at Agroscope in Ettenhausen (ET), Switzerland, and at the Leibniz Institute for Farm Animal Biology in Dummerstorf (DU), Germany, and with two selection lines of goats, namely Nigerian dwarf goats (61 non-lactating female goats) and dairy goats (59 non-lactating female goats). The Nigerian Dwarf goat is commonly kept as pet and zoo animal in Europe and not selected for productivity traits. We used dwarf goats bred at the research institute in DU. The only selection aim in this population was to avoid inbreeding. The potential milk yield of dwarf goats does likely not exceed 0.3 kg per day (Akinsoyinu et al., 1977). As it was common practice in DU, dwarf goat kids stayed with their dams for six weeks before they were weaned. Additionally, we used three of the most common high-producing dairy breeds in Switzerland and Germany, and their crossbreeds, namely Saanen ($n = 15$), Chamois Coloured ($n = 12$), Saanen × Chamois ($n = 3$) and Deutsche Edelziege ($n = 29$). These animals had a potential milk yield of up to 3 kg per day (Vacca et al., 2018). In accordance with common practice in the dairy goat industry, the dairy goat kids had been separated from their dam shortly after birth and were artificially raised.

At the Agroscope research station in ET, we housed 30 dwarf goats and 30 dairy goats (15 Saanen, 12 Chamois Coloured, 3 Saanen × Chamois crossbreeds). The dwarf goats were born between January and February 2017 in DU, Germany, and moved to ET in June 2017. The dairy goats were born between February and April 2017 on different Swiss farms and were moved to ET in June/July 2017. At the location in DU, we housed 31 dwarf goats (Nigerian Dwarf) and 29 dairy goats (Deutsche Edelziege). The dwarf goats were born between January and March 2018 in DU, except for eight animals. These were bought from the Zoo Osnabrück and the Wildpark Lüneburger Heide, Germany, due to a shortage of female animals in the facility's own breeding stock. All dairy goats were born on the same German farm in March 2018 (Gleistal-Mutterkuhhaltungs GmbH, Golmsdorf) and were moved to DU in July 2018.

All goats were moved to pens of 9–11 goats at the age of 7–8 months: three pens of dairy goats and three pens of dwarf goats at each location. The total area of each dwarf goat pen was 14 m^2 (approximately 3.6 m × 3.9 m), consisting of a deep-bedded straw area of 11 m^2 (approximately 2.8 m × 3.9 m) and a 0.5-m-elevated feeding place (1.4 m^2). The total area of each dairy goat pen was 17.7 m^2 (approximately 3.9 m × 4.55 m) consisting of a deep-bedded straw area of 13.4 m^2 (approximately

4.55 m × 2.95 m) and a 0.65-m-elevated feeding place (1.82 m^2). Hay was provided behind a feeding fence at the feeding place twice a day at around 8 a.m. and 4 p.m. in ET and at around 7 a.m. and 1 p.m. in DU. Each pen had one watering place and a mineral supply. Additional structures in the straw-bedded area included a wooden bench (for dairy: 2.4 m long, 0.6 m high, 0.62 m wide; for dwarf: 2.3 m long, 0.5 m high, 0.5 m wide) along the wall of the pen and a round wooden table (0.8 m high, 1.1 m in diameter) in the centre of the pen.

All animal care and experimental procedures were performed in accordance with all relevant Swiss legislative and regulatory requirements as well as the German welfare requirements for farm animals and the ASAB/ABS Guidelines for the Use of Animals in Research (ASAB & ABS, 2018). All procedures involving animal handling and treatment were approved by the Cantonal Veterinary Office, Thurgau, Switzerland (Approval No. TG04/17 – 29343) and the Committee for Animal Use and Care of the Ministry of Agriculture, Environment and Consumer Protection of the federal state of Mecklenburg-Vorpommern, Germany (Approval No. 7221.3-1.1-062/17).

Treatment groups and procedures

Three goats from each of the 12 pens (à 9–11 goats) were pseudo-randomly assigned to one of the three treatment groups: COG (*n* = 36), POS (*n* = 36) and ISO (*n* = 36). Except one pen, all pens housed 1–2 extra goats not assigned to a treatment group to replace others in case of e.g. disease or injury. In 44 test sessions, distributed over a period of 4–5 months, the COG group was exposed to cognitive tests in the form of discrimination and reversal learning tests and a cognitive test battery (see Appendix II: Table S1 for more details). During these tests, COG goats received food rewards from the experimenter for correct responses. The POS group was not exposed to cognitive tests but received a similar number of rewards as the individuals in the COG group (= median number of rewards received by COG group in the previous test session), provided by the experimenter in the test arena at pseudo-random times and over a similar period of being isolated as the COG group (= median time taken by COG group to finish all trials in the previous test session). Contrasting COG versus POS allows investigating the effect of the cognitive testing itself, disentangled from the effects of the positive association with the human and the isolation from the group during testing. Individuals allocated to the ISO treatment neither participated in cognitive tests nor did they receive rewards by the experimenters in the test arena. Instead, they were isolated over a similar period as the COG and the POS group in the same arena (= median time taken by COG group to finish all trials in the previous test session). Contrasting POS versus ISO allows investigating the effect of the positive association with the human, disentangled from the effect of isolation from the group during testing. To control for caloric intake, ISO animals received a similar amount of food as the COG and POS goats, but to avoid positive association with the human, the food was provided scattered over the floor of the waiting room where goats were kept before they were individually isolated in the test arena.

All goats in ET were between 15 and 18 months old when tested in the ABT (mean ± SD: dairy goats: 494 ± 3 d, dwarf goats: 537 ± 1 d) and in the PST (mean ± SD: dairy goats: 491 ± 3 d, dwarf goats: 533 ± 1 d). In DU, the goats were between 19 and 20 months old in the ABT (mean ± SD: dairy goats: 586 ± 0 d, dwarf goats: 616 ± 3 d) and in the PST (mean ± SD: dairy goats: 579 ± 0 d, dwarf goats: 608 ± 3 d).

A-not-B detour test

The ABT according to Osthaus et al. (Britta Osthaus et al., 2010) is a test that requires animals to learn to detour around a spatial barrier before (= A trials) and after (= B trials) the placement of the barrier is altered. Thus, the animal is required to suppress the once successful response (in A trials) and adapt to the new spatial setup (B trials). Due to the nature of the ABT, the animals are expected to show a spatial perseveration error in the B trials following a variable number of A trials with the degree of error depending on their level of impulse control. It is assumed that animals that more strongly suppress their previously learned response make fewer errors.

Test setup and procedure

The experiment was conducted in a large rectangular arena, which was divided by a movable fence into two same-sized compartments while leaving a gap on one or the other side of the arena (Fig. 10). The size of the arena differed slightly in ET and DU, but the angle between test subject and gap location was kept the same (approximately 15.2°) by adjusting the fence length accordingly (see Fig. 10). In ET, a mobile pen, familiar to the goats, was used to take the tested individuals in groups of 4–10 to the test arena. The mobile pen served as a holding area next to the test pen while individual goats were being tested. In DU, the test arena was much closer to the home pens of the goats, hence the goats were led individually to the test arena. In the arena, the goat was taken to the opposite side of the arena and restrained by Experimenter 1 (E1; Fig. 10 and Appendix III: Fig. S1). A second experimenter (E2) stayed at the side of the entrance, shaking a container, familiar to the goat, with dry pasta to motivate the goat to walk through the gap towards the food. The trial started when E1 released the goat. Each goat received eight trials: during the so-called ‘acquisition phase’, the goat received four trials with the gap on one side of the arena (= Trials A1–A4). Subsequently, the fence was moved, and the goat received four trials with the gap on the other side of the arena (= Trials B1–B4). If the tested goat moved through the gap, it was allowed to feed for approximately five seconds from the container before it was led back to the starting point through the same gap. Between the A and B trials, E2 switched the position of the gap, while E1 was holding the tested goat and covered its eyes. For half of the goats, the gap was first on the left side and for the other half it was first on the right side of the arena.

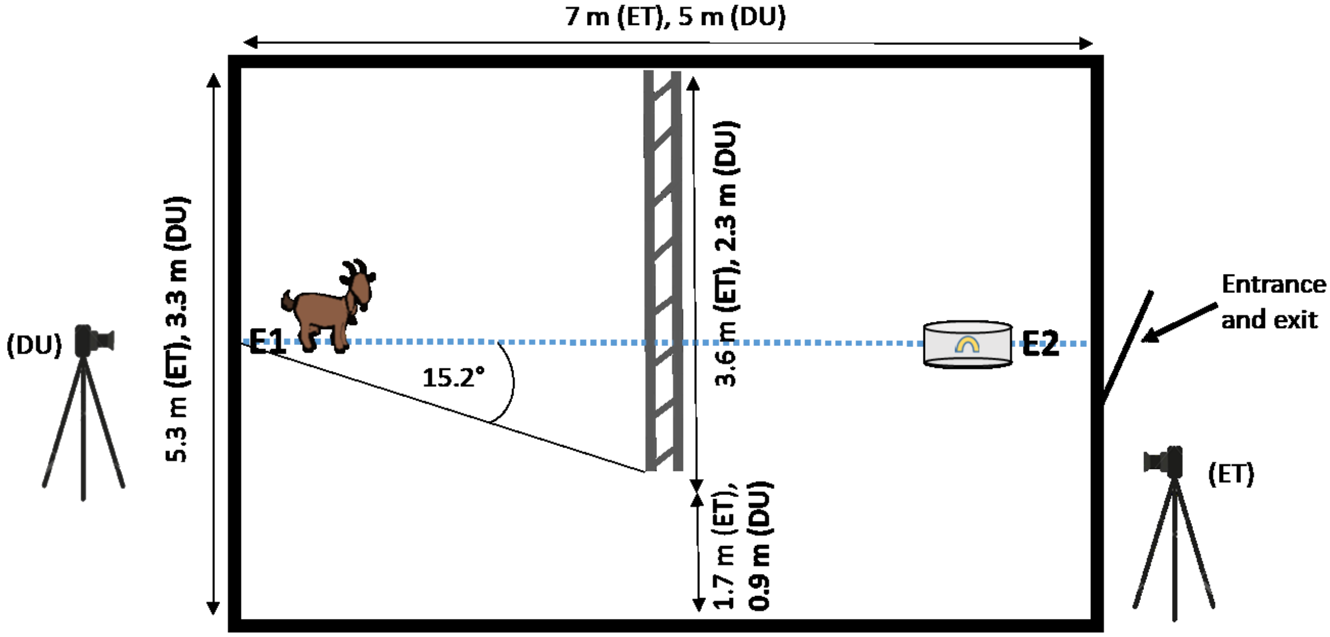


Figure 10: Schematic drawing of the test arena used for the A-not-B detour test in Ettenhausen (ET) and Dummerstorf (DU). Positions of the two experimenters are indicated with E1 and E2. The blue dotted line indicates the imaginary line that was used to determine whether a goat stepped towards the correct or the incorrect direction (= Accuracy).

Behavioural measures

Two parameters of performance were determined: 1.) *Accuracy*, a binary response being either correct (1) or incorrect (0) depending on whether the first step over an imaginary line (Fig. 10) was or was not directed towards the side of the arena where the gap was located and 2.) *Latency,* defined as the time (in seconds) from the first step until passing the gap in the fence with the shoulders. If a goat did not pass the gap in less than 60 seconds after it had been released, the trial was coded as *not available* and the animal was taken back to the starting point and the next trial began.

Problem-solving test

The setup and procedure of the PST are less standardised than the ABT and vary considerably between different studies. Experimental PST setups differ for example in the type of device used, such as containers, puzzle boxes, and tubes, as well as in the required opening techniques, including pulling a stick, sliding a lid, moving a lever, and rotating a lid (Benson-Amram & Holekamp, 2012; Cole & Quinn, 2012; Sol, Griffin, & Bartomeus, 2012; Thornton & Samson, 2012). The PST test in our study is an instrumental manipulation test that requires the animal to open a familiar food container covered with a lid novel to the animal.

Habituation phase

A round plastic container (11.5 cm tall, 34 cm in diameter) with a light wooden lid (38 cm in diameter, 0.33 kg) was used as manipulandum for the PST. The goats were already familiar with this container and the notion that it contains food from their general handling routines. During the habituation phase, the container was open with the lid leaned against the container. The goats were allowed to investigate the container and the lid, first in pairs, whereby the container was baited 10 times with one piece of pasta per goat. Once they fed out of the container repeatedly, they were allowed to investigate it individually until they had reliably eaten the reward out of the container in five consecutive trials.

Test setup and procedure

For the test phase, the container was covered with the wooden lid such that physical manipulation was necessary to open the container and access the food reward. Before the start of the test phase, each goat was individually led into the test arena (Fig. 11) and allowed to eat a piece of pasta from the open container. One experimenter (E1) then restrained the goat on one side of the test arena, while the other experimenter (E2) placed a piece of pasta into the container and covered it with the lid. Subsequently, the goat was released and allowed to approach the container. Each goat received five consecutive trials within a single day. A motivation trial with the lid open was administered between the test trials to maintain motivation.

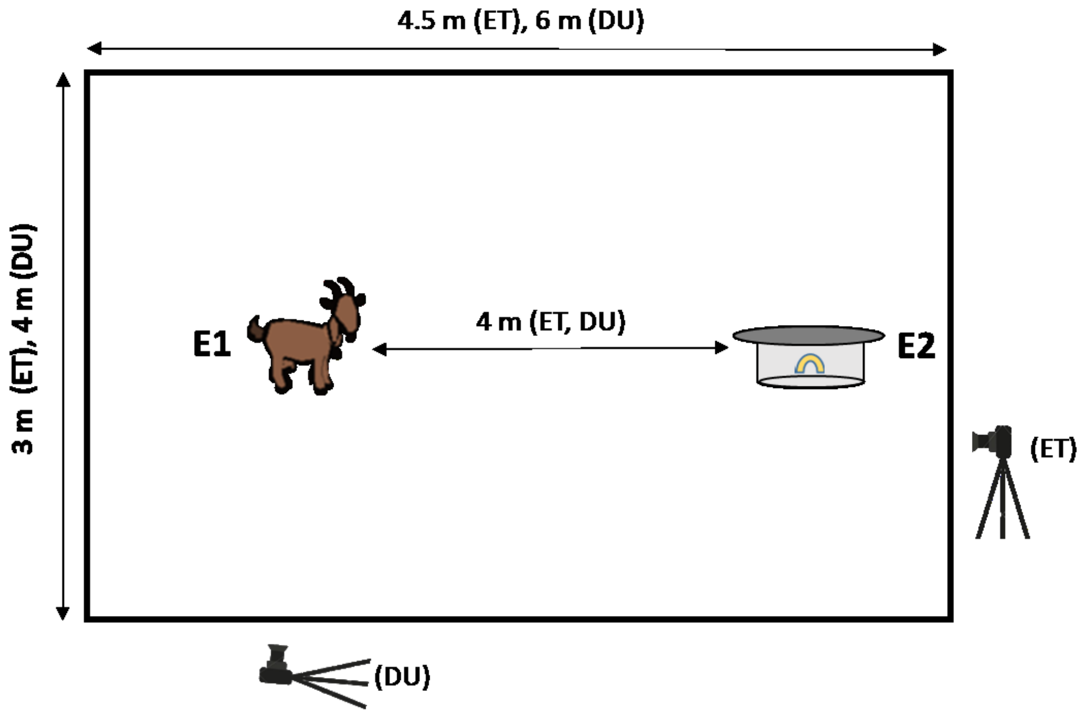


Figure 11: Schematic drawing of the test arena used for the problem-soling test in Ettenhausen (ET) and Dummerstorf (DU). Positions of the two experimenters are indicated with E1 and E2.

Behavioural measures

As measures of motivation to engage with the test, we recorded 1.) if the goat touched the lid or not (with any body part, = *Touched*) and 2.) the latency to touch (= *LatencyT*) measured from release to first touch of the lid (with any body part). As measures of proficiency, we recorded 1.) if the goat opened the container or not (= *Opened*) and 2.) the latency to open (= *LatencyO*) defined as the time from first touch with any body part to opening the lid. If a goat did not approach the container within a 1-m radius within 30 seconds after it had been released, the trial was coded as *not available* and the animal was taken back to the start for the next trial.

Data coding and statistical analysis

All trials were videotaped with a camcorder (ET: Sony HDR-CX240E; DU: Panasonic HDC-SD60), and most parameters were additionally recorded live by the experimenters. For the analysis of the ABT, we used the live-recorded data. For the PST, we decided in retrospect to code additional variables from video and therefore used the more comprehensive video-coded data for analysis.

To perform the reliability analysis for the ABT, an external person not familiar with the hypotheses coded 50% of the trials from videos. Inter-observer reliability was found to be very high (*Accuracy*: $\kappa = 0.966$, $z = 14.2$, $p < 0.001$; *Latency*: $r = 0.99$, $p < 0.001$). For the PST, only the variables *Opened* and *LatencyO* were coded live, thus we used these two measures to analyse inter-observer reliability and compared them with the corresponding data coded by an external person from video. Again, inter-observer reliability was found to be very high (*Opened*: $\kappa = 0.903$, $z = 10.4$, $p < 0.001$; *LatencyO*: $r = 0.98$, $p < 0.001$).

All statistical analyses were performed in R v4.0.3 (R Core Team, 2020). In the ABT and PST, seven dwarf goats and one dairy goat did not participate because they were too stressed during training for the preceding object-choice tests and thus could not receive their assigned treatment. In the ABT, we excluded nine animals because they jumped over or crawled under the fence, or due to latencies above 60 seconds. In addition, goats that showed an incorrect response in A3 and A4 trials were excluded from the analysis ($n = 19$) because they likely had not learnt the correct response by the end of the A trials and were thus, per definition, not able to show the perseveration error in the subsequent B trials. In the PST, we excluded 29 individuals during the training trials before the actual PST test because they did not meet the training criteria (= reliably eating the reward out of the open container within five trials) and four more animals that had missing trials due to human errors or technical failure. In sum, statistical analysis was performed on data of 28 COG, 22 POS and 27 ISO goats in the ABT and 23 COG, 24 POS and 20 ISO goats in the PST.

To test for the perseveration error in the ABT, we applied one-sided paired McNemar tests (*exact2x2* function from the R package exact2x2 (M. P. Fay, 2010)) and one-sided paired Wilcoxon signed-rank tests (base R function *wilcox.test*). McNemar tests were used to test whether B1 trials were less frequently correct than the corresponding A4 trials (binary variable *Accuracy*). Similarly, Wilcoxon signed-rank tests were used to test whether B1 trials were characterised by longer *Latency* as compared with the A4 trials.

To analyse the effects of the treatments (COG, POS, ISO) on the dependent variables in the ABT and the PST, we employed linear mixed-effects models from the R package lme4 (Bates, Mächler, et al., 2015). For binary responses, i.e. *Accuracy* in the ABT and *Touched* and *Opened* in the PST, the models were estimated as generalised linear mixed model (GLMM) with logit link by using the *glmer* function. For the continuous responses, i.e. *Latency* to cross the fence in the ABT and *LatencyT* and *LatencyO* in the PST, the models were estimated as ordinary linear mixed model by using the *lmer* function. We visually inspected residuals of all models by using the package DHARMa (Hartig, 2020). To achieve better normal distribution of residuals, we log2-transformed the latencies before model fitting.

For the ABT model, formulas in lme4 syntax were as follows:

response ~ 0 + Type:Treatment + Type:Treatment:I(Trial-1) + (1|SelectionLine) + (1|Site/Pen/Individual)

We included an intercept for each type of trial (= A and B trials) and treatment interaction individually (0 + Type:Treatment) and a slope for trial number for each type–treatment interaction (Type:Treatment:I[Trial-1]). The trial number was included as Trial-1 to render the intercept to correspond to Trial 1 instead of the non-meaningful Trial 0. Besides these fixed effects, a random intercept for selection line (1|SelectionLine) and for individual nested within the pen and within site (1|Site/Pen/Individual) was included to account for repeated testing and a potential effect of the affiliation to pen (A–F, U–Z) and site (ET, DU).

For the PST, the model formulas in lme4 syntax were as follows:

response ~ 0 + Treatment + Treatment:I(Trial-1) + (1|SelectionLine) + (1|Site/Pen/Individual)

Here, we included an intercept for each treatment (0 + Treatment) and a slope for trial for each treatment (Treatment:I[Trial-1]). Besides these fixed effects, a random intercept for individual nested within the pen and within site (1|Site/Pen/Individual) was included to account for repeated testing and potential effects of pen and site affiliation. Only dairy goats opened the container and could therefore be analysed with respect to the corresponding behavioural responses. In contrast to all other models, the models for *Opened* and *LatencyO* have thus no random intercept for selection line. To investigate differences in all behavioural responses between the treatments, we tested treatment contrasts for the fixed effects with the *glht* function from R package multcomp (Hothorn et al., 2008). The *p*-values for fixed-effect estimates and for the contrasts were obtained by using Wald *z*-tests (*summary.ghlt* function, multcomp package).

Results

A-not-B detour test (ABT)

In the ABT, we recorded *Accuracy* (Fig. 12) and *Latency* to cross the fence (Fig. 13). For both behavioural measures, the individual variability in B trials was large and no consistent patterns were apparent for treatment groups, selection lines, or for sites. A spatial perseveration error was indicated for all treatment groups: For B1 trials as compared with A4 trials, the McNemar test suggested a lower proportion of correct responses (COG: $p < 0.001$, POS: $p = 0.002$, ISO: $p < 0.001$) and the Wilcoxon signed-rank test a higher latency (COG: $p = 0.006$, POS: $p < 0.001$, ISO: $p = 0.009$).

According to the GLMM results, treatments did not differ in their *Accuracy*, neither in A (POS-COG: $p = 0.83$, ISO-COG: $p = 0.87$, ISO-POS: $p = 0.96$; Appendix III: Table S2) nor in B trials (POS-COG: $p = 0.13$, ISO-COG: $p = 0.19$, ISO-POS: $p = 0.81$; Appendix III: Table S2). Consistent across the three treatment groups, the probability for the goats to choose the correct side increased with increasing number of B trials, but with high statistical certainty only for the COG treatment, which showed the steepest increase in log odds over B trials (Treatment:I[Trial-1]: COG: est. = 0.73, $p < 0.001$, POS: est. = 0.41, $p = 0.07$, ISO: est. = 0.38, $p = 0.07$; Appendix III: Table S2). The estimated variance components for the random effects (Appendix III: Table S3) indicated large deviances among *Individuals* (SD = 0.76), among *Selection lines* (SD = 0.35) and among *Sites* (SD = 0.20). These deviances were within the range of the absolute values of the (statistically not supported) treatment contrasts (|est.| ≤ 0.87; Appendix III: Table S2).

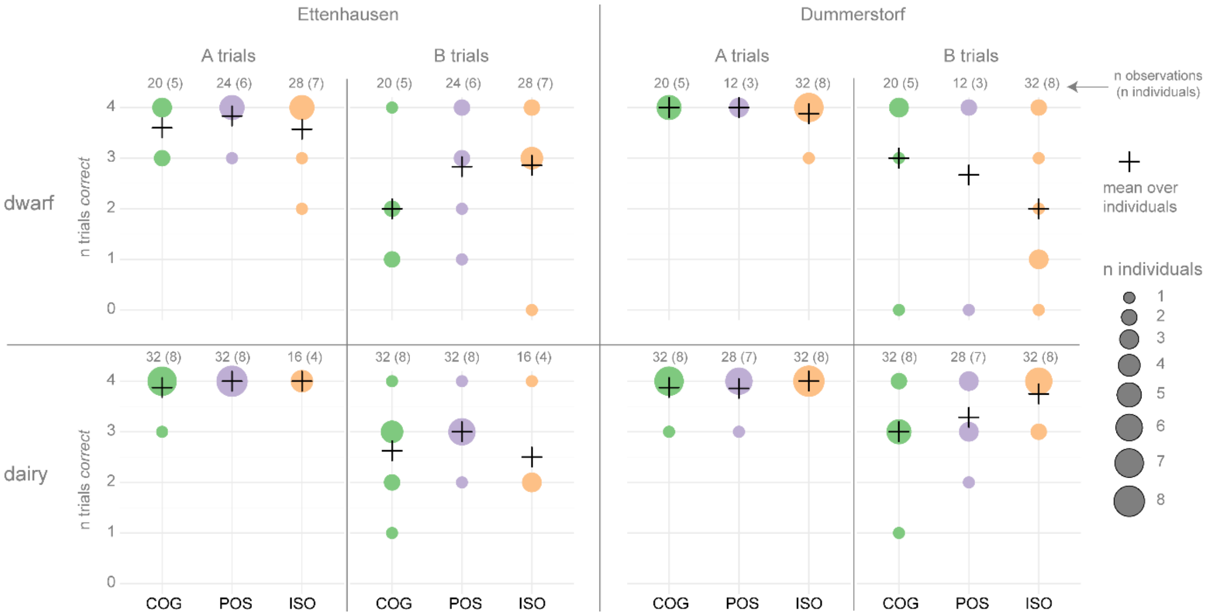


Figure 12: Accuracy shown as the number of correct trials for dwarf (top panels) and dairy goats (bottom panels) of the COG, POS, and ISO treatment groups during A trials and B trials in Ettenhausen (four left panels) and Dummerstorf (four right panels). The size of circles indicates the number of animals. As in the statistical analysis, goats that showed an incorrect response in A3 and A4 trials are excluded from this figure.

The linear mixed model for *Latency* to cross the fence did also not statistically detect treatment differences in A (POS-COG: *p* = 0.80, ISO-COG: *p* = 0.84, ISO-POS: *p* = 0.96; Appendix III: Table S4) or B trials (POS-COG: *p* = 0.19, ISO-COG: *p* = 0.67, ISO-POS: *p* = 0.37; Appendix III: Table S4). The estimation of the variance components indicated deviances among *Individuals* (SD = 0.59) and among *Residuals* (SD = 1.07; Appendix III: Table S5) that were larger than the absolute values of the (statistically not supported) treatment contrasts (|est.| ≤ 0.4; Appendix III: Table S4).

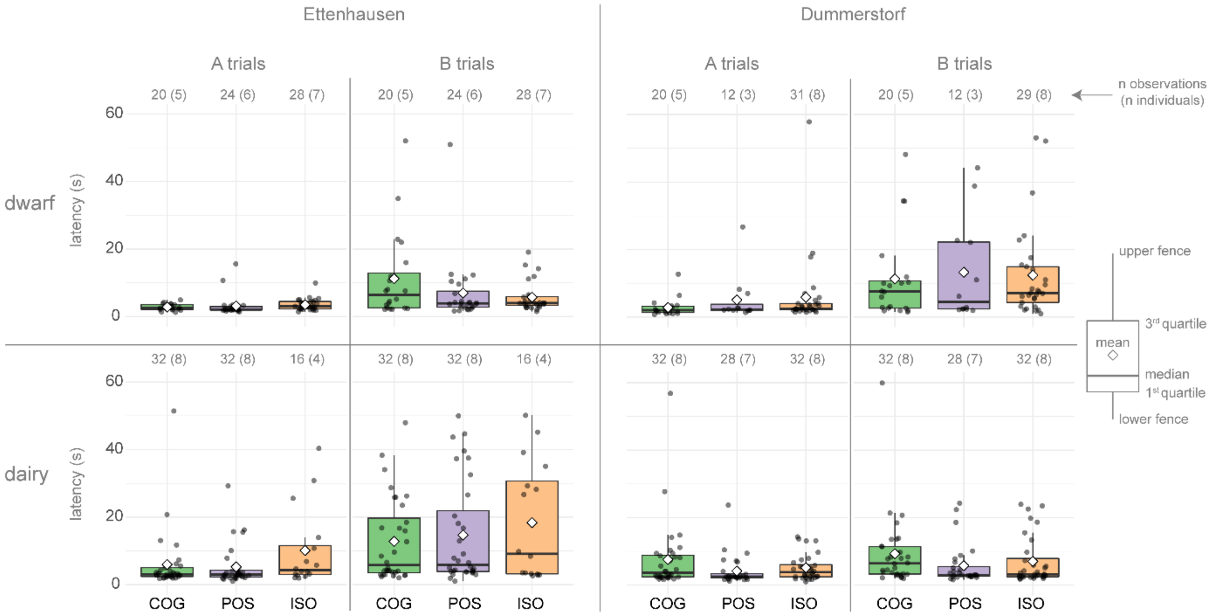


Figure 13: Latency to cross the fence for dwarf (top panels) and dairy goats (bottom panels) of the COG, POS and ISO treatment groups in Ettenhausen (four left panels) and Dummerstorf (four right panels). Distribution of individual latencies (jittered points) are summarised as boxplots. As in the statistical analysis, goats that showed an incorrect response in A3 and A4 trials are excluded from this figure.

Problem-solving test (PST)

In the PST, we recorded the occurrence of interactions with the lid as variables *Touched* and *Opened* (Fig. 14) and the corresponding latencies as *LatencyT* and *LatencyO* (Fig. 15). In all these behavioural measures, no consistent patterns with respect to the treatment groups were apparent, but the goats differed largely between selection lines and sites. The average number of touches of the treatment groups ranged from 0 to 2.5 for dwarf goats and from 3.0 to 4.4 for dairy goats depending on selection lines and site (Fig. 14a). In DU, dwarf goats from the ISO group never touched the container. For both selection lines, the average number of touches for each treatment group was consistently higher in ET than in DU. At both sites, none of the dwarf goats opened the container (data not shown in figures). The average number of trials for each treatment group in which a dairy goat opened the container was always smaller than 0.9 in DU, whereas it ranged from 3.2 to 3.9 in ET.

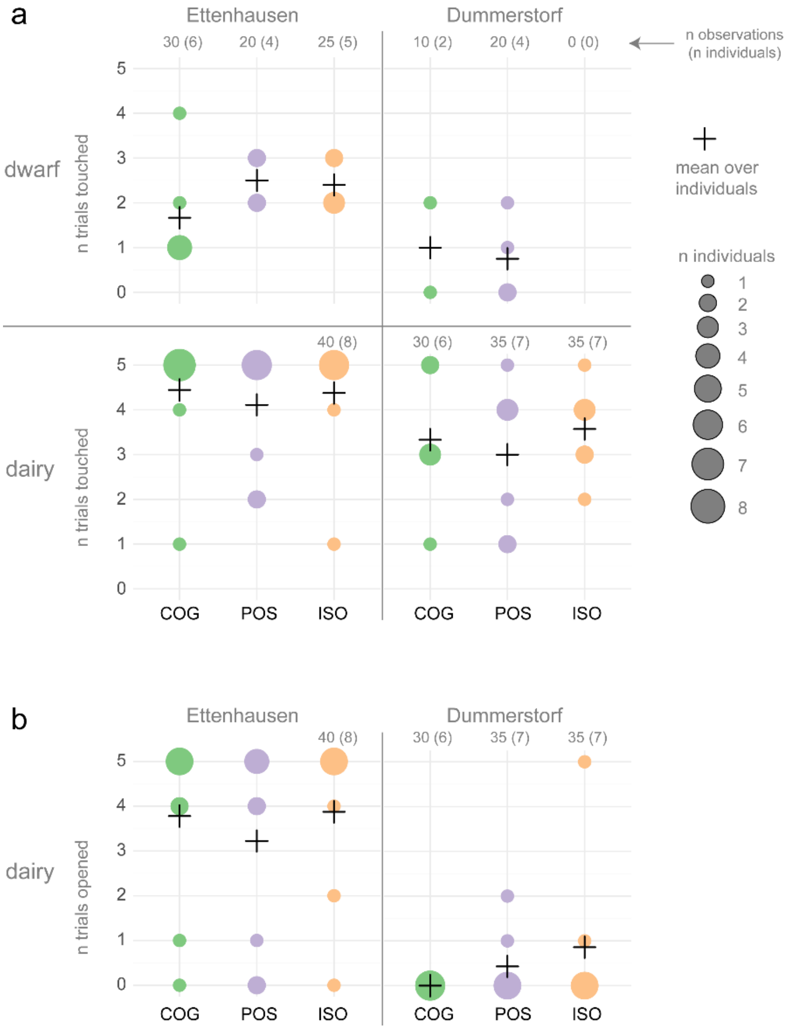


Figure 14: Number of trials in which dwarf (top panels) and dairy goats (bottom panels) of the COG, POS and ISO treatment groups a) touched and b) opened the container in Ettenhausen (left) and Dummerstorf (right). The cross indicates the mean, and the size of circles indicates the number of animals that chose the correct side. If not indicated otherwise, observations (n = 45) and animals (n = 9) were complete.

Consistent for all treatment groups, in ET the average *LatencyT* was around five seconds in dairy goats and 10 seconds or more in dwarf goats (Fig. 15a). Because dwarf goats in DU almost never touched the lid, corresponding latencies could be measured on only a few occasions. Only the average *LatencyT* for the POS dairy goats in DU was similar to the *LatencyT* of the corresponding treatment group in dairy goats in ET (Fig. 15a). In ET, shortly after the dairy goats touched the lid, most of them also opened it (*LatencyO*, Fig. 15b). In DU, the few dairy goats that opened the lid had on average a similar *LatencyO* as the dairy goats in ET.

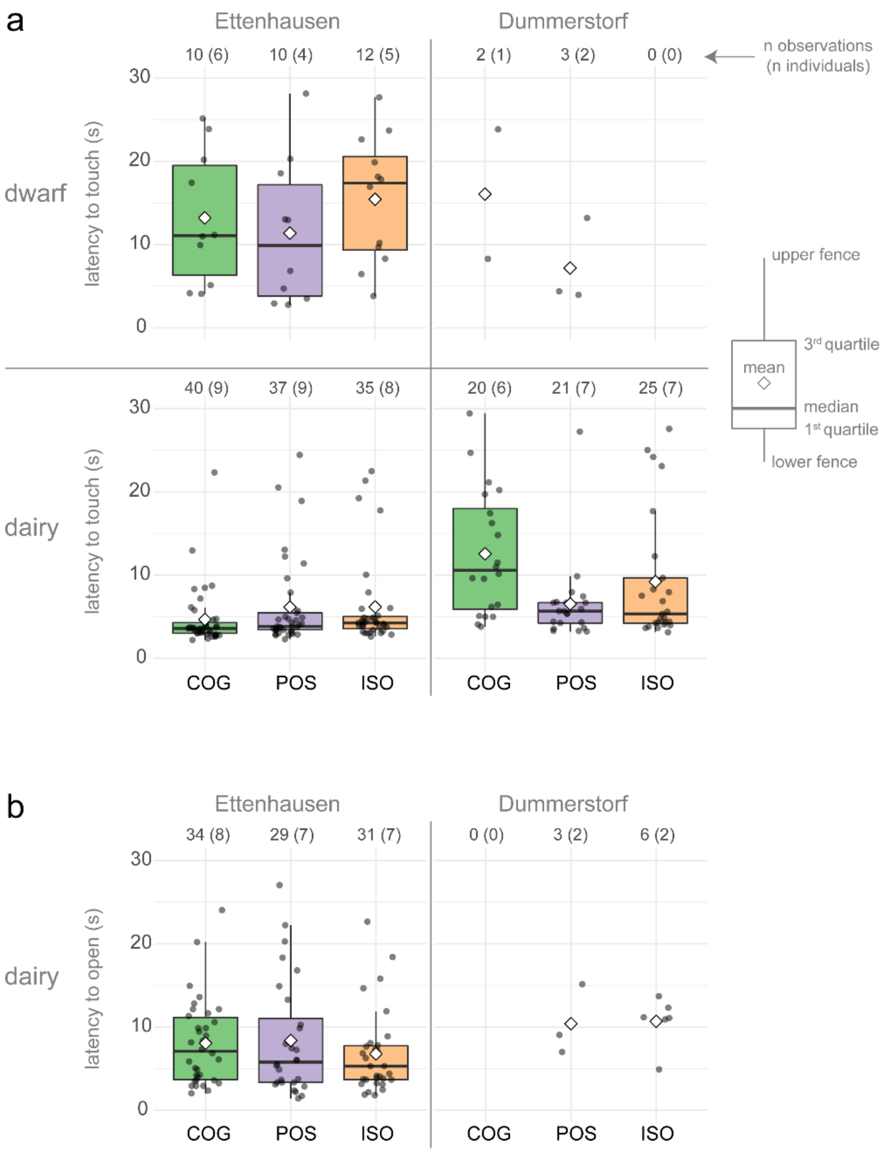


Figure 15: a) LatencyT to touch the container in dwarf (top panels) and dairy goats (bottom panels) and b) LatencyO to open the container in dairy goats in Ettenhausen (left) and Dummerstorf (right) from the COG, POS and ISO treatment groups. Distribution of individual latencies (jittered points) are summarised as boxplots.

In the GLMM for *Touched*, no effect of treatment on the probability of the animals to touch the container was detected (POS-COG: $p = 0.37$, ISO-COG: $p = 0.23$, ISO-POS: $p = 0.71$; Appendix III: Table S6). The probability for *Touched* decreased over trials in all groups, but with varying statistical certainty (Treatment:I[Trial-1]: COG: est. = −0.28, $p = 0.12$, POS: est. = −0.70, $p < 0.001$, ISO: est. = −0.63, $p = 0.003$; Appendix III: Table S6). Estimated variance components indicated deviances among *Selection lines* (SD = 1.51), among *Sites (*SD = 0.92) and among *Individuals* (SD = 1.10; Appendix III: Table S7) that were similar to or larger than the absolute values of the (statistically not supported) treatment contrasts (|est.| ≤ 0.98; Appendix III: Table S6). Similarly, in the GLMM for *Opened*, no effect of treatment on the probability of the goats to open the container was detected (POS-COG: $p = 0.65$, ISO-COG: $p = 0.58$, ISO-POS: $p = 0.33$; Appendix III: Table S8). The estimated variance components indicated deviances among *Individuals* (SD = 3.97) and among *Sites* (SD = 4.28; Appendix III: Table S9) that were larger than the absolute values of the (statistically not supported) treatment contrast (|est.| ≤ 2.35; Appendix III: Table S8).

The linear mixed model did not statistically detect differences in *LatencyT* between the COG, POS and ISO treatments (POS-COG: *p* = 0.18, ISO-COG: *p* = 0.93, ISO-POS: *p* = 0.21; Appendix III: Table S10). The model estimated small increases in *LatencyT* of 7–15% (0.1–0.2 log2 units) with every additional trial for all treatments (Treatment:I[Trial-1]: COG: est. = 0.12, *p* = 0.05, POS: est. = 0.21, *p* = 0.001, ISO: est. = 0.16, *p* = 0.01; Appendix III: Table S10). The estimation of variance components suggested deviances among *Selection lines* (SD = 0.74), among *Individuals* (SD = 0.48) and among *Residuals* (SD = 0.75; Appendix III: Table S11) that were larger than the absolute values of the (statistically not supported) treatment contrasts (|est.| ≤ 0.34; Appendix III: Table S10). For *LatencyO* in dairy goats, the linear mixed model suggested longer latency to open the lid in the POS versus the ISO treatment (POS-COG: *p* = 0.32, ISO-COG: *p* = 0.22, ISO-POS: *p* = 0.03; Appendix III: Table S12). With increasing trial number, *LatencyO* was estimated to decrease by 15–41% (0.2–0.5 log2 units) in all treatments (Treatment:I[Trial-1]: COG: est. = −0.35, *p* < 0.001, POS: est. = −0.51, *p* < 0.001, ISO: est. = −0.24, *p* = 0.02; Appendix III: Table S12). Estimated variance components indicated deviances among *Sites* (SD = 0.55) and among *Residuals* (SD = 0.84; Table S13) of similar magnitude as the absolute values of the estimates for the treatment contrasts (|est.| = 0.84; Appendix III: Table S12).

Discussion

We investigated whether participation in cognitive tests (here: visual discrimination and reversal learning as well as a cognitive test battery consisting of object-choice tests) over a period of 4–5 months affects the performance of goats in subsequent conceptually different cognitive tests, namely a spatial ABT and an instrumental PST. By comparing three treatment groups (COG, POS, ISO), we aimed to disentangle possible effects of the preceding cognitive test exposure (COG-POS) from effects of positive human–animal interactions (POS-ISO). We found only subtle differences between our treatment groups in terms of their behavioural flexibility in the ABT or their performance in the PST. Our results are thus in line with previous research suggesting that cognitive test exposure does not substantially alter performance in subsequent conceptually different tests (Müller, Riemer, Virányi, Huber, & Range, 2016; e.g. Tia et al., 2018; van Horik & Madden, 2016). In contrast to the treatments, individuals, selection lines, and sites accounted for large amounts of the variation in our data (see Appendix III: Tables S3, S5, S7, S9, S11, S13).

We hypothesised that cognitive training via object-choice tests does improve the performance in a subsequent spatial detour test (the ABT) and therefore expected the COG treatment group to be more behaviourally flexible in the ABT than the other treatment groups, i.e. to detour the fence more often correctly and faster. However, we did not find support for this hypothesis. All treatment groups were similarly affected by the switch of the gap and had longer latencies and fewer correct responses in B1 than in A4 trials, indicating a spatial perseveration error. The improvement of correct responses over trials was only statistically certain for COG goats. This finding may indicate that cognitive stimulation (COG treatment) affected the goats' ability to better inhibit their initially learned response in the B trials over time. The lack of a treatment effect between POS and ISO suggests that the additional positive human–animal interaction in the COG and POS groups had not substantially affected the goats' behaviour in the spatial ABT. As discussed in Langbein et al. (Jan Langbein, Krause, & Nawroth, 2018), maybe the daily standard handling during husbandry procedures had already altered the behaviour of the goats of all treatment groups, interfering with the effect of the additional human contact during the application of the treatments.

In the PST, we hypothesised that the COG treatment enhances goats' problem-solving ability and thus that these goats would touch and open a covered container more often and faster than POS goats. Our findings did not support this hypothesis because no effect of COG on any of the variables in the PST was found. This result is in contrast to studies on dogs which showed that high levels of training generally improved the dogs' problem-solving abilities and their probability to interact with novel objects (Marshall-Pescini et al., 2016, 2008; B. Osthaus et al., 2003). However, trained dogs often experience different management conditions than non-trained dogs, and they may show improved performance due to other reasons than cognitive test experience *per se*. In addition, we expected that POS goats also perform better than goats from the ISO treatment. However, POS dairy goats showed an even longer latency to open the lid than ISO dairy goats. In our study, the probability to touch the container over trials decreased, suggesting that the goats may have lost interest in the container with increasing trial number. Again, this development occurred similarly in all treatment groups. The motivation to explore and learn is a strong predictor of problem-solving success (Wat, Banks, & McArthur, 2020), but it may decrease with increasing exposure to the novel item (Gifford, Cloutier, & Newberry, 2007; Trickett, Guy, & Edwards, 2009), given the problem is not kept challenging enough by regular modification (Meehan & Mench, 2007). In sum, these results suggest that cognitive test exposure (COG) did not substantially affect the outcomes of the ABT and the PST, even though exposure to the cognitive testing was longer than in other studies (Tia et al., 2018; e.g. van Horik & Madden, 2016).

To increase the heterogeneity of our sample and thus the external validity of our findings, we tested two selection lines of goats (dairy goats and dwarf goats) at two research sites (Voelkl et al., 2020, 2018; Würbel, 2017). In the current study, we found a high variance of *Site* and *Selection line* which highlights the importance of effects of location as well as phenotypic variation on cognitive test results. We found site differences within dwarf goats although subjects from both sites originated from mainly the same population, indicating that other factors than genetics should be considered. Because we used almost the same experimental setup and the same human experimenters at both sites, this variation might for example be caused by different previous experiences with animal care staff at the two research sites (Boivin, Garel, Mante, & Le Neindre, 1998; Hausberger, Roche, Henry, & Visser, 2008), size of the farms where the goats were bred (Mattiello et al., 2010), rearing history, i.e. human- or dam-reared (Lyons, 1989), or isolation and handling at a young age (Boivin & Braastad, 1996).

Possible explanations for the phenotypic variations may be genetic and/or environmental factors. Van Horik and Madden (van Horik & Madden, 2016) found that inherent motivational traits, such as the motivation to approach an apparatus or the persistency to solve a test, best predict the success in a PST. These traits have often been linked to personality traits. Dogs described as active, playful and bold were found to be more likely to learn complex behaviours and perform well in situations requiring persistence (Svartberg, 2002). Similarly, bold and exploratory animals were found to be particularly likely to be innovative (Boogert, Reader, & Laland, 2006; Overington, Cauchard, Côté, & Lefebvre, 2011; Webster & Lefebvre, 2001) and to overcome novel challenges faster compared with less bold animals (Reader & Laland, 2003). Learning speed in reversal tests was also found to be correlated with individual variation in exploration, but this relationship differed between tests and age (Zidar et al., 2018). Interestingly, personality seems to affect learning performances especially if individuals are in a state of stress (Valenchon et al., 2013). Thus, the differences we found in the PST in dwarf goats compared with dairy goats may also be explained by other factors such as different levels of stress, which were found to be linked to the motivation to explore (Matzel et al., 2006).

Additionally, fear towards the experimenter may have affected performance in the PST (Brajon et al., 2016) by reducing the motivation of stressed individuals to approach the container positioned next to the experimenter. Although both selection lines were handled in a similar manner, dairy goats might have been more inclined to approach humans during training and test sessions compared with dwarf goats. Research in chicken and sheep suggests that selection for high productivity has reduced stress reactivity towards humans (C. Lindqvist & Jensen, 2008; Schütz & Jensen, 2001). If we apply this assumption to our study, it seems that the selection for high productivity in dairy goats may have decreased fear towards humans and as a result also increased the goats' inclination to approach and manipulate the container positioned next to the experimenter. As it is common for the dairy industry, the dairy goats used in this study had been separated from their dam right after birth. In contrast, dwarf goats had been allowed to stay with their mothers for six weeks. Early separation from the dam and rearing by humans has been shown to increase tameness scores in goats (Lyons et al., 1988).

When comparing the variances in performances of the two tests applied in this study, we found that the ABT showed less variation in factors such as *Site* and *Selection line* compared with the PST (see Appendix III: Tables S3, S5, S7, S9, S11, S13). In contrast to the PST, in the ABT both, dwarf and dairy goats, were able to solve the problem (i.e., to detour the fence). This finding may indicate that the standardised ABT setup is better suited for ungulate species (Nawroth, Baciadonna, et al., 2016; Britta Osthaus et al., 2013) than the PST setup that was used in our study. Nevertheless, both tests seem to assess biologically relevant skills for goats as a species. Goats show highly selective feeding behaviour adapted to seasonal changes in plant abundance (Aldezabal & Garin, 2000). Hence, they need to have the ability to manipulate specific plants and to overcome spatial barriers to access feed sources that are out of reach. Indeed, domestic goats were shown to possess good spatial learning abilities in a maze learning paradigm (J Langbein, 2012) and were found to be capable of opening containers in other studies (Briefer et al., 2014; Jan Langbein et al., 2018; Nawroth, Brett, et al., 2016). The manipulation test in this study might have induced different levels of neophobia between selection lines and thus likely would have required higher levels of habituation to the setup. Even though the suitability of different test paradigms cannot be conclusively answered with our study, our findings suggest that caution must be taken when comparing different selection lines or phenotypes of a species if general claims about their cognitive capacities are to be made.

Conclusion

Our results suggest that cognitive testing *per se* and exposure to a testing environment via human-presented object-choice tests and isolation do not appear to notably affect the performance in subsequent conceptually different cognitive tests in goats. Furthermore, we found that variability induced by differences in phenotype and research site can be considerably greater than the effect of treatments when investigating the performance in spatial detour tests and instrumental problem-solving test. We therefore suggest that further cognitive research should conduct multi-lab studies to control for factors associated with testing on single sites or specific phenotypes of a species.

Acknowledgements

We thank the staff of the Agroscope Research Station in Ettenhausen, Switzerland, and of the Leibniz Institute for the Biology of Farm Animals in Dummerstorf, Germany, for assisting in this experiment and for caring for the animals. We additionally want to thank Dieter Sehland and Jonas Barkhau for coding our videos. This work was supported by an SNF grant (310030E-170537) to Nina Keil and a DFG grant (LA 1187/6-1) to Jan Langbein.

General Discussion

Captive animals live in restrictive environments with little opportunities to perform species-specific behaviour (Averós et al., 2010; Morgan & Tromborg, 2007). However, animals are highly motivated to engage in certain activities such as exploration, the acquisition of resources or problem solving, and their welfare may be impaired if they lack opportunities to do so (Hughes & Duncan, 1988; Hughes, 1997; Poole, 1992; Spinka & Wemelsfelder, 2017). The provision of opportunities to solve problems and control some aspects of the environment allows animals to engage evolved cognitive skills and can improve their welfare (Clark, 2017). In Chapter II, we assessed if goats have a need for cognitive stimulation and thus are willing to execute a cognitive task to receive a reward when the same reward is simultaneously available for free. By giving goats the choice between free food and opening a little sliding door to receive food, we showed that goats will invest effort in order to receive a valuable resource. In contrast to previous research, suggesting that the selection for high productivity has decreased the motivation to work for food (Lindqvist & Jensen, 2009; Schütz & Jensen, 2001), we found that dwarf and dairy goats are equally interested to work for a reward. In Chapter III and IV, we assessed the effect of long-term exposure to cognitive tests/experiments on measures of stress reactivity and behavioural flexibility in goats. We found that the participation in cognitive tests had only very subtle effects on measures of stress and behavioural flexibility. In this general discussion, I will critically examine our current results and previous findings on animals' motivation for cognitive stimulation. Then, I will discuss potential factors that might constitute the positive effects of cognitive enrichment in general and the lack of effects in our study as well as their implications for goat welfare. Due to the high variation between research sites and selection lines in our findings, I further highlight why we should account for variation in cognitive tests. Finally, I will give a brief outlook on future perspectives of the field of animal cognition and welfare.

Do animals want to be cognitively stimulated?

Seventy years ago, Harlow et al. (1950) found that captive rhesus macaques (*Macaca mulatta*) would explore mechanical "puzzles" for several hours a day and learn to open them without any food reward. The author concluded that the animals seem to have a drive to manipulate "which may be as primary and as important as the homeostatic drives."(Harlow et al., 1950). Since then, evidence has been accumulating that not only primates, but also other species are willing to voluntarily invest time and effort to solve challenging tasks or access a resource, even if they have the choice to get the same resource for free, a phenomenon called contrafreeloading (CFL, reviewed by Inglis et al., 1997). In Chapter II, we showed that goats would voluntarily perform in a cognitive task and choose to open a closed door rather than feeding from an open door (Rosenberger et al., 2020). Although many possible explanations for the occurrence of CFL exist, it seems likely that CFL is partly driven by an intrinsic need to gather information and to manipulate the environment. Both behaviours are a key for survival in the wild and thus might be an adaptive behaviour in animals in captivity. It may be an exaggeration to say that animals “want” to be cognitively enriched, but the stimulation of cognitive processes involved in information seeking, exploring and manipulating the environment possibly represent a behavioural need of animals.

White’s Competence Theory (White, 1959) postulates the need of animals to control and modify their surroundings. The CFL task we provided offered a possibility to gain control for the goats. As many goats opted to open the closed door several times, it is likely that the animals found the manipulation of the door rewarding. In previous studies, it has been also suggested that learning has reinforcing properties especially when making improvements in learning (Hagen & Broom, 2004). As goats were taught how to open the door before the CFL experiment started, we can exclude that it was the positive reinforcement of learning that prompted the goats to open the door. In other words, learning was already finished before the test and thus the task might not have been a cognitive challenge any longer. The more likely explanation for why goats chose to open the closed door is that the execution of the task was stimulating the need to manipulate the environment. Manipulating the environment may give the animal a feeling of control, which is an important aspect of animal welfare (Meehan & Mench, 2007).

The motivation to participate in (cognitive) experiments is depending on many factors. Participation may be hampered if a task is too difficult causing frustration (Meehan & Mench, 2007), if it is not appropriate or ecologically relevant for a species, or if the test environment is too stressful. Thus, participation rates might be a good indicator whether the provided task and experimental setup is adequate for a species but also the individual. A low participation rate in previous CFL literature often caused the exclusion of many animals from the experiment (Meagher et al., 2020). This may result in low sample sizes and selection bias towards certain individuals of a species and thus may reduce the validity of test results. In our study on CFL, we had very high participation rates, which suggests that the task was suitable and the test environment adequate. In contrast, in the PST in Chapter IV, hardly any dwarf goats, and very few dairy goats choose to open the lid of a container. Hence, one might infer that goats showed low motivation for this cognitive task. However, it is also possible that the motivation to manipulate the container was hampered by high stress levels during the test (Matzel et al., 2006). In case that participation rates are low, the response to “not participate” should be incorporated in the statistical analysis anyhow as it is a valid response by the animal (e.g. by means of IRTrees as suggested in Chapter II).

Whether and how fast a cognitive learning device finds acceptance by an animal may be partly determined by genetics, e.g. selection line. It has been proposed that the motivation for energy-demanding behaviours such as foraging or CFL was reduced in chicken selected for high productivity (Lindqvist & Jensen, 2008; Schütz & Jensen, 2001). In contrast, we found that dairy goats were faster than dwarf goats to choose the closed door and thus perform CFL in initial trials. Dwarf goats, which were not selected for productivity, took longer to reach similar levels in the extent of CFL like dairy goats. Another possible explanation for this initial reservation to manipulate the closed door in dwarf goats is that they were more fearful during the first trials than dairy goats. Animals selected for high productivity often show reduced stress reactivity in various stressful situations (Campler et al., 2009; Colpoys et al., 2014; Kilgour & Szantar-Coddington, 1995; Lindqvist & Jensen, 2008; Romeyer & Bouissou, 1992; Schütz & Jensen, 2001). Potentially higher stress levels of dwarf goats may have caused them to avoid the more difficult option in the beginning, until they settled and habituated to the testing procedure.

We cannot conclusively answer if the motivation behind CFL is a need for cognitive stimulation, the need to gather information, or the experience of a feeling of control. Either way, goats appear to have a high motivation to manipulate their environment, which may have intrinsic reinforcing properties. Our results confirm the findings by Langbein et al. (2009) who proposed that goats search for cognitive challenges beyond the motivation to get the reward and support the existing evidence that solving a challenging task can have intrinsic value for an animal. The provision of such challenges may even fulfil a behavioural need by stimulating cognitive abilities evolved to enhance survival in the wild. Thus, the incorporation of opportunities to engage species-specific cognitive skills into captive housing conditions may be an integral part of good animal husbandry.

What constitutes the positive effects of cognitive enrichment?

Captive husbandry conditions often lack opportunities for the stimulation of cognitive processes and may thus cause boredom in animals (Meagher, 2019). The provision of cognitive enrichment aims to engage animals in learning tasks and by this, provide opportunities to take control over their surroundings. Previous literature stated that cognitive stimulation by means of learning devices has positive effects on several measures of animal welfare (Oesterwind et al., 2016; Puppe et al., 2007; Zebunke et al., 2013). In our case, cognitive stimulation was provided in the form of human-administered cognitive tests. In contrast to previous studies, we could not find a substantial effect of cognitive stimulation in the studies in Chapter III and IV on welfare measured by stress reactivity, or on behavioural flexibility, respectively. Our results raise the questions: what caused the welfare-relevant behavioural changes described in previous studies applying cognitive enrichment? Moreover, what may be different in the way we provided cognitive stimulation?

Currently, there are many different approaches to providing cognitive enrichment. For instance, various types of cognitive enrichment devices have been used in the past, from “puzzle feeders” to touch screen; these have been provided to individual animals as well as to a group and sometimes they involve human contact, but often they do not. This is making it very difficult to disentangle which aspect of the provided cognitive stimulation caused the positive effects on welfare in previous studies. For instance, Zebunke et al. (2013) used a call-feeding station based on auditory conditioning to cognitively enrich pigs. The pigs first had to learn to associate an acoustic signal with their name in order to get a reward and at a later stage they additionally had to press a button to receive the food. The authors found that the enriched pigs were less fearful and more explorative in a combined open field/novel object test and concluded that successful coping with the call-feeding station caused this result. However, as the enrichment device also reduced agonistic interactions in the home pen of pigs, changes in behaviour may have been caused indirectly by reduced aggression rather than cognitive stimulation *per se*. Therefore, it appears that welfare parameters might be indirectly affected by the exposure to this specific task. In addition, time of exposure to cognitive stimulation may play a role. In the current thesis, COG animals in Chapter III and IV were exposed to 44 cognitive test days per animal spread over a period of 4-5 months. This exposure might have been too short or too irregularly to produce detectable differences between treatment groups. In contrast, Zebunke et al. (2013) enriched pigs daily, but only over 7 weeks and found positive effects on feeding behaviour and behaviour in stressful situations. Puppe et al. (2007) had used the same call-feeding station before and reported a reduction in open-field activity as well as a reduced excitement and fear behaviour in enriched compared to control pigs. They suggested that the call-feeding station may induces positive appraisal in pigs by the association of successful coping with a demanding behavioural task. Again, other behavioural patterns recorded within the housing environment, namely locomotor behaviour and belly nosing, were affected as well by the enrichment. The causational pathways of these effects, i.e. whether they are directly or indirectly associated with the cognitive stimulation, are not clear. In a study by Oesterwind et al. (2016), cognitive stimulation was provided in the form of a learning device where goats had to solve a visual four-choice discrimination task. The learning device was provided in a separate compartment of the enriched pens, which goats had to enter individually. The authors found that cognitively enriched goats were more curious and explorative towards a novel object in a novel object test than non-enriched goats. While goats interacted with the learning device they were inevitably separated from their conspecifics for a short period of time. This part-time separation may allowed enriched goats to habituate to isolation, in contrast to control animals, which have not been separated from their group in a similar manner. In addition, goats had visual and olfactory contact in the compartment containing the learning device which may made the separation less stressful (Siebert et al., 2011). As a result, it is difficult to determine if enriched goats were more explorative due to the cognitive stimulation or due to habituation to isolation (or an interaction of both). It becomes apparent from these studies that not only cognitive stimulation itself, but also many confounding factors associated with cognitive enrichment, such as human contact or separation from the group, could indirectly cause changes in behaviour as well.

One possible explanation for the lack of an effect on stress reactivity or behavioural flexibility in our tests is that our administered cognitive stimulation did not affect the behaviour of enriched goats substantially or that confounding effects associated with our study design hampered a detection of potential effects. In the standardised cognitive tests we used to enrich COG goats, a confounding factor that could have indirectly affected goat behaviour was the human-animal interaction and the separation of goats during testing. Using this approach, we did not find an effect of cognitive stimulation (COG) or human interaction (POS) on stress reactivity in an NA, NO, NH or weighing test. The previously mentioned studies applying cognitive stimulation successfully (Oesterwind et al., 2016; Puppe et al., 2007; Zebunke et al., 2013) incorporated the cognitive enrichment device in the home pen where the animals could actively decide to enter a test compartment and to participate or not. Thus, the animals could execute a certain amount of control. The acquisition of control has been suggested to be an important determinant of animal welfare (e.g. Broom, 1991) and a central part of cognitive enrichment (Clark, 2017). If opportunities to take control are lacking, cognitive enrichment may lose its purpose and thus will not produce positive effects. Thus, a lack of control in the way cognitive stimulation was provided in Chapter III and IV might explain the absence of a treatment effect. Goats in the COG treatment were taken into the test room by an experimenter and therefore could not decide by themselves when to seek cognitive stimulation. In addition, they had no control over being isolated, which is a major stressor particularly for social animals. Nonetheless, goats could still actively decide not to participate in the tests once they were in the test room. Further, it is possible that certain aspects of the cognitive testing procedure such as unpredictable training events were experienced as negative (Doyle et al., 2011; Galhardo et al., 2011). Autonomous enrichment devices may be perceived as very predictable for the animal. They can be used under the animal's terms regarding time of the day or frequency of use. COG as applied in Chapter III and IV was not always predictable for goats. Although tests started at the same time on test days, breaks in between test blocks made it impossible for the animals to predict when tests would start again. Also during testing itself, things were not always predictable for goats. For instance, the order of goats to be tested was randomised making it hard for the animal to anticipate its turn. These necessary steps and factors of cognitive testing might have been stressful for test subjects and reduced controllability and predictability for goats of all treatment groups. If cognitive testing had an effect at all, the high unpredictability may caused a flooring effect (by e.g. decreasing welfare) that prevented us to detect differences.

A second potential explanation for the lack of an effect of cognitive testing in Chapter III and IV on welfare measures is that the applied tests were not targeting ecologically relevant behavioural needs in order to be enriching. Different enrichment types are potentially stimulating/satisfying a variety of different behaviours such as foraging or exploration. In their study on pigs, Puppe et al. (2007) suggest that the call-feeding station may have satisfied a motivational need to forage for food in pigs. In the standardised tests we used to test COG goats, the goats had to discriminate between differently coloured pots, and interpret human-given physical and social cues to receive a food reward. In order to receive the reward, they had to indicate the location with their snouts through a metal fence, but they were not able to access the food without the help of the experimenter. This might not be an ecologically relevant task for goats as they usually browse for food in shrubs or graze and are not dependent on cooperation with conspecifics or even humans. Thus, the cognitive tests could not satisfy a behavioural need in this respect. This explanation is rather unlikely though, as the benefits of cognitive stimulation have been documented when providing animals with both, naturalistic challenges, such as species-specific foraging opportunities (de Jonge et al., 2008), as well as artificial problem-solving opportunities via the provision of computerised devices (Jan Langbein et al., 2004; Oesterwind et al., 2016). For instance, Langbein et al. (2004) and Oesterwind et al. (2016) used visual discrimination tasks presented on a touch screen as enrichment and found changes of vagal activity and behaviour indicative of a positive effect of such cognitive challenges. We also assessed visual discrimination learning in addition to physical and social cognition, but tasks were presented by a human experimenter. Hence, other factors than the type of enrichment applied likely played a more important role.

A third potential explanation for the lack of an effect of COG could be the very high amount of environmental enrichment and stimulation of all treatment groups. The applied test design, as well as the goats' living conditions and the frequent handling during tests might be seen as enrichment itself and thus may caused a ceiling effect in the welfare parameters we measured. An individual from a complex, enriched home environment may not find a task as rewarding as would an individual from a less stimulating or barren home environment. Various studies of farm animals have demonstrated that not only cognitive, but also structural enrichment can cause animals to be less fearful and more exploratory (Beattie et al., 2000; Hillmann et al., 2003; Oesterwind et al., 2016; Puppe et al., 2007). All treatment groups were kept in enriched home pens with climbing and hiding opportunities, and hay and straw being provisioned *ad libitum*. In addition, they were regularly taken out of their pens and exposed to new environmental settings, novel objects (i.e. activity loggers, ECG monitors), novel smells, and sounds, especially during test times. Taken together, all these factors may have raised the environmental enrichment and stimulation to a level where additional cognitive stimulation could not affect goats substantially any more.

Cognitive tests and consequences for welfare

It has been suggested that engagement in cognitive processes and thus the satisfaction of cognitive needs may be an important aspect of welfare for a range of species (Duncan & Petherick, 1991; Meehan & Mench, 2007; Nawroth et al., 2019; Spinka & Wemelsfelder, 2017). Participation in standardised cognitive tests as applied in Chapter III and IV does not appear to affect goats' response to different stressors or behavioural flexibility substantially. First, these results may suggest that cognitive tests commonly applied in animal (cognitive) research do not negatively affect goats' responses to future stressful situations or challenging tasks. The lack of finding an effect of the COG treatment may suggest that e.g. the type of cognitive tests used were not appropriate to stimulate relevant cognitive processes in goats. Another possible reason might be that the cognitive tests provided to the goats did not stimulate biologically relevant behaviours and thus had no obvious benefits for goats. Whether it is the aspect of 'thinking' or 'doing' that is enriching and as a result improving welfare, has yet to be determined. In light of these findings, the introduction of cognitive enrichment does not seem to be enriching for animals *per se*. Many factors such as the type of test, its difficulty, its biological/ecological relevance and its design, i.e. testing in isolation, are determining whether the cognitive stimulation provided is enriching or not.

Biological variation – a blessing and a curse?

In research, high reproducibility of a study enables the comparability of results of independent replicate studies (Goodman et al. 2016). The common approach in laboratory animal research to achieve reproducibility is strict standardization of both the test subjects and also their environment (Laukens, Brinkman, Raes, De Vos, & Vandenabeele, 2015; Willmann et al., 2012). However, this approach neglects the unavoidable existence of biological variation, caused by genetic differences, but also ontogeny, age, early experience, and social status (Voelkl et al., 2020). In their paper, Voelkl et al. (2020) propose a deliberate heterogenization of environmental factors to increase reproducibility of experimental studies.

For this thesis, we tested two selection lines of goats at two research sites. This approach allowed us to investigate how much variation was caused by selection line and by site. Genetic (i.e. selection for productivity) or environmental factors (i.e. housing conditions) may cause inter-individual variation. It has been suggested that it is necessary to appropriately incorporate differences between individuals rather than deriving conclusions from group level to prevent ecological fallacy (Robinson, 1950). Understanding the causes and consequences of individual differences in cognitive tests will improve the interpretation of animals' performances in these tests and allow accommodating better for their cognitive needs.

Variation caused by genetics – differences between selection lines

Genetic differences can lead to variability at the individual but also group level and may explain why some individuals are more likely to participate in and benefit from cognitive enrichment than others. Domestication and the selection for certain production traits in farm animals has also led to side-effects such as changes in behaviour and physiology (reviewed by Rauw et al., 1998). The literature suggests that the selection for high feed efficiency simultaneously led to a reduced stress response described in terms of physiological stress indicators such as lower cortisol levels (e.g., Knott, Cummins, Dunshea, & Leury, 2008; Sadler, Johnson, Lonergan, Nettleton, & Dekkers, 2011) and also in terms of behavioural stress indicators such as shorter duration of freezing and less escape attempts in pigs (Colpoys et al., 2014). Accordingly, we used two selection lines of goats differing in their grade of domestication and expected that dairy goats selected for high productivity will be less reactive towards different stressors and more explorative than dwarf goats that were not selected for high productivity. In both, the CFL task as well as the PST, dwarf goats took longer to approach the test device than dairy goats, suggesting higher stress reactivity in dwarf goats. However, we did not find general differences in reactivity in a NO, a WH or a NH test between the two selection lines of goats. Contrary to our expectations, dwarf goats were even less reactive to isolation in a novel arena test than dairy goats (Chapter III). Thus, stress reactivity appears to be expressed in a differential and context-depending manner, likely because the various situations in which we measured it included a variety of different stressors (separation from group, human presence, etc). However, no overall trend towards lower stress reactivity of dairy goats compared to dwarf goats could be observed.

Another trait that facilitates the approach of a novel problem is risk-taking behaviour. Risk-taking behaviour is described as a behaviour that may lead to negative consequences (Oliver & Hyde, 1993). High reactivity towards stress can affect and even reduce risk-taking behaviour (Maner et al., 2007). In Chapter IV in the PST, we found that dwarf goats touched a closed container less often than dairy goats. As the container was covered with a novel lid unknown to the animals, the approach of the container and subsequent touch required the animal to take a risk. Negative affective states such as fear play an important role in the decision-making process in humans and animals (e.g. Bechara, 2003; Mendl, Burman, Parker, & Paul, 2009). Fearful animals might touch the container less often and thus take lower risks by staying away from the container. Animals with low fear levels may be taking higher risks and touch the container more often. As risk-taking behaviour is a highly heritable trait (Van Oers, Drent, De Goede, & Van Noordwijk, 2004), it is possible that dairy goats were unintentionally selected for higher risk-taking behaviour as a by-product of the selection for high feed efficiency. In the model for the PST, we found that most variance was explained by *Individual* and *Selection line* rather than *Treatmen*t. In other words, the exposure to different treatments caused less variation in the variable *Touched* than genetic factors. In addition, risk-taking behaviour correlates with an individual's motivation to explore (Van Oers et al., 2004). If dairy goats have been unintentionally selected to take higher risks, they should also be more explorative. In the NO in Chapter III, we assessed the motivation to explore a novel object, amongst other things. We found that only ISO dwarf goats were less explorative than ISO dairy goats, while COG and POS dwarf goats were as explorative as COG and POS dairy goats. Voelkl et al. (2020) state that ".., treatment effects can be assessed and interpreted meaningfully only against biological variation — including gene × environment interactions". Thus, the combination of exposure to cognitive testing (COG), human-animal interaction (POS) and genetic differences in stress reactivity may explain the variation between treatment groups best. However, in the absence of statistically meaningful treatment effects, this is only one of many possible explanations.

Stress is known to influence many cognitive processes such as memory formation, i.e. how and what we learn in humans (Schwabe, Wolf, & Oitzl, 2010), and in animals (Mendl, 1999). In the ABT (Chapter IV), the *Accuracy* to detour a fence was used as a measure for behavioural flexibility of goats. In this parameter, we found that most variation in the data was explained by *Individual* rather than *Treatment*, *Site* or *Selection line*. This was also the case for the *Latency* to detour the fence and suggests that individual differences in reactivity towards stress could explain differences in the goats' ability and speed to make accurate decisions. In pigs, already mild stressors such as isolation, a novel food source, or a novel spatial environment have been shown to impair spatial memory, i.e the relocation of food sources (Mendl, Laughlin, & Hitchcock, 1997). Regolin et al. (1995) reported that the ability of chicks to succeed in an object permanence task was masked under stress, namely when tested in social isolation. In contrast, Brust et al. (2013) found that more fearful and active zebra finches are more flexible and perform better in reversal learning, as they are more attentive to their environment than less active birds. I suggest that individual variation in stress reactivity should be accounted for to implement cognitive enrichment successfully and to assess cognitive abilities correctly.

In addition, individuals may differ in their motivation to solve a cognitive task and in their abilities. Although the provision of cognitive enrichment has been suggested to engage cognitive skills and improve welfare (Clark, 2017), not every individual may be equally interested in cognitive stimulation.

For instance, Clark et al. (2013) found that two out of six subjects in a group of male dolphins were "high users" and used a cognitive device significantly more than their groupmates. Similarly, in the CFL study in Chapter II, all except of four goats chose to open the closed door at least once. However, inter-individual variation in the extent of performing CFL was high and ranged from one to seven of ten trials (Fig. 2, Chapter II). This indicates that the need for cognitive stimulation seems to be highly variable between individuals and can affect the motivation to use cognitive enrichment devices. In addition, individuals may differ in the level of stress that they experience when being confronted with a cognitive challenge. An animal that is successful in a task will likely have a rewarding experience, which will increase the probability to engage in the same task again. In contrast, an animal that cannot succeed due to impaired cognitive skills may be frustrated and does not want to repeat the task. In a study by Yamanashi and Matsuzawa (2010), individual chimpanzees differed in their reactions to making an error and to the task difficulty. While three out of the six chimpanzees exhibited higher rates of self-directed behaviours, an indicator for negative emotions, after incorrect trials and in difficult tasks, the other three chimpanzees did not. These results suggest that cognitive tests can be used to cognitively enrich animals, but their success is dependent on the appropriateness of the task and the individuals' motivation to engage in it.

Variation caused by the environment - differences between sites

Although variation in behavioural traits of an animal is often genetically determined, it is constantly modulated by the interaction of genetic traits with the environment (Boissy & Bouissou, 1995). In the current thesis, the environmental factor *Site* caused a lot of variation in the behavior of goats in tests, which highlights the importance of multi-lab/site approaches in cognitive research. For instance, in the PST, we found that dwarf goats in DU almost never touched the container, while all dwarf goats in ET touched it at least once. Dwarf goats across both sites mainly originated from the same dwarf goat population, thus environmental factors rather than genetics likely affected behaviour. In dairy goats, the average number of touches was also consistently higher in ET than in DU, again indicating that variation was caused by environmental factors. However, as mentioned earlier, the most variation in the variable *Touched* was explained by *Individual*. Thus, it is also possible that individual variation and site variation were interfering with each other. At both sites, none of the dwarf goats opened the container, while almost all dairy goats in ET, but hardly any dairy goat in DU did so. In order to open the container, goats had to approach the experimenter who was standing next to the container. One explanation could be that previous experiences with animal care staff at the two research sites differed. These experiences may caused differences in stress reactivity towards humans in goats in ET compared to DU (Boivin et al., 1998; Hausberger et al., 2008). This is supported by the fact that we found a high variance in selection lines between research sites in terms of sociability towards a novel human in the NH (Chapter III). Further, differences in the size of the farms where the goats were bred (Mattiello et al., 2010), their rearing history, i.e. human- or dam-reared (Lyons, 1989), or habituation to isolation and handling at a young age may affected goats differently at the two sites (Boivin & Braastad, 1996).

Final Conclusions

To conclude, we could show that domestic goats are willing to work for food even if the same food is available for free, which might reflect a need for stimulation of species-specific cognitive processes. Thus, the incorporation of opportunities to engage evolved cognitive skills into captive housing conditions may be an integral part of good animal husbandry. Further, we did not find an effect of long-term cognitive testing *per se* or human-animal interaction on responses to different stressors or goats' detour or problem-solving performance. Possible explanations may be that the feeling of control was lacking in the way cognitive stimulation was provided in Chapter III and IV. If opportunities to take control are lacking, cognitive enrichment may lose its purpose and will not produce positive effects. It is also possible that the level of environmental enrichment was so high that the additional cognitive stimulation did not affect goats substantially or in a biologically meaningful way. The selection for high productivity did not reduce the willingness to perform CFL in goats, but it seems that it decreased stress reactivity towards humans and as a result also increased the goats' inclination to approach a human and improved the performance in a problem-solving task. In addition to differences between selection lines, we also found that individual goats differ in their motivation to solve challenging tasks and in their reactivity towards stressors. These differences between selection lines and individuals may be also used to inform handling procedures and husbandry conditions for goats. In sum, we found high variation between individuals when systematically varying the genotype (different selection lines) of the target species, and by introducing uncontrolled variation in the form of a multi-site approach (as suggested by Voelkl et al., 2020). It becomes evident that heterogenization of both test subjects and the environment is not only a necessity for external validity and reproducible results, but also beneficial for the interpretation of test results.

Outlook and perspectives

Farm animals such as domestic goats have been found to seek cognitive stimulation, albeit to various extents. Its provision can improve animal welfare, likely by targeting a behavioural need. This need should be considered in the design of husbandry systems for farm animals by means of opportunities to engage species-specific cognitive capacities. Further, cognitive tests commonly used in comparative cognitive research such as object-choice tasks do not seem to negatively affect welfare of goats. However, we could not find substantial benefits of cognitive testing on goat welfare either. This suggests that more research is needed to disentangle possible confounding factors, e.g., testing individuals in isolation or the human-animal-relationship from effects of cognitive stimulation *per se.*

Future research in comparative cognition/applied animal welfare should thrive to consider certain aspects:

Separation from the group is a potential stressor for animals. In future studies, we should thrive to design tests in such a way that animals can be tested either in a group or at least close to it with visual contact. This will not only reduce stress for the animal but also lead to more valid test results, as stress can impair many cognitive functions. Thus, testing animals in a social setting can lead to increased validity of tests and improve animal welfare.

Individual animals differ in their motivation, experiences, and skills. Thus, modular devices should be applied that allow adapting the level of difficulty of a task to the individuals' skills and needs. This may also help to avoid frustration in animals that are not able to solve a task or boredom in the very skilled ones, as only appropriate challenges will have beneficial effects on test subjects and produce lasting effects for animal welfare.

Large differences in the type and provision of cognitive enrichment devices are making comparisons between studies very difficult. Currently, there is no consensus whether it is the 'thinking' or 'doing' aspect of cognitive stimulation that is enriching and targeting a need and as a result improving welfare. To determine what constitutes the positive effects of cognitive enrichment, more studies investigating controllability, predictability, and the reinforcing properties of learning itself are needed.

By increasing the heterogeneity of our test design, we were able to explain variation caused by individuals, selection lines, and sites in our studies. To increase external validity and to improve reproducibility of future studies in the fields of animal research, different phenotypes and a multi-lab/site approach should be favoured.

References

Akinsoyinu, A. O., Mba, A. U., & Olubajo, F. O. 0. (1977). Studies on milk yield and composition of the West African dwarf goat in Nigeria. *Journal of Dairy Research*, *44*(1), 57–62. https://doi.org/10.1017/S0022029900019920

Aldezabal, A., & Garin, I. (2000). Browsing preference of feral goats (Capra hircus L.) in a Mediterranean mountain scrubland. *Journal of Arid Environments*, *44*(1), 133–142. https://doi.org/10.1006/jare.1999.0573

Alonso, M. E., González-Montaña, J. R., & Lomillos, J. M. (2020). Consumers' concerns and perceptions of farm animal welfare. *Animals*, *10*(3), 1–13. https://doi.org/10.3390/ani10030385

ASAB, & ABS. (2018). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour*, *135*, 1–11. https://doi.org/10.1016/j.anbehav.2017.10.001

Aschwanden, J., Gygax, L., Wechsler, B., & Keil, N. M. (2008a). Cardiac activity in dairy goats whilst feeding side-by-side at two different distances and during social separation. *Physiology and Behavior*, *95*(5), 641–648. https://doi.org/10.1016/j.physbeh.2008.09.016

Aschwanden, J., Gygax, L., Wechsler, B., & Keil, N. M. (2008b). Cardiac activity in dairy goats whilst feeding side-by-side at two different distances and during social separation. *Physiology and Behavior*, *95*(5), 641–648. https://doi.org/10.1016/j.physbeh.2008.09.016

Aschwanden, J., Gygax, L., Wechsler, B., & Keil, N. M. (2009). Loose housing of small goat groups: Influence of visual cover and elevated levels on feeding, resting and agonistic behaviour. *Applied Animal Behaviour Science*, *119*(3–4), 171–179. https://doi.org/10.1016/j.applanim.2009.04.005

Averós, X., Brossard, L., Dourmad, J. Y., de Greef, K. H., Edge, H. L., Edwards, S. A., & Meunier-Salaün, M. C. (2010). A meta-analysis of the combined effect of housing and environmental enrichment characteristics on the behaviour and performance of pigs. *Applied Animal Behaviour Science*, *127*(3–4), 73–85. https://doi.org/10.1016/j.applanim.2010.09.010

Baciadonna, L., Duepjan, S., Briefer, E. F., de la Torre, M. P., & Nawroth, C. (2018). Looking on the bright side of livestock emotions-the potential of their transmission to promote positive welfare. *Frontiers in Veterinary Science*, *5*(SEP), 1–6. https://doi.org/10.3389/fvets.2018.00218

Baker, K. C. (2004). Benefits of positive human interaction for socially housed chimpanzees. *Animal Welfare*, *13*(2), 239–245.

Bassett, L., & Buchanan-Smith, H. M. (2007). Effects of predictability on the welfare of captive animals. *Applied Animal Behaviour Science*, *102*(3–4), 223–245. https://doi.org/10.1016/j.applanim.2006.05.029

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., Walker, S., Haubo, R., Christensen, B., … Scheipl, F. (2015). Package “lme4.” *Convergence*, *12*(1), 2.

Bean, D., Mason, G. J., & Bateson, M. (1999). Contrafreeloading in starlings: testing the information hypothesis. *Behaviour*, *136*(10–11), 1267–1282.

Beattie, V. E., O’Connell, N. E., Kilpatrick, D. J., & Moss, B. W. (2000). Influence of environmental enrichment on welfare-related behavioural and physiological parameters in growing pigs. *Animal Science*, *70*(3), 443–450. https://doi.org/10.1017/S1357729800051791

Bechara, A. (2003). Risky business: Emotion, decision-making, and addiction. *Journal of Gambling Studies*, *19*(1), 23–51. https://doi.org/10.1023/A:1021223113233

Beilharz, R. G., Luxford, B. G., & Wilkinson, J. L. (1993). Quantitative genetics and evolution: Is our understanding of genetics sufficient to explain evolution? *Journal of Animal Breeding and Genetics*, *110*(1–6), 161–170. https://doi.org/10.1111/j.1439-0388.1993.tb00728.x

Bekoff, M. (2000). Animal emotions: Exploring passionate natures. *BioScience*, *50*(10), 861–870. https://doi.org/10.1641/0006-3568(2000)050[0861:AEEPN]2.0.CO;2

Benson-Amram, S., & Holekamp, K. E. (2012). Innovative problem solving by wild spotted hyenas. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1744), 4087–4095. https://doi.org/10.1098/rspb.2012.1450

Boissy, A., & Bouissou, M. F. (1988). Effects of early handling on heifers’ subsequent reactivity to humans and to unfamiliar situations. *Applied Animal Behaviour Science*, *20*(3–4), 259–273. https://doi.org/10.1016/0168-1591(88)90051-2

Boissy, A., & Bouissou, M. F. (1995). Assessment of individual differences in behavioural reactions of heifers exposed to various fear-eliciting situations. *Applied Animal Behaviour Science*, *46*(1–2), 17–31. https://doi.org/10.1016/0168-1591(95)00633-8

Boissy, A., Bouix, J., Orgeur, P., Poindron, P., Bibé, B., & Le Neindre, P. (2005). Genetic analysis of emotional reactivity in sheep: Effects of the genotypes of the lambs and of their dams. *Genetics Selection Evolution*, *37*(4), 381–401. https://doi.org/10.1051/gse:2005007

Boissy, A., & Le Neindre, P. (1997). Behavioral, cardiac and cortisol responses to brief peer separation and reunion in cattle. *Physiology and Behavior*, *61*(5), 693–699. https://doi.org/10.1016/S0031-9384(96)00521-5

Boissy, A., & Lee, C. (2014). How assessing relationships between emotions and cognition can improve farm animal welfare. *OIE Revue Scientifique et Technique*, *33*(1), 103–110. https://doi.org/10.20506/rst.33.1.2260

Boissy, A., Manteuffel, G., Jensen, M. B., Moe, R. O., Spruijt, B., Keeling, L. J., … Aubert, A. (2007). Assessment of positive emotions in animals to improve their welfare. *Physiology and Behavior*, *92*(3), 375–397. https://doi.org/10.1016/j.physbeh.2007.02.003

Boivin, X., & Braastad, B. O. (1996). Effects of handling during temporary isolation after early weaning on goat kids' later response to humans. *Applied Animal Behaviour Science*, *48*(1–2), 61–71. https://doi.org/10.1016/0168-1591(95)01019-X

Boivin, X., Garel, J. P., Mante, A., & Le Neindre, P. (1998). Beef calves react differently to different handlers according to the test situation and their previous interactions with their caretaker. *Applied Animal Behaviour Science*, *55*(3–4), 245–257. https://doi.org/10.1016/S0168-1591(97)00050-6

Boivin, X., Neindre, P. Le, & Chupin, J. M. (1992). Establishment of cattle-human relationships. *Applied Animal Behaviour Science*, *32*(4), 325–335. https://doi.org/10.1016/S0168-1591(05)80025-5

Boivin, X., Tournadret, H., & Le Neindre, P. (2000). Hand-feeding and gentling influence early-weaned lambs' attachment responses to their stockperson. *Journal of Animal Science*, *78*(4), 879–884. https://doi.org/10.2527/2000.784879x

Boogert, N. J., Reader, S. M., & Laland, K. N. (2006). The relation between social rank, neophobia and individual learning in starlings. *Animal Behaviour*, *72*(6), 1229–1239. https://doi.org/10.1016/j.anbehav.2006.02.021

Brajon, S., Laforest, J. P., Schmitt, O., & Devillers, N. (2016). A preliminary study of the effects of individual response to challenge tests and stress induced by humans on learning performance of weaned piglets (Sus scrofa). *Behavioural Processes*, *129*, 27–36. https://doi.org/10.1016/j.beproc.2016.05.007

Brando, S. I. C. A. (2012). Animal Learning and Training. Implications for Animal Welfare. *Veterinary Clinics of North America - Exotic Animal Practice*, *15*(3), 387–398. https://doi.org/10.1016/j.cvex.2012.06.008

Bray, E. E., MacLean, E. L., & Hare, B. A. (2014). Context specificity of inhibitory control in dogs. *Animal Cognition*, *17*(1), 15–31. https://doi.org/10.1007/s10071-013-0633-z

Breuer, K., Hemsworth, P. H., & Coleman, G. J. (2003). The effect of positive or negative handling on the behavioural and physiological responses of nonlactating heifers. *Applied Animal Behaviour Science*, *84*(1), 3–22. https://doi.org/10.1016/S0168-1591(03)00146-1

Briefer, E. F., Haque, S., Baciadonna, L., & McElligott, A. G. (2014). Goats excel at learning and remembering a highly novel cognitive task. *Frontiers in Zoology*, *11*(1). https://doi.org/10.1186/1742-9994-11-20

Brockhurst, J., Cheleuitte-Nieves, C., Buckmaster, C. L., Schatzberg, A. F., & Lyons, D. M. (2015). Stress inoculation modeled in mice. *Translational Psychiatry*, *5*(3), 1–5. https://doi.org/10.1038/tp.2015.34

Broom, D. M. (1986). Indicators of poor welfare. *British Veterinary Journal*, *142*, 524–526.

Broom, D. M. (1991). Animal welfare: concepts and measurement. *Journal of Animal Science*, *69*(10), 4167–4175. https://doi.org/10.2527/1991.69104167x

Broom, Donald M. (2008). Welfare assessment and relevant ethical decisions: Key concepts. *Annual Review of Biomedical Sciences*, *10*, 79–90. https://doi.org/10.5016/1806-8774.2008.v10pT79

Broom, Donald M. (2019). Land and water usage in beef production systems. *Animals*, *9*(6), 1–13. https://doi.org/10.3390/ani9060286

Brust, V., Wuerz, Y., & Krüger, O. (2013). Behavioural Flexibility and Personality in Zebra Finches. *Ethology*, *119*(7), 559–569. https://doi.org/10.1111/eth.12095

Brydges, N. M., & Braithwaite, V. A. (2008). Measuring animal welfare: What can cognition contribute? *Annual Review of Biomedical Sciences*, *10*, 91–103. https://doi.org/10.5016/1806-8774.2008.v10pT91

Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: Caveats and guidelines. *Ethology*, *116*(5), 472–480. https://doi.org/10.1111/j.1439-0310.2010.01758.x

Campler, M., Jöngren, M., & Jensen, P. (2009). Fearfulness in red junglefowl and domesticated White Leghorn chickens. *Behavioural Processes*, *81*(1), 39–43. https://doi.org/10.1016/j.beproc.2008.12.018

Carbonaro, D. A., Friend, T. H., Dellmeier, G. R., & Nuti, L. C. (1992). Behavioral and physiological responses of dairy goats to food thwarting. *Physiology and Behavior*, *51*(2), 303–308. https://doi.org/10.1016/0031-9384(92)90145-R

Carlstead, K., & Shepherdson, D. (2000). Alleviating stress in zoo animals with environmental enrichment. In G. P. M. & J. A. Mench (Ed.), *The biology of animal stress: Basic principles and implications for animal welfare* (pp. 337–354). London, UK: CABI Publishing.

Cerrone, M. (2019). Keepers as social companions: Tactile communication and social enrichment for captive apes. *Sign Systems Studies*, *47*(3–4), 453–479. https://doi.org/10.12697/SSS.2019.47.3-4.06

Clark, F. E. (2011). Great ape cognition and captive care: Can cognitive challenges enhance well-being? *Applied Animal Behaviour Science*, *135*(1–2), 1–12. https://doi.org/10.1016/j.applanim.2011.10.010

Clark, F. E. (2017). Cognitive enrichment and welfare: Current approaches and future directions. *Animal Behavior and Cognition*, *4*(1), 52–71. https://doi.org/10.12966/abc.05.02.2017

Clark, F. E., & Smith, L. J. (2013). Effect of a cognitive challenge device containing food and non-food rewards on chimpanzee well-being. *American Journal of Primatology*, *75*(8), 807–816. https://doi.org/10.1002/ajp.22141

Clarke, C. H., & Jones, R. B. (2000). Responses of adult laying hens to video images presented repeatedly outside the home cage. *Applied Animal Behaviour Science*, *67*(1–2), 97–110. https://doi.org/10.1016/S0168-1591(99)00107-0

Cole, E. F., & Quinn, J. L. (2012). Personality and problem-solving performance explain competitive ability in the wild. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1731), 1168–1175. https://doi.org/10.1098/rspb.2011.1539

Colpoys, J. D., Abell, C. E., Young, J. M., Keating, A. F., Gabler, N. K., Millman, S. T., … Johnson, A. K. (2014). Effects of genetic selection for residual feed intake on behavioral reactivity of castrated male pigs to novel stimuli tests. *Applied Animal Behaviour Science*, *159*, 34–40. https://doi.org/10.1016/j.applanim.2014.06.013

Croney, C. C., Gardner, B., & Baggot, S. (2004). Beyond Animal Husbandry. *Essays in Philosophy*, *5*(2), 391–403. https://doi.org/10.5840/eip20045213

D’Eath, R. B., Roehe, R., Turner, S. P., Ison, S. H., Farish, M., Jack, M. C., & Lawrence, A. B. (2009). Genetics of animal temperament: Aggressive behaviour at mixing is genetically associated with the response to handling in pigs. *Animal*, *3*(11), 1544–1554. https://doi.org/10.1017/S1751731109990528

Da Costa, A. P., Leigh, A. E., Man, M. S., & Kendrick, K. M. (2004). Face pictures reduce behavioural, autonomic, endocrine and neural indices of stress and fear in sheep. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1552), 2077–2084. https://doi.org/10.1098/rspb.2004.2831

Dall, S. R. X., Giraldeau, L. A., Olsson, O., McNamara, J. M., & Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution*, *20*(4), 187–193. https://doi.org/10.1016/j.tree.2005.01.010

Dawkins, M. S. (2000). Animal minds and animal emotions. *American Zoologist*, *40*(6), 883–888. https://doi.org/10.1093/icb/40.6.883

de Azevedo, C. S., Cipreste, C. F., & Young, R. J. (2007). Environmental enrichment: A GAP analysis. *Applied Animal Behaviour Science*, *102*(3–4), 329–343. https://doi.org/10.1016/j.applanim.2006.05.034

De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-Based Item Response Models of the GLMM Family. *JSS Journal of Statistical Software*, *48*. Retrieved from http://www.jstatsoft.org/

De Bona, S., Valkonen, J. K., López-Sepulcre, A., & Mappes, J. (2015). Predator mimicry, not conspicuousness, explains the efficacy of butterfly eyespots. *Proceedings of the Royal Society B: Biological Sciences*, *282*(20150202). https://doi.org/10.1098/rspb.2015.0202

de Jonge, F. H., Tilly, S. L., Baars, A. M., & Spruijt, B. M. (2008). On the rewarding nature of appetitive feeding behaviour in pigs (Sus scrofa): Do domesticated pigs contrafreeload? *Applied Animal Behaviour Science*, *114*(3–4), 359–372. https://doi.org/10.1016/j.applanim.2008.03.006

De Paula Vieira, A., de Passillé, A. M., & Weary, D. M. (2012). Effects of the early social environment on behavioral responses of dairy calves to novel events. *Journal of Dairy Science*, *95*(9), 5149–5155. https://doi.org/10.3168/jds.2011-5073

De Paula Vieira, A., von Keyserlingk, M. a G., & Weary, D. M. (2010). Effects of pair versus single housing on performance and behavior of dairy calves before and after weaning from milk. *Journal of Dairy Science*, *93*(7), 3079–3085. https://doi.org/10.3168/jds.2009-2516

de Resende, L. S., Gomes, K. C. P., Andriolo, A., Genaro, G., Remy, G. L., & de Almeida Ramos, V. (2011). Influence of cinnamon and catnip on the stereotypical pacing of oncilla cats (Leopardus tigrinus) in captivity. *Journal of Applied Animal Welfare Science*, *14*(3), 247–254. https://doi.org/10.1080/10888705.2011.576981

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

Dicko, M. S., & Sikena, L. K. (1992). Feeding behaviour, quantitative and qualitative intake of browse by domestic ruminants. Legume trees and other fodder trees as protein sources for livestock. *Food and Agriculture Organization*, 129–144.

Dodd, C. L., Pitchford, W. S., Hocking Edwards, J. E., & Hazel, S. J. (2012). Measures of behavioural reactivity and their relationships with production traits in sheep: A review. *Applied Animal Behaviour Science*, *140*(1–2), 1–15. https://doi.org/10.1016/j.applanim.2012.03.018

Doyle, R. E., Lee, C., Deiss, V., Fisher, A. D., Hinch, G. N., & Boissy, A. (2011). Measuring judgement bias and emotional reactivity in sheep following long-term exposure to unpredictable and aversive events. *Physiology and Behavior*, *102*(5), 503–510. https://doi.org/10.1016/j.physbeh.2011.01.001

Dudde, A., Krause, E. T., Matthews, L. R., & Schrader, L. (2018). More than eggs - relationship between productivity and learning in laying hens. *Frontiers in Psychology*, *9*(OCT), 1–10. https://doi.org/10.3389/fpsyg.2018.02000

Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *6*(5), 178–190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the social brain. *Science*, *317*(5843), 1344–1347. https://doi.org/10.1126/science.1145463

Duncan, I. J., & Petherick, J. C. (1991). The implications of cognitive processes for animal welfare. *Journal of Animal Science*, *69*(12), 5017–5022. https://doi.org/10.2527/1991.69125017x

Ernst, K., Tuchscherer, M., Kanitz, E., Puppe, B., & Manteuffel, G. (2006). Effects of attention and rewarded activity on immune parameters and wound healing in pigs. *Physiology and Behavior*, *89*(3), 448–456. https://doi.org/10.1016/j.physbeh.2006.07.001

Ernst, Katrin, Puppe, B., Schön, P. C., & Manteuffel, G. (2005). A complex automatic feeding system for pigs aimed to induce successful behavioural coping by cognitive adaptation. *Applied Animal Behaviour Science*, *91*(3–4), 205–218. https://doi.org/10.1016/j.applanim.2004.10.010

Fay, C., & Miller, L. (2015). Utilizing Scents as Environmental Enrichment: Preference Assessment and Application with Rothschild Giraffe. *Animal Behavior and Cognition*, *2*(3), 285–291. https://doi.org/10.12966/abc.08.07.2015

Fay, M. P. (2010). Two-sided exact tests and matching confidence intervals for discrete data. *R Journal*, *2*(1), 53–58. https://doi.org/10.32614/rj-2010-008

Finkemeier, M. A., Langbein, J., & Puppe, B. (2018). Personality research in mammalian farm animals: Concepts, measures, and relationship to welfare. *Frontiers in Veterinary Science*, *5*(131). https://doi.org/10.3389/fvets.2018.00131

Finkemeier, M. A., Oesterwind, S., Nürnberg, G., Puppe, B., & Langbein, J. (2019). Assessment of personality types in Nigerian dwarf goats (Capra hircus) and cross-context correlations to behavioural and physiological responses. *Applied Animal Behaviour Science*, *217*(February), 28–35. https://doi.org/10.1016/j.applanim.2019.05.004

Forkman, B., Boissy, a., Meunier-Salaün, M. C., Canali, E., & Jones, R. B. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology and Behavior*, *92*, 340–374. https://doi.org/10.1016/j.physbeh.2007.03.016

Franks, B. (2019). What do animals want? *Animal Welfare*, *28*(1), 1–10. https://doi.org/10.7120/09627286.28.1.001

Fureix, C., & Meagher, R. K. (2015). What can inactivity (in its various forms) reveal about affective states in non-human animals? A review. *Applied Animal Behaviour Science*, *171*, 8–24. https://doi.org/10.1016/j.applanim.2015.08.036

Gabor, V., & Gerken, M. (2014). Shetland ponies (Equus caballus) show quantity discrimination in a matching-to-sample design. *Animal Cognition*, *17*(6), 1233–1243. https://doi.org/10.1007/s10071-014-0753-0

Gaillard, C., Meagher, R. K., Von Keyserlingk, M. a G., & Weary, D. M. (2014). Social housing improves dairy calves' performance in two cognitive tests. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0090205

Galhardo, L., Vital, J., & Oliveira, R. F. (2011). The role of predictability in the stress response of a cichlid fish. *Physiology and Behavior*, *102*(3–4), 367–372. https://doi.org/10.1016/j.physbeh.2010.11.035

Gieling, E. T., Nordquist, R. E., & van der Staay, F. J. (2011). Assessing learning and memory in pigs. *Animal Cognition*, *14*(2), 151–173. https://doi.org/10.1007/s10071-010-0364-3

Gifford, A. K., Cloutier, S., & Newberry, R. C. (2007). Objects as enrichment: Effects of object exposure time and delay interval on object recognition memory of the domestic pig. *Applied Animal Behaviour Science*, *107*(3–4), 206–217. https://doi.org/10.1016/j.applanim.2006.10.019

Graham, L., Wells, D. L., & Hepper, P. G. (2005). The influence of olfactory stimulation on the behaviour of dogs housed in a rescue shelter. *Applied Animal Behaviour Science*, *91*(1–2), 143–153. https://doi.org/10.1016/j.applanim.2004.08.024

Graunke, K. L., Nürnberg, G., Repsilber, D., Puppe, B., & Langbein, J. (2013a). Describing Temperament in an Ungulate: A Multidimensional Approach. *PLoS ONE*, *8*(9), 1–12. https://doi.org/10.1371/journal.pone.0074579

Graunke, K. L., Nürnberg, G., Repsilber, D., Puppe, B., & Langbein, J. (2013b). Describing Temperament in an Ungulate: A Multidimensional Approach. *PLoS ONE*, *8*(9), 1–12. https://doi.org/10.1371/journal.pone.0074579

Grignard, L., Boissy, A., Boivin, X., Garel, J. P., & Le Neindre, P. (2000). The social environment influences the behavioural responses of beef cattle to handling. *Applied Animal Behaviour Science*, *68*(1), 1–11. https://doi.org/10.1016/S0168-1591(00)00085-X

Grimberg-Henrici, C. G. E., Vermaak, P., Elizabeth Bolhuis, J., Nordquist, R. E., & van der Staay, F. J. (2016). Effects of environmental enrichment on cognitive performance of pigs in a spatial holeboard discrimination task. *Animal Cognition*, *19*(2), 271–283. https://doi.org/10.1007/s10071-015-0932-7

Hagen, K., & Broom, D. M. (2004). Emotional reactions to learning in cattle. *Applied Animal Behaviour Science*, *85*(3–4), 203–213. https://doi.org/10.1016/j.applanim.2003.11.007

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65.

Harlow, H. F., Harlow, M. K., & Meyer, D. R. (1950). Learning motivated by a manipulation drive. *Journal of Experimental Psychology*, *40*(2), 228–234. https://doi.org/10.1037/h0056906

Hartig, F. (2020). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. Retrieved from https://cran.r-project.org/package=DHARMa

Hausberger, M., Roche, H., Henry, S., & Visser, E. K. (2008). A review of the human-horse relationship. *Applied Animal Behaviour Science*, *109*(1), 1–24. https://doi.org/10.1016/j.applanim.2007.04.015

Held, S., Baumgartner, J., Kilbride, A., Byrne, R. W., & Mendl, M. (2005). Foraging behaviour in domestic pigs (Sus scrofa): Remembering and prioritizing food sites of different value. *Animal Cognition*, *8*(2), 114–121. https://doi.org/10.1007/s10071-004-0242-y

Hemsworth, P. H. (2003). Human-animal interactions in livestock production. *Applied Animal Behaviour Science*, *81*(3), 185–198. https://doi.org/10.1016/S0168-1591(02)00280-0

Hemsworth, P. H., Mellor, D. J., Cronin, G. M., & Tilbrook, A. J. (2015). Scientific assessment of animal welfare. *New Zealand Veterinary Journal*, *63*(1), 24–30.

https://doi.org/10.1080/00480169.2014.966167

Hemsworth, P. H., Verge, J., & Coleman, G. J. (1996). Conditioned approach-avoidance responses to humans: The ability of pigs to associate feeding and aversive social experiences in the presence of humans with humans. *Applied Animal Behaviour Science*, *50*(1), 71–82. https://doi.org/10.1016/0168-1591(96)01065-9

Hessle, A., Rutter, M., & Wallin, K. (2008). Effect of breed, season and pasture moisture gradient on foraging behaviour in cattle on semi-natural grasslands. *Applied Animal Behaviour Science*, *111*(1–2), 108–119. https://doi.org/10.1016/j.applanim.2007.05.017

Hillmann, E., Von Hollen, F., Bünger, B., Todt, D., & Schrader, L. (2003). Farrowing conditions affect the reactions of piglets towards novel environment and social confrontation at weaning. *Applied Animal Behaviour Science*, *81*(2), 99–109. https://doi.org/10.1016/S0168-1591(02)00254-X

Hofmann, R. R. (1989). Evolutionary steps of ecophysiological adaptation and diversification of ruminants: a comparative view of their digestive system*’**. *Oecologia*, *78*, 443–457.

Hölscher, C. (1999). Stress impairs performance in spatial water maze learning tasks. *Behavioural Brain Research*, *100*(1–2), 225–235. https://doi.org/10.1016/S0166-4328(98)00134-X

Hothorn, T., Bretz, F., Westfall, P., & Heiberger, R. M. (2008). Multcomp: simultaneous inference for general linear hypotheses. *R Package Version*, 1--3.

Hughes, B. O., & Duncan, I. J. H. (1988). The notion of ethological “need”, models of motivation and animal welfare. *Anim. Behav*, *36*, 1696–1707.

Hughes, R. N. (1997). Intrinsic exploration in animals: Motives and measurement. *Behavioural Processes*, *41*(3), 213–226. https://doi.org/10.1016/S0376-6357(97)00055-7

Ickes, B. R., Pham, T. M., Sanders, L. A., Albeck, D. S., Mohammed, A. H., & Granholm, A. C. (2000). Long-term environmental enrichment leads to regional increases in neurotrophin levels in rat brain. *Experimental Neurology*, *164*(1), 45–52. https://doi.org/10.1006/exnr.2000.7415

Inglis, I., Forkman, B., & Lazarus, J. (1997). Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behaviour*, *53*(6), 1171–1191. https://doi.org/10.1006/anbe.1996.0320

Inglis, I. R., & Ferguson, N. J. . J. K. (1986). Starlings search for food rather than eat freely-available, identical food. *Animal Behaviour*, *34*(2), 614–617. https://doi.org/10.1016/S0003-3472(86)80136-1

Inglis, Ian R., Langton, S., Forkman, B., & Lazarus, J. (2001). An information primacy model of exploratory and foraging behaviour. *Animal Behaviour*, *62*(3), 543–557. https://doi.org/10.1006/anbe.2001.1780

Jacobson, S. L., Kwiatt, A. C., Ross, S. R., & Cronin, K. A. (2019a). The effects of cognitive testing on the welfare of zoo-housed Japanese macaques (Macaca fuscata). *Applied Animal Behaviour Science*, *212*(November 2018), 90–97. https://doi.org/10.1016/j.applanim.2018.12.014

Jacobson, S. L., Kwiatt, A. C., Ross, S. R., & Cronin, K. A. (2019b). The effects of cognitive testing on the welfare of zoo-housed Japanese macaques (Macaca fuscata). *Applied Animal Behaviour Science*, *212*, 90–97. https://doi.org/10.1016/j.applanim.2018.12.014

Jago, J. G., Krohn, C. C., & Matthews, L. R. (1999). The influence of feeding and handling on the development of the human-animal interactions in young cattle. *Applied Animal Behaviour Science*, *62*(2–3), 137–151. https://doi.org/10.1016/S0168-1591(98)00219-6

Jensen, G. D. (1963). Preference for bar pressing over" freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology*, *65*(5), 451–454.

Jensen, P. (2006). Domestication - From behaviour to genes and back again. *Applied Animal Behaviour Science*, *97*(1), 3–15. https://doi.org/10.1016/j.applanim.2005.11.015

Jensen, P., & Toates, F. M. (1993). Who needs “behavioural needs”? Motivational aspects of the needs of animals. *Applied Animal Behaviour Science*, *37*(2), 161–181. https://doi.org/10.1016/0168-1591(93)90108-2

Jones, N. A. R., Webster, M., & Vea Salvanes, A. G. (2021). Physical enrichment research for captive fish: time to focus on the DETAILS . *Journal of Fish Biology*, 0–2. https://doi.org/10.1111/jfb.14773

Kabadayi, C., Bobrowicz, K., & Osvath, M. (2018). The detour paradigm in animal cognition. *Animal Cognition*, *21*(1), 21–35. https://doi.org/10.1007/s10071-017-1152-0

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Keltner, D., & Lerner, J. (2010). Emotion. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (pp. 317–352). John Wiley & Sons. https://doi.org/10.5840/philstudies19812835

Kihslinger, R. L., Lema, S. C., & Nevitt, G. A. (2006). Environmental rearing conditions produce forebrain differences in wild Chinook salmon Oncorhynchus tshawytscha. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *145*(2), 145–151. https://doi.org/10.1016/j.cbpa.2006.06.041

Kilgour, R. J., & Szantar-Coddington, M. R. (1995). Arena behaviour of ewes selected for superior mothering ability differs from that of unselected ewes. *Animal Reproduction Science*, *37*(2), 133–141. https://doi.org/10.1016/0378-4320(94)01332-G

Knott, S. A., Cummins, L. J., Dunshea, F. R., & Leury, B. J. (2008). Rams with poor feed efficiency are highly responsive to an exogenous adrenocorticotropin hormone (ACTH) challenge. *Domestic Animal Endocrinology*, *34*(3), 261–268. https://doi.org/10.1016/j.domaniend.2007.07.002

Kogan, L. R., Schoenfeld-Tacher, R., & Simon, A. A. (2012). Behavioral effects of auditory stimulation on kenneled dogs. *Journal of Veterinary Behavior: Clinical Applications and Research*, *7*(5), 268–275. https://doi.org/10.1016/j.jveb.2011.11.002

Langbein, J., Siebert, K., & Nuernberg, G. (2008). Concurrent recall of serially learned visual discrimination problems in dwarf goats (Capra hircus). *Behavioural Processes*, *79*(3), 156–164. https://doi.org/10.1016/j.beproc.2008.07.004

Langbein, J. (2012). Investigations on training, recall and reversal learning of a Y-maze by dwarf goats (Capra hircus): the impact of lateralisation. *Behavioural Processes*, *89*(3), 304–310. https://doi.org/10.1016/j.beproc.2011.12.013

Langbein, Jan. (2018). Motor self-regulation in goats (*Capra aegagrus hircus*) in a detour-reaching task. *PeerJ*, *6*, e5139. https://doi.org/10.7717/peerj.5139

Langbein, Jan, Krause, A., & Nawroth, C. (2018). Human-directed behaviour in goats is not affected by short-term positive handling. *Animal Cognition*, *21*(6), 795–803. https://doi.org/10.1007/s10071-018-1211-1

Langbein, Jan, Nürnberg, G., & Manteuffel, G. (2004). Visual discrimination learning in dwarf goats and associated changes in heart rate and heart rate variability. *Physiology and Behavior*, *82*(4), 601–609. https://doi.org/10.1016/j.physbeh.2004.05.007

Langbein, Jan, Nürnberg, G., Puppe, B., & Manteuffel, G. (2006). Self-controlled visual discrimination learning of group-housed dwarf goats (Capra hircus): Behavioral strategies and effects of relocation on learning and memory. *Journal of Comparative Psychology*, *120*(1), 58–66. https://doi.org/10.1037/0735-7036.120.1.58

Langbein, Jan, Siebert, K., Nuernberg, G., & Manteuffel, G. (2007). The impact of acoustical secondary reinforcement during shape discrimination learning of dwarf goats (Capra hircus). *Applied Animal Behaviour Science*, *103*(1–2), 35–44. https://doi.org/10.1016/j.applanim.2006.04.019

Langbein, Jan, Siebert, K., & Nürnberg, G. (2009). On the use of an automated learning device by group-housed dwarf goats: Do goats seek cognitive challenges? *Applied Animal Behaviour Science*, *120*(3–4), 150–158. https://doi.org/10.1016/j.applanim.2009.07.006

Langbein, Jan, Siebert, K., Nürnberg, G., & Manteuffel, G. (2007). Learning to Learn During Visual Discrimination in Group Housed Dwarf Goats (Capra hircus). *Journal of Comparative Psychology*, *121*(4), 447–456. https://doi.org/10.1037/0735-7036.121.4.447

Laukens, D., Brinkman, B. M., Raes, J., De Vos, M., & Vandenabeele, P. (2015). Heterogeneity of the gut microbiome in mice: Guidelines for optimizing experimental design. *FEMS Microbiology Reviews*, *40*(1), 117–132. https://doi.org/10.1093/femsre/fuv036

Le Neindre, P., Poindron, P., Trillat, G., & Orgeur, P. (1993). Influence of breed on reactivity of sheep to humans. *Genetics Selection Evolution*, *25*(5), 447–458. https://doi.org/10.1051/gse:19930504

Leal, M., & Powell, B. J. (2012). Behavioural flexibility and problem-solving in a tropical lizard. *Biology Letters*, *8*, 28–30. https://doi.org/10.1098/rsbl.2011.0480

Leggio, M. G., Mandolesi, L., Federico, F., Spirito, F., Ricci, B., Gelfo, F., & Petrosini, L. (2005). Environmental enrichment promotes improved spatial abilities and enhanced dendritic growth in the rat. *Behavioural Brain Research*, *163*(1), 78–90. https://doi.org/10.1016/j.bbr.2005.04.009

Lensink, B. J., Raussi, S., Boivin, X., Pyykkönen, M., & Veissier, I. (2001). Reactions of calves to handling depend on housing condition and previous experience with humans. *Applied Animal Behaviour Science*, *70*(3), 187–199. https://doi.org/10.1016/S0168-1591(00)00152-0

Lindqvist, C. E. S., Schütz, K. E., & Jensen, P. (2002). Red jungle fowl have more contrafreeloading than White Leghorn layers: Effect of food deprivation and consequences for information gain. *Behaviour*, *139*(9), 1195–1209. https://doi.org/10.1163/15685390260437335

Lindqvist, C., & Jensen, P. (2008). Effects of age, sex and social isolation on contrafreeloading in red junglefowl (Gallus gallus) and White Leghorn fowl. *Applied Animal Behaviour Science*, *114*(3–4), 419–428. https://doi.org/10.1016/j.applanim.2008.03.002

Lindqvist, C., & Jensen, P. (2009). Domestication and stress effects on contrafreeloading and spatial learning performance in red jungle fowl (Gallus gallus) and White Leghorn layers. *Behavioural Processes*, *81*(1), 80–84. https://doi.org/10.1016/j.beproc.2009.02.005

López-Sepulcre, A., De Bona, S., Valkonen, J. K., Umbers, K. D. L., & Mappes, J. (2015). Item Response Trees: A recommended method for analyzing categorical data in behavioral studies. *Behavioral Ecology*, *26*(5), 1268–1273. https://doi.org/10.1093/beheco/arv091

Lu, C. D. (1988). Grazing Behavior and Diet Selection of Goats. *Small Ruminant Research*, *1*, 205–216.

Lyons, D. M. (1989). Individual differences in temperament of dairy goats and the inhibition of milk ejection. *Applied Animal Behaviour Science*, *22*(3–4), 269–282. https://doi.org/10.1016/0168-1591(89)90022-1

Lyons, D. M., & Price, E. O. (1987). Relationships between heart rates and behavior of goats in encounters with people. *Applied Animal Behaviour Science*, *18*(3–4), 363–369. https://doi.org/10.1016/0168-1591(87)90230-9

Lyons, D. M., Price, E. O., & Moberg, G. P. (1988). Individual differences in temperament of domestic dairy goats: constancy and change. *Animal Behaviour*, *36*(5), 1323–1333. https://doi.org/10.1016/S0003-3472(88)80201-X

MacDonald, S. E., & Ritvo, S. (2016). Comparative cognition outside the laboratory. *Comparative Cognition and Behavior Reviews*, *11*(1), 49–61. https://doi.org/10.3819/ccbr.2016.110003

Maier, S. F., & Seligman, M. E. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, *105*(1), 3–46. https://doi.org/10.1037//0096-3445.105.1.3

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, *42*(4), 665–675. https://doi.org/10.1016/j.paid.2006.08.016

Manteuffel, G., Langbein, J., & Puppe, B. (2009). Increasing farm animal welfare by positively motivated instrumental behaviour. *Applied Animal Behaviour Science*, *118*(3–4), 191–198. https://doi.org/10.1016/j.applanim.2009.02.014

Marshall-Pescini, S., Frazzi, C., & Valsecchi, P. (2016). The effect of training and breed group on problem-solving behaviours in dogs. *Animal Cognition*, *19*(3), 571–579. https://doi.org/10.1007/s10071-016-0960-y

Marshall-Pescini, S., Valsecchi, P., Petak, I., Accorsi, P. A., & Previde, E. P. (2008). Does training make you smarter? The effects of training on dogs' performance (Canis familiaris) in a problem solving task. *Behavioural Processes*, *78*(3), 449–454. https://doi.org/10.1016/J.BEPROC.2008.02.022

Mattiello, S., Battini, M., Andreoli, E., Minero, M., Barbieri, S., & Canali, E. (2010). Avoidance distance test in goats: A comparison with its application in cows. *Small Ruminant Research*, *91*(2–3), 215–218. https://doi.org/10.1016/j.smallrumres.2010.03.002

Matzel, L. D., Townsend, D. A., Grossman, H., Han, Y. R., Hale, G., Zappulla, M., … Kolata, S. (2006). Exploration in outbred mice covaries with general learning abilities irrespective of stress reactivity, emotionality, and physical attributes. *Neurobiology of Learning and Memory*, *86*(2), 228–240. https://doi.org/10.1016/j.nlm.2006.03.004

McGowan, R. T. S. S., Robbins, C. T., Alldredge, J. R., & Newberry, R. C. (2010). Contrafreeloading in grizzly bears: Implications for captive foraging enrichment. *Zoo Biology*, *29*(4), 484–502. https://doi.org/10.1002/zoo.20282

McKinley, J., Buchanan-Smith, H. M., Bassett, L., & Morris, K. (2003). Training common marmosets (Callithrix jacchus) to cooperate during routine laboratory procedures: Ease of training and time investment. *Journal of Applied Animal Welfare Science*, *6*(3), 209–220. https://doi.org/10.1207/S15327604JAWS0603_06

Meagher, R. K. (2019). Is boredom an animal welfare concern? *Animal Welfare*, *28*(1), 21–32. https://doi.org/10.7120/09627286.28.1.021

Meagher, Rebecca K., Strazhnik, E., von Keyserlingk, M. A. G., & Weary, D. M. (2020). Assessing the motivation to learn in cattle. *Scientific Reports*, *10:6847*. https://doi.org/10.1038/s41598-020-63848-1

Meehan, C. L., & Mench, J. A. (2007). The challenge of challenge: Can problem solving opportunities enhance animal welfare? *Applied Animal Behaviour Science*, *102*(3–4), 246–261. https://doi.org/10.1016/j.applanim.2006.05.031

Mehlhorn, J., & Caspers, S. (2021). The Effects of Domestication on the Brain and Behavior of the Chicken in the Light of Evolution. *Brain, Behavior and Evolution*, 1–15. https://doi.org/10.1159/000516787

Mellor, D. J. (2015). Enhancing animal welfare by creating opportunities for positive affective engagement. *New Zealand Veterinary Journal*, *63*(1), 3–8. https://doi.org/10.1080/00480169.2014.926799

Mellor, D. J. (2016). Updating animalwelfare thinking: Moving beyond the “five freedoms” towards “A lifeworth living.” *Animals*, *6*(3). https://doi.org/10.3390/ani6030021

Mellor, D. J., Beausoleil, N. J., Littlewood, K. E., McLean, A. N., McGreevy, P. D., Jones, B., & Wilkins, C. (2020). The 2020 five domains model: Including human–animal interactions in assessments of animal welfare. *Animals*, *10*(10), 1–24. https://doi.org/10.3390/ani10101870

Mench, J. (1998). Why it is important to understand animal behavior. *ILAR Journal*, *39*(1), 20–26. https://doi.org/10.1093/ilar.39.1.20

Mendl, M. (1999). Performing under pressure: Stress and cognitive function. *Applied Animal Behaviour Science*, *65*(3), 221–244. https://doi.org/10.1016/S0168-1591(99)00088-X

Mendl, M., Burman, O. H. P., Parker, R. M. A., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, *118*(3–4), 161–181. https://doi.org/10.1016/j.applanim.2009.02.023

Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1696), 2895–2904. https://doi.org/10.1098/rspb.2010.0303

Mendl, M., Laughlin, K., & Hitchcock, D. (1997). Pigs in space: Spatial memory and its susceptibility to interference. *Animal Behaviour*, *54*(6), 1491–1508. https://doi.org/10.1006/anbe.1997.0564

Meyer, S., Puppe, B., & Langbein, J. (2010). Cognitive enrichment in zoo and farm animals - Implications for animal behaviour and welfare. *Berliner Und Munchener Tierarztliche Wochenschrift*, *123*(11–12), 446–44656. https://doi.org/10.2376/0005-9366-123-446

Millot, S., Nilsson, J., Fosseidengen, J. E., Bégout, M. L., Fernö, A., Braithwaite, V. A., & Kristiansen, T. S. (2014). Innovative behaviour in fish: Atlantic cod can learn to use an external tag to manipulate a self-feeder. *Animal Cognition*, *17*(3), 779–785. https://doi.org/10.1007/s10071-013-0710-3

Mitchell, D., Scott, D. W., & Williams, K. D. (1973). Preference for Earned Food. *Behavioral Biology*, *9*, 613–624.

Mohr, E., Langbein, J., & Nürnberg, G. (2002). Heart rate variability: A noninvasive approach to measure stress in calves and cows. *Physiology & Behavior*, *75*(1–2), 251–259. https://doi.org/10.1016/S0031-9384(01)00651-5

Morand-Fehr, P., Hervieu, J., & Sauvant, D. (1980). Contribution à la description de la prise alimentaire de la chèvre. *Reproduction Nutrition Développement*, *20*(5B), 1641–1644.

Morgan, K. N., & Tromborg, C. T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science*, *102*(3–4), 262–302. https://doi.org/10.1016/j.applanim.2006.05.032

Morton, F. B., Lee, P. C., & Buchanan-Smith, H. M. (2013). Taking personality selection bias seriously in animal cognition research: A case study in capuchin monkeys (Sapajus apella). *Animal Cognition*, *16*(4), 677–684. https://doi.org/10.1007/s10071-013-0603-5

Müller, C. A., Riemer, S., Virányi, Z., Huber, L., & Range, F. (2016). Inhibitory control, but not prolonged object-related experience appears to affect physical problem-solving performance of pet dogs. *PLoS ONE*, *11*(2), 1–19. https://doi.org/10.1371/journal.pone.0147753

Murphey, R. M., Duarte, F. A. M., & Torres Penedo, M. C. (1981). Responses of cattle to humans in open spaces: Breed comparisons and approach-avoidance relationships. *Behavior Genetics*, *11*(1), 37–48. https://doi.org/10.1007/BF01065826

Nawroth, C. (2017). Invited review: Socio-cognitive capacities of goats and their impact on human–animal interactions. *Small Ruminant Research*, *150*, 70–75. https://doi.org/10.1016/j.smallrumres.2017.03.005

Nawroth, C., Baciadonna, L., & McElligott, A. G. (2016). Goats learn socially from humans in a spatial problem-solving task. *Animal Behaviour*, *121*, 123–129. https://doi.org/10.1016/j.anbehav.2016.09.004

Nawroth, C., Brett, J. M., & McElligott, A. G. (2016). Goats display audience-dependent human-directed gazing behaviour in a problem-solving task. *Biology Letters*, *12*(7). https://doi.org/10.1098/rsbl.2016.0283

Nawroth, C., Langbein, J., Coulon, M., Gabor, V., Oesterwind, S., Benz-Schwarzburg, J., & von Borell, E. (2019). Farm animal cognition-linking behavior, welfare and ethics. *Frontiers in Veterinary Science*, *6*(FEB), 1–16. https://doi.org/10.3389/fvets.2019.00024

Nawroth, C., von Borell, E., & Langbein, J. (2015). Object permanence in the dwarf goat (Capra aegagrus hircus): Perseveration errors and the tracking of complex movements of hidden objects. *Applied Animal Behaviour Science*, *167*, 20–26. https://doi.org/10.1016/j.applanim.2015.03.010

Nawroth, C., Von Borell, E., & Langbein, J. (2014). Exclusion performance in dwarf goats (Capra aegagrus hircus) and sheep (Ovis orientalis aries). *PLoS ONE*, *9*(4), 1–8. https://doi.org/10.1371/journal.pone.0093534

Neave, H. W., Costa, J. H. C., Weary, D. M., & von Keyserlingk, M. A. G. (2018). Personality is associated with feeding behavior and performance in dairy calves. *Journal of Dairy Science*, *101*(8), 7437–7449. https://doi.org/10.3168/jds.2017-14248

Neuringer, A. (1969). Animals Respond for Food in the Presence of Free Food. *Science*, *166*, 399–401.

Newberry, R. C. (1995). Environmental enrichment: Increasing the biological relevance of captive environments. *Applied Animal Behaviour Science*, *44*(2–4), 229–243. https://doi.org/10.1016/0168-1591(95)00616-Z

Nicol, C. J. (1996). Farm animal cognition. *Animal Science*, *62*(3), 375–391. https://doi.org/10.1017/S1357729800014934

Oesterwind, S., Nürnberg, G., Puppe, B., & Langbein, J. (2016). Impact of structural and cognitive enrichment on the learning performance, behavior and physiology of dwarf goats (Capra aegagrus hircus). *Applied Animal Behaviour Science*, *177*, 34–41. https://doi.org/10.1016/j.applanim.2016.01.006

Ogura, T. (2011). Contrafreeloading and the value of control over visual stimuli in Japanese macaques (Macaca fuscata). *Animal Cognition*, *14*(3), 427–431. https://doi.org/10.1007/s10071-010-0377-y

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin*, *114*(1), 29–51. https://doi.org/10.1037/0033-2909.114.1.29

Osborne, S. R. (1977). The free food (contrafreeloading) phenomenon: A review and analysis. *Animal Leaming & Behavior*, *5*(3), 221–235.

Osthaus, B., Lea, S. E. G., & Slater, A. M. (2003). Training influences problem-solving abilities in dogs (canis lupus familiaris) . *Proceedings of the British Society of Animal Science*, *2003*, 103–103. https://doi.org/10.1017/s175275620001262x

Osthaus, Britta, Marlow, D., & Ducat, P. (2010). Minding the gap: spatial perseveration error in dogs. *Animal Cognition*, *13*(6), 881–885. https://doi.org/10.1007/s10071-010-0331-z

Osthaus, Britta, Proops, L., Hocking, I., & Burden, F. (2013). Spatial cognition and perseveration by horses, donkeys and mules in a simple A-not-B detour task. *Animal Cognition*, *16*(2), 301–305. https://doi.org/10.1007/s10071-012-0589-4

Overington, S. E., Cauchard, L., Côté, K. A., & Lefebvre, L. (2011). Innovative foraging behaviour in birds: What characterizes an innovator? *Behavioural Processes*, *87*(3), 274–285. https://doi.org/10.1016/j.beproc.2011.06.002

Parker, K. J., Buckmaster, C. L., Schatzberg, A. F., & Lyons, D. M. (2004). Prospective investigation of stress inoculation in young monkeys. *Archives of General Psychiatry*, *61*(9), 933–941. https://doi.org/10.1001/archpsyc.61.9.933

Patt, A., Gygax, L., Wechsler, B., Hillmann, E., Palme, R., & Keil, N. M. (2013). Factors influencing the welfare of goats in small established groups during the separation and reintegration of individuals. *Applied Animal Behaviour Science*, *144*(1–2), 63–72. https://doi.org/10.1016/j.applanim.2012.11.009

Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: The utility of a cognitive approach. *Neuroscience and Biobehavioral Reviews*, *29*(3), 469–491. https://doi.org/10.1016/j.neubiorev.2005.01.002

Peterson, R. A., & Cavanaugh, J. E. (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *JOURNAL OF APPLIED STATISTICS*, *2020*, 2312–2327. https://doi.org/10.1080/02664763.2019.1630372

Platt, D. M., & Novak, M. A. (1997). Videostimulation as enrichment for captive rhesus monkeys (Macaca mulatta). *Applied Animal Behaviour Science*, *52*(1–2), 139–155. https://doi.org/10.1016/S0168-1591(96)01093-3

Poole, T. B. (1992). The nature and evolution of behavioural needs in mammals. *Animal Welfare*, *1*(3), 203–220.

Powell, D. M. (1995). Preliminary Evaluation of Environmental. *Animal Welfare*, *4*(November 1995), 361–370.

Price, E. (1999). Behavioral development in animals undergoing domestication. *Applied Animal Behaviour Science*, *65*, 245–271.

Price, E. O. (1984). THE QUARTERLY of BIJOLOGY. *THE QUARTERLY of BIOLOGY*, *59*(1), 1–33.

Price, E. O., & Thos, J. (1980). Behavioral responses to short-term social isolation in sheep and goats. *Applied Animal Ethology*, *6*(4), 331–339. https://doi.org/10.1016/0304-3762(80)90133-9

Prut, L., & Belzung, C. (2003). The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: A review. *European Journal of Pharmacology*, *463*(1–3), 3–33. https://doi.org/10.1016/S0014-2999(03)01272-X

Puppe, B., Ernst, K., Schön, P. C., & Manteuffel, G. (2007). Cognitive enrichment affects behavioural reactivity in domestic pigs. *Applied Animal Behaviour Science*, *105*(1–3), 75–86. https://doi.org/10.1016/j.applanim.2006.05.016

Pyke, G. H., Pulliam, H. R., Charnov, E., & Charnov, E. L. 1977. (1977). *Optimal foraging: A selective review of theory and tests Recommended Citation Optimal foraging: A selective review of theory and tests. Quarterly Review of Biology 52:137-154*. Retrieved from https://digitalrepository.unm.edu/biol_fsp

R Core Team. (2014). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing. Vienna, Austria. Retrieved from http://www.r-project.org/

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from http://www.r-project.org/

Raîche, G., Walls, T. A., Magis, D., Riopel, M., & Blais, J. G. (2013). Non-graphical solutions for Cattell's scree test. *Methodology*, *9*(1), 23–29. https://doi.org/10.1027/1614-2241/a000051

Raoult, C. M. C., Osthaus, B., Hildebrand, A. C. G., McElligott, A. G., & Nawroth, C. (2021). Goats show higher behavioural flexibility than sheep in a spatial detour task. *Royal Society Open Science*, *8*(3), 201627. https://doi.org/10.1098/rsos.201627

Rasmussen, J. H., Rosenberger, K., & Langbein, J. (2020). EasieRR: An open-source software for non-invasive heart rate variability assessment. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13393

Rauw, W. M., Kanis, E., Noordhuizen-Stassen, E. N., & Grommers, F. J. (1998). Undesirable side effects of selection for high production efficiency in farm animals: A review. *Livestock Production Science*, *56*(1), 15–33. https://doi.org/10.1016/S0301-6226(98)00147-X

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(7), 4436–4441. https://doi.org/10.1073/pnas.062041299

Reader, S. M., & Laland, K. N. (2003). *Animal Inovation*. *Animal Innovation* (10th ed.). Oxford: Oxford University Press.

Regolin, L., Vallortigara, G., & Zanforlin, M. (1995). Detour behaviour in the domestic chick: Searching for a disappearing prey or a disappearing social partner. *Animal Behaviour*, *50*(1), 203–211. https://doi.org/10.1006/anbe.1995.0232

Revelle, Wi. (2020). How to: Use the pysch packkage for Factor Analysys and data reduction, 96. Retrieved from https://cran.r-project.org/web/packages/psychTools/vignettes/factor.pdf

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*, 351–357.

Romeyer, A., & Bouissou, M. F. (1992). Assessment of fear reactions in domestic sheep, and influence of breed and rearing conditions. *Applied Animal Behaviour Science*, *34*(1–2), 93–119. https://doi.org/10.1016/S0168-1591(05)80060-7

Rosenberger, K., Simmler, M., Nawroth, C., Langbein, J., & Keil, N. (2020). Goats work for food in a contrafreeloading task. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-78931-w

Ruby, S., & Buchanan-Smith, H. M. (2015). The effects of individual cubicle research on the social interactions and individual behavior of brown capuchin monkeys (Sapajus apella). *American Journal of Primatology*, *77*(10), 1097–1108. https://doi.org/10.1002/ajp.22444

Sadler, L. J., Johnson, A. K., Lonergan, S. M., Nettleton, D., & Dekkers, J. C. M. (2011). The effect of selection for residual feed intake on general behavioral activity and the occurrence of lesions in Yorkshire gilts. *Journal of Animal Science*, *89*(1), 258–266. https://doi.org/10.2527/jas.2009-2595

Salvanes, A. G. V., & Braithwaite, V. A. (2005). Exposure to variable spatial information in the early rearing environment generates asymmetries in social interactions in cod (Gadus morhua).

Behavioral Ecology and Sociobiology, *59*(2), 250–257. https://doi.org/10.1007/s00265-005-0031-x

Salvanes, A. G. V., Moberg, O., Ebbesson, L. O. E., Nilsen, T. O., Jensen, K. H., & Braithwaite, V. A. (2013). Environmental enrichment promotes neural plasticity and cognitive ability in fish. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1767), 1–7. https://doi.org/10.1098/rspb.2013.1331

Sánchez, M. M., Hearn, E. F., Do, D., Rilling, J. K., & Herndon, J. G. (1998). Differential rearing affects corpus callosum size and cognitive function of rhesus monkeys. *Brain Research*, *812*(1–2), 38–49. https://doi.org/10.1016/S0006-8993(98)00857-9

Sasson-Yenor, J., & Powell, D. M. (2019). Assessment of contrafreeloading preferences in giraffe (Giraffa camelopardalis). *Zoo Biology*, *38*(5), 414–423. https://doi.org/10.1002/zoo.21513

Sauer, F. J., Hermann, M., Ramseyer, A., Burger, D., Riemer, S., & Gerber, V. (2019). Effects of breed, management and personality on cortisol reactivity in sport horses. *BioRxiv*, 1–19. https://doi.org/10.1101/739847

Schapiro, S. J., Bloomsmith, M. A., Porter, L. M., & Suarez, S. A. (1996). Enrichment effects on rhesus monkeys successively housed singly, in pairs, and in groups. *Applied Animal Behaviour Science*, *48*(3–4), 159–171. https://doi.org/10.1016/0168-1591(96)01038-6

Schrijver, N. C. A., Pallier, P. N., Brown, V. J., & Würbel, H. (2004). Double dissociation of social and environmental stimulation on spatial learning and reversal learning in rats. *Behavioural Brain Research*, *152*(2), 307–314. https://doi.org/10.1016/j.bbr.2003.10.016

Schütz, K. E., & Jensen, P. (2001). Effects of resource allocation on behavioural strategies: A comparison of red junglefowl (Gallus gallus) and two domesticated breeds of poultry. *Ethology*, *107*(8), 753–765. https://doi.org/10.1046/j.1439-0310.2001.00703.x

Schwabe, L., Wolf, O. T., & Oitzl, M. S. (2010). Memory formation under stress: Quantity and quality. *Neuroscience and Biobehavioral Reviews*, *34*(4), 584–591. https://doi.org/10.1016/j.neubiorev.2009.11.015

Selye, H. (1975). Stress and distress. *Comprehensive Therapy*, *1*(8), 9–13.

Shepherdson, D. J., Mellen, J. D., & Hutchins, M. (1999). *Second Nature: Environmental Enrichment for Captive Animals*. Washington and London: Smithsonian Institution Press.

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour*, *61*(2), 277–286. https://doi.org/10.1006/anbe.2000.1606

Siebert, K., Langbein, J., Schön, P. C., Tuchscherer, A., & Puppe, B. (2011). Degree of social isolation affects behavioural and vocal response patterns in dwarf goats (Capra hircus). *Applied Animal Behaviour Science*, *131*(1–2), 53–62. https://doi.org/10.1016/j.applanim.2011.01.003

Singh, D., & Query, W. T. (1971). Preference for work over “freeloading” in children. *Psychonomic*

Science, *24*(2), 77–79.

Sneddon, L. U., Braithwaite, V. A., & Gentle, M. J. (2003). Novel Object Test: Examining Nociception and Fear in the Rainbow Trout. *Journal of Pain*, *4*(8), 431–440. https://doi.org/10.1067/S1526-5900(03)00717-X

Sol, D., Griffin, A. S., & Bartomeus, I. (2012). Consumer and motor innovation in the common myna: The role of motivation and emotional responses. *Animal Behaviour*, *83*(1), 179–188. https://doi.org/10.1016/j.anbehav.2011.10.024

Souders, D. J., Boot, W. R., Blocker, K., Vitale, T., Roque, N. A., & Charness, N. (2017). Evidence for narrow transfer after short-term cognitive training in older adults. *Frontiers in Aging Neuroscience*, *9*(FEB), 1–10. https://doi.org/10.3389/fnagi.2017.00041

Spinka, M., & Wemelsfelder, F. (2017). Environmental challenge and animal agency. *Animal Welfare*, 27–43. https://doi.org/10.1079/9781845936594.0027

Stephens, D. W., Krebs, J. R. (1986). *Foraging Theory*. Princeton, New Jersey: Princeton University Press.

Studnitz, M., Jensen, M. B., & Pedersen, L. J. (2007). Why do pigs root and in what will they root?. A review on the exploratory behaviour of pigs in relation to environmental enrichment. *Applied Animal Behaviour Science*, *107*(3–4), 183–197. https://doi.org/10.1016/j.applanim.2006.11.013

Svartberg, K. (2002). Shyness-boldness predicts performance in working dogs. *Applied Animal Behaviour Science*, *79*(2), 157–174. https://doi.org/10.1016/S0168-1591(02)00120-X

Tallet, C., Sy, K., Prunier, A., Nowak, R., Boissy, A., & Boivin, X. (2014). Behavioural and physiological reactions of piglets to gentle tactile interactions vary according to their previous experience with humans. *Livestock Science*, *167*(1), 331–341. https://doi.org/10.1016/j.livsci.2014.06.025

Tallet, C., Veissier, I., & Boivin, X. (2009). How does the method used to feed lambs modulate their affinity to their human caregiver? *Applied Animal Behaviour Science*, *119*(1–2), 56–65. https://doi.org/10.1016/j.applanim.2009.03.012

Thornton, A., & Samson, J. (2012). Innovative problem solving in wild meerkats. *Animal Behaviour*, *83*(6), 1459–1468. https://doi.org/10.1016/j.anbehav.2012.03.018

Tia, B., Viaro, R., & Fadiga, L. (2018). Tool-use training temporarily enhances cognitive performance in long-tailed macaques (Macaca fascicularis). *Animal Cognition*, *21*(3), 365–378. https://doi.org/10.1007/s10071-018-1173-3

Toyoshima, M., Yamada, K., Sugita, M., & Ichitani, Y. (2018). Social enrichment improves social recognition memory in male rats. *Animal Cognition*, *21*(3), 345–351. https://doi.org/10.1007/s10071-018-1171-5

Trickett, S. L., Guy, J. H., & Edwards, S. A. (2009). The role of novelty in environmental enrichment for the weaned pig. *Applied Animal Behaviour Science*, *116*(1), 45–51.

https://doi.org/10.1016/j.applanim.2008.07.007

Vacca, G. M., Stocco, G., Dettori, M. L., Pira, E., Bittante, G., & Pazzola, M. (2018). Milk yield, quality, and coagulation properties of 6 breeds of goats: Environmental and individual variability. *Journal of Dairy Science*, *101*(8), 7236–7247. https://doi.org/10.3168/jds.2017-14111

Valenchon, M., Lévy, F., Prunier, A., Moussu, C., Calandreau, L., & Lansade, L. (2013). Stress Modulates Instrumental Learning Performances in Horses (Equus caballus) in Interaction with Temperament. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0062324

Van Horik, J. O., Langley, E. J. G., Whiteside, M. A., Laker, P. R., Beardsworth, C. E., & Madden, J. R. (2018). Do detour tasks provide accurate assays of inhibitory control? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1875). https://doi.org/10.1098/rspb.2018.0150

van Horik, J. O., & Madden, J. R. (2016). A problem with problem solving: Motivational traits, but not cognition, predict success on novel operant foraging tasks. *Animal Behaviour*, *114*, 189–198. https://doi.org/10.1016/j.anbehav.2016.02.006

Van Oers, K., Drent, P. J., De Goede, P., & Van Noordwijk, A. J. (2004). Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1534), 65–73. https://doi.org/10.1098/rspb.2003.2518

Van Os, J. M. C., Mintline, E. M., DeVries, T. J., & Tucker, C. B. (2018). Domestic cattle (Bos taurus taurus) are motivated to obtain forage and demonstrate contrafreeloading. *PLoS ONE*, *13*(3). https://doi.org/10.1371/journal.pone.0193109

Veissier, I., & le Neindre, P. (1992). Reactivity of Aubrac heifers exposed to a novel environment alone or in groups of four. *Applied Animal Behaviour Science*, *33*(1), 11–15. https://doi.org/10.1016/S0168-1591(05)80079-6

Voelkl, B., Altman, N. S., Forsman, A., Forstmeier, W., Gurevitch, J., Jaric, I., … Würbel, H. (2020). Reproducibility of animal research in light of biological variation. *Nature Reviews Neuroscience*, *21*(7), 384–393. https://doi.org/10.1038/s41583-020-0313-3

Voelkl, B., Vogt, L., Sena, E. S., & Würbel, H. (2018). Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biology*, *16*(2). https://doi.org/10.1371/journal.pbio.2003693

Waiblinger, S., Menke, C., & Fölsch, D. W. (2003). Influences on the avoidance and approach behaviour of dairy cows towards humans on 35 farms. *Applied Animal Behaviour Science*, *84*(1), 23–39. https://doi.org/10.1016/S0168-1591(03)00148-5

Washburn, D. A., & Rumbaugh, D. M. (1991). Impaired performance from brief social isolation of rhesus monkeys (Macaca mulatta): a multiple video-task assessment. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *105*(2), 145–151. https://doi.org/10.1037/0735-7036.105.2.145

Wat, K. K. Y., Banks, P. B., & McArthur, C. (2020). Linking animal personality to problem-solving performance in urban common brushtail possums. *Animal Behaviour*, *162*, 35–45. https://doi.org/10.1016/j.anbehav.2020.01.013

Webster, S. J., & Lefebvre, L. (2001). Problem solving and neophobia in a columbiform-passeriform assemblage in Barbados. *Animal Behaviour*, *62*(1), 23–32. https://doi.org/10.1006/anbe.2000.1725

Wechsler, B., & Lea, S. E. G. (2007). Adaptation by learning: Its significance for farm animal husbandry. *Applied Animal Behaviour Science*, *108*(3–4), 197–214. https://doi.org/10.1016/j.applanim.2007.03.012

Wells, D. L. (2004). A review of environmental enrichment for kennelled dogs, Canis familiaris. *Applied Animal Behaviour Science*, *85*(3–4), 307–317. https://doi.org/10.1016/j.applanim.2003.11.005

Wells, D. L. (2009). Sensory stimulation as environmental enrichment for captive animals: A review. *Applied Animal Behaviour Science*, *118*(1–2), 1–11. https://doi.org/10.1016/j.applanim.2009.01.002

Wells, D. L., & Egli, J. M. (2004). The influence of olfactory enrichment on the behaviour of captive black-footed cats, Felis nigripes. *Applied Animal Behaviour Science*, *85*(1–2), 107–119. https://doi.org/10.1016/j.applanim.2003.08.013

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297.

Whitehouse, J., Micheletta, J., Powell, L. E., Bordier, C., & Waller, B. M. (2013). The impact of cognitive testing on the welfare of group housed primates. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0078308

Willmann, R., De Luca, A., Benatar, M., Grounds, M., Dubach, J., Raymackers, J. M., & Nagaraju, K. (2012). Enhancing translation: Guidelines for standard pre-clinical experiments in mdx mice. *Neuromuscular Disorders*, *22*(1), 43–49. https://doi.org/10.1016/j.nmd.2011.04.012

Wischner, D., Kemper, N., & Krieter, J. (2009). Nest-building behaviour in sows and consequences for pig husbandry. *Livestock Science*, *124*(1–3), 1–8. https://doi.org/10.1016/j.livsci.2009.01.015

Wood-Gush, D. G. M., & Vestergaard, K. (1989). Exploratory behavior and the welfare of intensively kept animals. *Journal of Agricultural Ethics*, *2*(2), 161–169. https://doi.org/10.1007/BF01826929

Würbel, H. (2017). More than 3Rs: The importance of scientific validity for harm-benefit analysis of animal research. *Lab Animal*, *46*(4), 164–166. https://doi.org/10.1038/laban.1220

Yamanashi, Y. (2010). Emotional consequences when chimpanzees (Pan troglodytes) face challenges: *Animal Welfare*, *19*(1), 25–30.

Yamanashi, Y., & Hayashi, M. (2011). Assessing the effects of cognitive experiments on the welfare of captive chimpanzees (Pan troglodytes) by direct comparison of activity budget between wild and captive chimpanzees. *American Journal of Primatology*, *73*(12), 1231–1238. https://doi.org/10.1002/ajp.20995

Young, R. J., & Lawrence, A. B. (2003). Do domestic pigs in controlled environments contrafreeload? *Journal of Applied Animal Welfare Science*, *6*(4), 309–318. https://doi.org/10.1207/s15327604jaws0604_5

Zebunke, M., Puppe, B., & Langbein, J. (2013). Effects of cognitive enrichment on behavioural and physiological reactions of pigs. *Physiology and Behavior*, *118*, 70–79. https://doi.org/10.1016/j.physbeh.2013.05.005

Zidar, J., Balogh, A., Favati, A., Jensen, P., Leimar, O., Sorato, E., & Løvlie, H. (2018). The relationship between learning speed and personality is age- and task-dependent in red junglefowl. *Behavioral Ecology and Sociobiology*, *72*(10). https://doi.org/10.1007/s00265-018-2579-2

Zobel, G., & Nawroth, C. (2020). Current state of knowledge on the cognitive capacities of goats and its potential to inform species-specific enrichment. *Small Ruminant Research*. https://doi.org/10.1016/j.smallrumres.2020.106208

Zobel, G., Neave, H. W., & Webster, J. (2019). Understanding natural behavior to improve dairy goat (Capra hircus) management systems. *Translational Animal Science*, *3*(1), 212–224. https://doi.org/10.1093/tas/txy145

Acknowledgements

Danke! Thank you!

Christian, vielen Dank für die großartige und freundschaftliche Zusammenarbeit, für das Teilen deines grossen Wissens zum Training und kognitiven Testen von Nutztieren (besonders Ziegen), und für deine unendliche Geduld und dein stets offenes Ohr. Danke, für die Einführung in die Welt von Twitter und dafür, dass du mir gezeigt hast, wie nützlich Social media in unserem Job sein kann. Von dir durfte ich lernen wie man mit Herz und Seele Forschung betreibt, Netzwerke knüpft und seine Ergebnisse dann in die Welt trägt, (und wie man nebenbei beim Gärtnern entspannt ☺)!

Nina, ich danke dir für die unzähligen Gespräche, wenn ich in der Klemme steckte, und für deine positive und aufmunternde Art. Deine tolle Fähigkeit, den roten Faden in meinen Texten zu finden, hat mich gelehrt, besser und klarer zu schreiben. Durch dich habe ich auch gelernt, dass es immer eine Lösung gibt und, dass man alles schaffen kann, wenn man es nur will!

Jan, danke für die Hilfe und Unterstützung beim Verfassen diverser Manuskripte, besonders beim Schreiben des EasieRR-Manuskriptes. Vielen Dank für deinen wertvollen Beitrag von Know-how in Sachen physiologischer Stressmessung sowie der Aufnahme und Analyse von Herzratenvariabilität. Danke auch, dass ich so nett in Rostock und am FBN willkommen geheissen wurde und die Möglichkeit bekam, das Leben an der Ostsee kennen und lieben zu lernen.

Michi, ich kann dir gar nicht genug danken für deine Unterstützung in statistischen Angelegenheiten, deinen Input zu meinen Manuskripten und deine unglaubliche Geduld und Zeit, die du mir geschenkt hast. Ohne dich wäre ich nie so schnell und mit derselben Qualität meiner Arbeiten ans Ziel gekommen! Merci vielmals!

Jeppe, thank you so much for encouraging me to develop “EasieRR” together, to publish it and for always having my back!! Working together with you never felt like work and was always lots of fun! Thanks for letting me be your roommate, for making so many great memories together, e.g. cooking, sight-seeing, jogging, and introducing me to Denmark.

Camille, ein riesen Dankeschön für die herzliche Aufnahme ins Gästehaus, und die vielen gemeinsamen Ausflüge, die mir den Einstieg ins Leben in der Schweiz sehr erleichtert haben. Danke für deine unglaublich liebenswerte Art, dein allzeit offenes Ohr für Probleme, und deine Hilfsbereitschaft in so vielen Situationen. Ohne dich wäre meine Zeit in Tänikon nicht dieselbe gewesen.

Conny, danke für deine Hilfsbereitschaft in medizinischen Notfällen bei den Geissen (und Lilly ☺), für tolle Gespräche und Ausflüge zusammen und dafür, dass du stets ein Zimmer mit mir geteilt hast bei Konferenzen etc.!

Gästehaus Tänikon, danke an alle Leute (besonders Judith, Ranu, Loris, Andi, Leonie, Sophia, Maria, Mariana), die mich wie Familie in meiner Zeit im Gästehaus aufgenommen haben. Ich werde die vielen tollen Ausflüge, Kochsessions und sportlichen Aktivitäten wie Uni-Hockey niemals vergessen!

Annika K., danke für die vielen wertvollen Infos und das Teilen deines Know-hows betreffend Aufnahme und Analyse von EKG-Daten, deine Zeit für meine Fragen und für deine fröhliche hilfsbereite Umgangsweise.

Katrin S., danke vielmals für die tolle Unterstützung bei der Auswertung meiner Videos in Observer, für deine Hilfe mit Excel-Problemen, deine Aushilfe, wenn es Not an der Frau gab und die Vermietung deiner schönen Wohnung in Rostock, die ich mit vielen unvergesslichen Momenten verbinde.

Richi, danke für deine Hilfe mit technischen und handwerklichen Problemen, guten Gesprächen und der ständigen Bereitstellung von Gummistiefeln ;).

Christina R., vielen herzlichen Dank für deine kurzfristige Bereitschaft meine Doktorarbeit zu lesen, deine Flexibilität und deinen wertvollen Input und Vorschläge zur Verbesserung!

Jana, danke für deine grossartige Unterstützung als studentische Hilfskraft bei der Datenaufnahme in Dummerstorf und deine Freundschaft. Deine unfassbar positive Art hat mir so manchen Tag verschönert, und die Arbeit mit dir immer lustig gemacht. Danke auch, dass du (und Anna) mich in die «Surf-Community» eingeführt habt!

Charlotte, thank you so much for your help with ECG data analysis as a DAAD intern. I really enjoyed sharing my apartment in Rostock with you and I will never forget the many times we went stand-up paddling together, our bike trips to the beach and gardening together.

Heather, thank you so much for your help with the analysis of behavioural data obtained from stress tests and with PCA analysis. You not only proofed to be a great tutor during my MSc internship in Canada, but also a great friend and support during my PhD. Thanks for teaching me so much about personality tests and PCA analysis!

Philipp, ohne deine Hingabe bei der Auswertung der MSR Daten mit R, wären sie wohl heute noch unberührt auf meinem PC. Danke für deine Motivation und Selbstständigkeit bei der Arbeit und die Aufnahme von «hilfsbedürftigen» Ziegen!

Barbara, danke für die Unterstützung und den Beistand in vielen Situationen, egal ob per Telefon oder Whatsapp, du warst immer erreichbar und die schönen gemeinsamen Abendessen und Ausflüge. Danke auch für deine Hilfe beim «Babysitting» von Lilly, wenn es mal Not an der Frau gab!

Roxy, danke für deine offene Art, dein fröhliches Gemüt und deine Hilfsbereitschaft. Dank deiner Kompetenz und Hingabe weiss ich «meine» Ziegen in guten Händen.

Mama und Papa, danke, dass ihr immer an mich geglaubt habt, mich meinen Weg gehen lassen und mich dabei immer finanziell unterstützt habt. Ich weiss, die vielen Abschiede im Laufe meines Studiums waren schwierig, aber ihr habt mich trotzdem immer bestärkt, das zu machen was ich liebe und mir ein «Dahoam» geschenkt, wenn ich es brauchte.

Rafi, danke für die Besuche mit Andi in der Schweiz, die Mitbringsel aus der «Hoamat», für die vielen Telefonate und für dein offenes Ohr, wenn es mir schlecht ging/geht. Besser eine grossartige Schwester in der Ferne als eine schlechte in der Nähe ;)

Lara, danke dir für dein grosses Verständnis, deine Bewunderung für mich und deine Bestärkung in meinen Entscheidungen. Es tut mir leid, dass du mich durch meine Berufswahl nur wenig in deiner Nähe hattest.

Oma R., Opa D. und Oma D., ich möchte euch für all die mentale und grosszügige finanzielle Unterstützung, sowie viele Geschenke aus der Heimat danken. Ohne euch hätte sich mein ganzes Studium wohl um einiges schwieriger gestaltet und ich wäre nicht da, wo ich jetzt bin!

Yves, danke für deine Geduld in stressigen Zeiten, für deinen Zuspruch, wenn ich am Verzweifeln war und für dein unendliches Verständnis für «PhD-bedingte» Situationen. Selbst als ich ins 1000-kmentfernte Rostock ziehen musste, nachdem wir uns gerade erst kennen gelernt hatten, hast du dich nicht beschwert und mich immer besucht und unterstützt. Danke für deine Liebe und Loyalität!

Appendix I for Chapter II

Goats work for food in a contrafreeloading task.

Table S1: IRTree GLMM of behavioural responses. Intercepts correspond to Trial 1 (see Methods). This table is an extended version of Table 1 in the paper. Results with p-value ≤ 0.05 are given in bold.

	Node 1: Participation *p* participation								**Node 2: Door type** *p* closed door							
	Dwarf				Dairy				Dwarf				Dairy			
	est.	s.e.	z	p	est.	s.e.	z	p	est.	s.e.	z	p	est.	s.e.	z	p
Intercept	**2.43**	**0.64**	**3.77**	**<0.001**	**3.32**	**0.69**	**4.82**	**<0.001**	**-0.85**	**0.29**	**-2.92**	**0.003**	-0.27	0.26	-1.05	0.29
(Trial -1)	0.03	0.07	0.39	0.69	**-0.15**	**0.07**	**-2.07**	**0.04**	**0.11**	**0.05**	**2.14**	**0.03**	0.00	0.05	0.03	0.97
	Dwarf - Dairy								**Dwarf – Dairy**							
Contrast in…	est.		s.e.		z		p		est.		s.e.		Z		p	
Intercept	0.90		0.90		0.99		0.32		0.58		0.38		1.51		0.13	
(Trial – 1)	-0.18		0.10		-1.77		0.08		-0.11		0.07		-1.55		0.12	
	Node 3: Approach time closed door *p* long approach time								**Node 4: Approach time open door** *p* long approach time							
	Dwarf				Dairy				Dwarf				Dairy			
	est.	s.e.	z	p	est.	s.e.	z	p	est.	s.e.	z	p	est.	s.e.	z	P
Intercept	0.25	0.78	0.32	0.75	-0.72	0.65	-1.11	0.27	0.55	0.48	1.15	0.25	**0.99**	**0.48**	**2.08**	**0.04**
(Trial -1)	0.07	0.10	0.71	0.48	0.04	0.08	0.57	0.57	**-0.21**	**0.08**	**-2.65**	**0.008**	-0.07	0.07	-0.95	0.34
	Dwarf - Dairy								**Dwarf - Dairy**							
Contrast in…	est.		s.e.		z		p		est.		s.e.		z		p	
Intercept	-0.98		1.01		-0.97		0.33		0.45		0.66		0.68		0.50	
(Trial – 1)	-0.03		0.12		-0.21		0.84		0.15		0.10		1.38		0.17	

Table S2: Random effect variance components and correlations of the IRTree GLMM.

Random effects						
Groups	**Name**	**Variance**	**Std. Dev.**	**Corr**		
Obs	(Intercept)	0.00	0.00			
ClosedSide:(Individual:Pen)	NodeN1	0.25	0.50			
	NodeN2	0.22	0.47	0.34		
	NodeN3	0.09	0.30	0.33	-0.77	
	NodeN4	0.15	0.39	1.00	0.39	0.28
Individual:Pen	NodeN1	2.90	1.70			
	NodeN2	0.00	0.0.03	1.00		
	NodeN3	1.00	1.00	-0.33	-0.33	
	NodeN4	1.14	1.07	-0.77	-0.77	0.86
Pen	Node 1	0.36	0.60			
	Node 2	0.01	0.10	1.00		
	Node 3	0.64	0.80	1.00	1.00	
	Node 4	0.10	0.31	1.00	1.00	1.00

Table S3: Random effect variance components of the LMM with continuous approach time as response.

Random effects			
Groups	**Name**	**Variance**	**Std. Dev.**
ClosedSide:(Individual:Pen)	(Intercept)	0.02	0.14
Individual:Pen	(Intercept)	0.13	0.36
Pen	(Intercept)	0.03	0.18
Residual		0.53	0.73

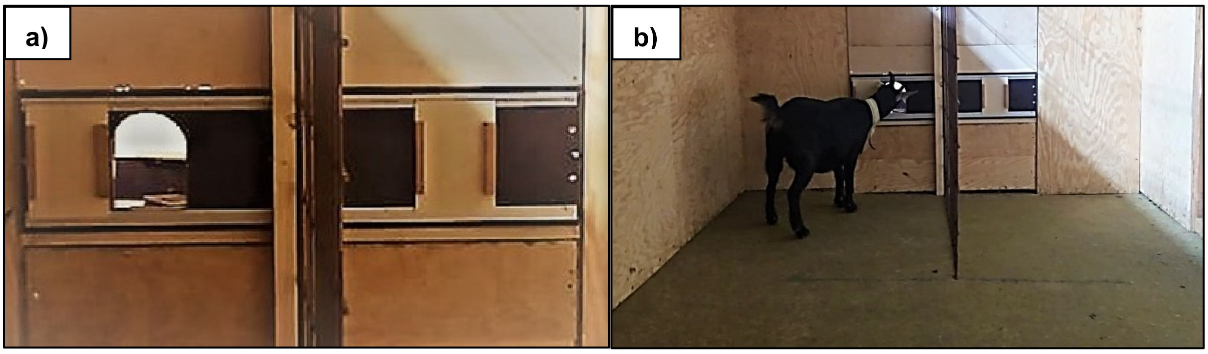


Figure S1: a) Test apparatus with left sliding door open (= free reward) and right sliding door closed (= work). b) Test setup
with dwarf goat feeding from the open door.

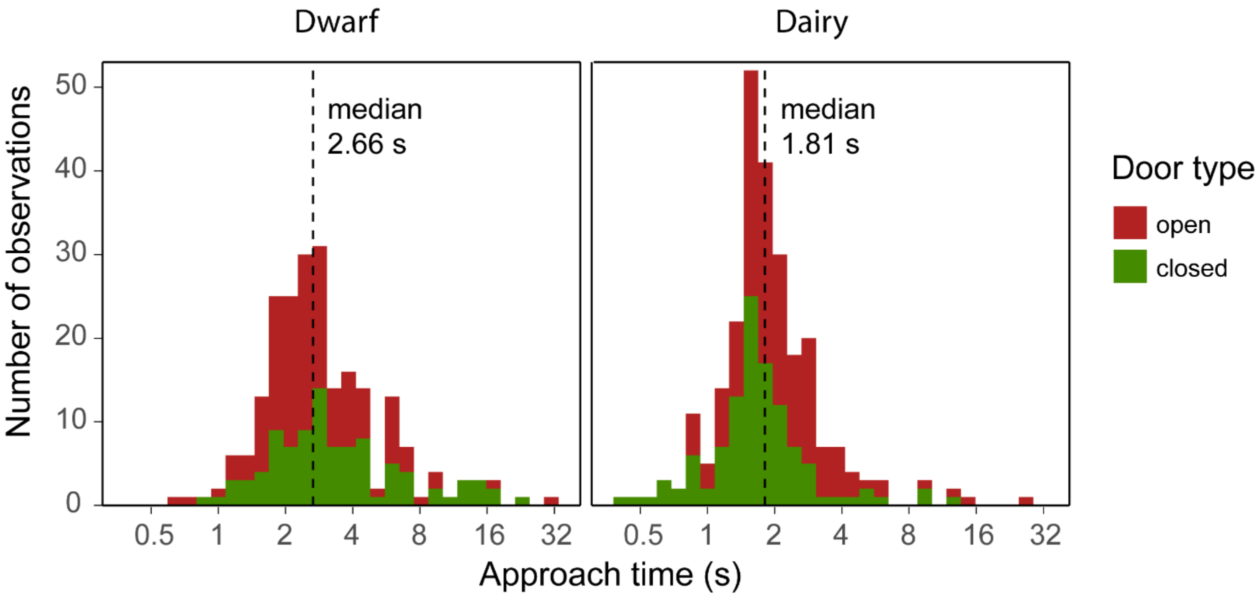


Figure S2: Distribution of approach times of dwarf and dairy goats towards the open and closed doors.

Appendix II for Chapter III

Reactivity of domesticated goats towards different stressors following long-term cognitive test exposure

Description of habituation to isolation and shaping to the cognitive test procedures

All animals were habituated to the test arena at first the whole group of a pen (2 d; 20 min per day), then in groups of two (4–6 d), and finally individually (7–13 d). For the COG and POS treatment, the habituation continued until each animal confidently took a reward (a piece of uncooked pasta) from the sliding board. Only COG goats continued with the following 'shaping' phase (10 trials per session) where the animals were trained to indicate a choice. A plant saucer (2 cm high, Ø 14 cm) with a reward was placed in the middle of the sliding board (4 trials). As soon as the animal stuck its nose through the middle of the grid, it received the reward from the experimenter. In the next step, the reward was covered with a plant pot (light brown, 10.5 cm high, Ø 12 cm) before the animal could make a choice (6 trials). The goal of the shaping phase was achieved, if at the end of the shaping phase, the animals showed no signs of stress during the manipulation of and feeding from the sliding board. This phase was followed by further training sessions of 10 trials, where 2 pots with saucers were positioned on the left and right side of the sliding board (30 cm distance). The experimenter baited, visible to the goat, one of the saucers, and then covered both saucers with pots. The goat received the reward only if it chose the baited pot. The rewarded pot was presented in a pseudo-randomised manner on each side. The training criterion required that the goat chose the baited side in at least 8 out of 10 trials over two consecutive sessions to proceed to the cognitive tests (see Supplementary Table S1).

Table S1: Detailed description of handling and cognitive tests used in the COG, POS and ISO treatments. NA = not applicable.

Treatment	Test name	Task description (all of type object choice)	Rewards	Time isolated in test arena	Sessions (1 session per day)
COG 1st phase	Visual-discrimination task	Object choice between a black or a white cup	Dependent on number of trials solved, but a max. of 14 (= 12 trials plus 2 motivational trials)	Until all trials solved or a max. of 10 minutes	20 sessions of 12 trials
	Reversal-learning task	Object choice with reversed cup colours			
COG 2nd phase	Cognitive test battery: Direct visual exclusion	Of two cups, only one is baited with food, but both are lifted providing direct visual information regarding the location of a food reward	Dependent on number of trials solved, but a max. of 14 (= 12 trials plus 2 motivational trials)	Until all trials solved or a max. of 10 minutes	12 sessions of 12 trials
	Cognitive test battery: Indirect visual exclusion	Of two cups, only one is baited, and only the empty cup is lifted providing indirect visual information regarding the location of a food reward			
	Cognitive test battery: Direct auditory exclusion	Of two cups, only one is baited, but both are shaken providing direct auditory information regarding the location of a food reward			
	Cognitive test battery: Indirect auditory exclusion	Of two cups, only one is shaken providing indirect auditory information regarding the location of a food reward			
	Cognitive test battery: Transposition task	One of two cups is baited in full view of the individual before the left cup is moved to the right side and the right cup to the left side			
COG 3rd phase	Cognitive test battery: Sustained pointing	The experimenter is positioned in the middle between two cups and points at the baited cup until individual makes a choice	Dependent on number of trials solved, but a max. of 14 (= 12 trials plus 2 motivational trials)	Until all trials solved or a max. of 10 minutes	12 sessions of 12 trials
	Cognitive test battery: Momentary pointing	The experimenter is positioned in the middle between two cups and will point at the baited cup for about one second			

	Cognitive test battery: Pointing vs. body position	The experimenter is positioned behind the empty cup and points with ipsilateral arm and finger at the baited cup			
	Cognitive test battery: Body orientation	The experimenter is positioned in the middle between two cups and orients body and head towards the baited cup			
	Cognitive test battery: Marking of correct cup	The experimenter is positioned in the middle between two cups and places a marker on top of the baited cup			
POS	NA	NA	Median reward delivery of all ‘test’ individuals from previous test session	Mean test time of all ‘test’ individuals from previous test session	20 sessions
ISO	NA	NA	No rewards	Mean test time of all ‘test’ individuals from previous test session	20 sessions

Table S2: LMER model and random effect variance components for the rotated component "active" in the novel arena test (NA).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.13	0.52	1.00	-0.24	0.50	0.98
POS	0.30	0.51	0.94	-0.23	0.50	0.98
ISO	-0.11	0.50	1.00	0.13	0.49	1.00
Random effects	*variance*	*sd*	*n groups*			
Pen	0.13	0.36	12			
Site	0.36	0.60	2			
Residual	0.73	0.85	87			
Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.44	0.35	0.22	0.01	0.31	0.98
ISO - POS	-0.41	0.33	0.22	0.36	0.30	0.22
ISO - COG	0.03	0.33	0.94	0.37	0.30	0.22

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.11	0.40	0.789	-0.53	0.39	0.17	0.24	0.37	<0.001

Table S3: LMER model and random effect variance components for the rotated component "reactive to isolation" in the novel arena test (NA).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.53	0.29	0.30	0.49	0.26	0.28
POS	-0.41	0.28	0.57	0.48	0.26	0.29
ISO	-0.91	0.27	0.005	0.48	0.25	0.28
Random effects	*variance*	*sd*	*n groups*			
Pen	0.19	0.44	12			
Site	0.00	0.00	2			
Residual	0.55	0.74	87			
Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.12	0.31	0.69	-0.01	0.27	0.96
ISO - POS	-0.50	0.29	0.08	0.00	0.26	0.99
ISO - COG	-0.38	0.29	0.19	-0.02	0.26	0.95

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	1.02	0.39	0.009	0.89	0.38	0.02	1.38	0.37	<0.001

Table S4: LMER model and random effect variance components for the rotated component "active" in the novel object test (NO).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	0.30	0.35	0.91	-0.36	0.32	0.75
POS	0.65	0.37	0.31	-0.53	0.33	0.39
ISO	0.38	0.33	0.71	-0.20	0.32	0.98
Random effects	*variance*	*sd*	*n groups*			
Pen	0.05	0.22	12			
Site	0.10	0.32	2			
Residual	0.77	0.88	89			
Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.35	0.38	0.35	-0.17	0.31	0.57
ISO - POS	-0.27	0.36	0.45	0.34	0.30	0.27
ISO - COG	0.09	0.34	0.80	0.16	0.30	0.58

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.66	0.36	0.06	-1.18	0.38	<0.001	-0.58	0.33	0.08

Table S5: LMER model and random effect variance components for the rotated component "explorative" in the novel object test (NO).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.35	0.37	0.84	0.13	0.35	1.00
POS	0.44	0.39	0.71	0.29	0.35	0.90
ISO	-0.62	0.35	0.28	0.31	0.34	0.86
Random effects	*variance*	*sd*	*n groups*			
Pen	0.00	0.00	12			
Site	0.14	0.38	2			
Residual	0.83	0.91	89			
Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*

POS - COG	0.80	0.39	0.04	0.17	0.32	0.60
ISO - POS	-1.06	0.37	0.00	0.02	0.31	0.96
ISO - COG	-0.27	0.35	0.44	0.18	0.31	0.55

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	0.48	0.34	0.16	-0.15	0.37	0.68	0.93	0.31	<0.001

Table S6: LMER model and random effect variance components for the rotated component “sociable towards a novel human” in the novel human test (NH).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.23	0.69	0.97	0.14	0.69	1.00
POS	-0.32	0.69	0.89	0.11	0.69	1.00
ISO	-0.51	0.69	0.68	0.26	0.68	0.94
Random effects	*variance*	*sd*	*n groups*			
Pen	0.02	0.12	12			
Site	0.87	0.93	2			
Residual	0.49	0.70	89			
Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	-0.09	0.28	0.73	-0.03	0.26	0.92
ISO - POS	-0.19	0.26	0.46	0.15	0.25	0.55
ISO - COG	-0.28	0.27	0.29	0.12	0.24	0.61

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	0.37	0.28	0.184	0.44	0.27	0.11	0.78	0.25	0.002

Table S7: LMER model and random effect variance components for the rotated component “active” in the novel human test (NH).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.44	0.39	0.65	0.33	0.37	0.82
POS	-0.53	0.38	0.46	0.71	0.38	0.20
ISO	-0.59	0.37	0.33	0.60	0.37	0.31
Random effects	*variance*	*sd*	*n groups*			
Pen	0.03	0.16	12			
Site	0.18	0.43	2			
Residual	0.64	0.80	89			

Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	-0.09	0.32	0.78	0.38	0.29	0.20
ISO - POS	-0.07	0.29	0.82	-0.10	0.29	0.72
ISO - COG	-0.16	0.31	0.61	0.27	0.28	0.33

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	0.77	0.32	0.02	1.23	0.32	<0.001	1.20	0.29	<0.001

Table S8: LMER model and random effect variance components for the rotated component “reactive towards handling” in the weighing test (WH).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	-0.04	0.55	1.00	-0.47	0.54	0.71
POS	0.74	0.55	0.38	0.01	0.55	1.00
ISO	0.03	0.54	1.00	0.05	0.54	1.00

Random effects	*variance*	*sd*	*n groups*
Pen	0.06	0.25	12
Site	0.49	0.70	2
Residual	0.64	0.80	94

Treament contrasts	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est*	*s.e.*	*p*
POS - COG	0.78	0.31	0.01	0.48	0.28	0.09
ISO - POS	-0.71	0.30	0.02	0.04	0.28	0.90
ISO - COG	0.06	0.29	0.83	0.52	0.27	0.06

Selection line contrasts	*COG*			*POS*			*ISO*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*
Dairy-Dwarf	-0.43	0.32	0.18	-0.73	0.33	0.03	0.02	0.31	0.95

Table S9: LMER model and random effect variance components for the variable “entering score” in the weighing Test (WH).

Fixed effects	*Dwarf*			*Dairy*		
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*
COG	2.77	0.24	<0.001	2.51	0.22	<0.001
POS	2.73	0.25	<0.001	2.38	0.23	<0.001
ISO	2.58	0.23	<0.001	2.28	0.21	<0.001

Random effects	*variance*		*sd*		*n groups*					
Pen	0.00		0.00		12					
Site	0.00		0.00		2					
Residual	0.82		0.91		94					
Treament contrasts	*Dwarf*						*Dairy*			
	est.		*s.e.*		*p*		*est*		*s.e.*	*p*
POS - COG	-0.04		0.35		0.92		-0.14		0.32	0.66
ISO - POS	-0.15		0.34		0.65		-0.10		0.31	0.76
ISO - COG	-0.19		0.33		0.57		-0.24		0.31	0.44
Selection line contrasts	*COG*			*POS*			*ISO*			
	est.	*s.e.*	*p*	*est.*	*s.e.*	*p*	*est.*	*s.e.*	*p*	
Dairy-Dwarf	-0.25	0.33	0.44	-0.36	0.34	0.29	-0.30	0.31	0.33	

Appendix III for Chapter IV

Performance of goats in a detour and a problem-solving test following long-term cognitive test exposure

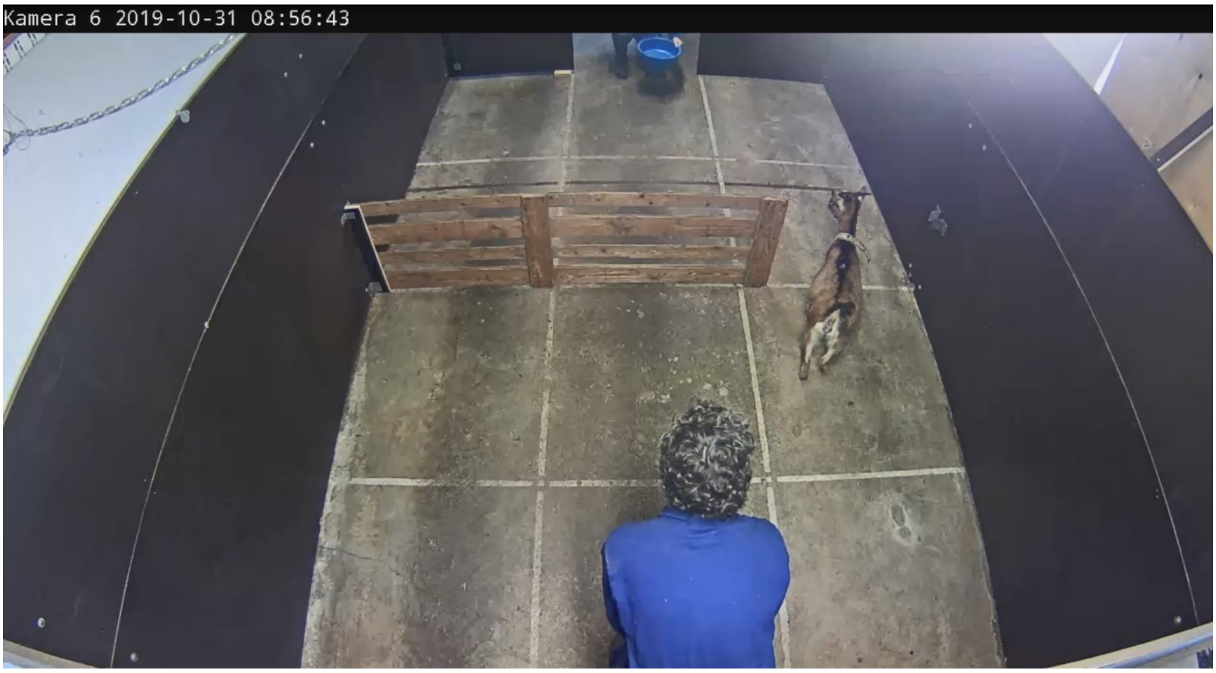


Figure S1: Camera view (in Dummerstorf) of test arena with dwarf goat moving from Experimenter 1 to Experimenter 2 while crossing the fence.

Table S2: GLMER model for Accuracy in the four A trials and the four B trials.

Fixed effects									
A trials					**B trials**				
	est.	**s.e.**	***z***	***p***		**est.**	***s.e.***	***z***	**p**
COG	2.41	0.72	3.36	<0.001	COG	-0.24	0.49	-0.50	0.62
POS	2.62	0.82	3.21	<0.001	POS	0.63	0.51	1.23	0.22
ISO	2.56	0.72	3.56	<0.001	ISO	0.49	0.48	1.01	0.31
COG:I(Trial-1)	1.09	0.65	1.67	0.09	COG:I(Trial-1)	0.73	0.22	3.27	<0.001
POS:I(Trial-1)	15.87	193.52	0.08	0.93	POS:I(Trial-1)	0.41	0.23	1.80	0.07
ISO:I(Trial-1)	1.09	0.65	1.67	0.09	ISO:I(Trial-1)	0.38	0.21	1.82	0.07
Contrast in…	**est.**	**s.e.**	**z**	**p**	**Contrast in…**	**est.**	**s.e.**	**z**	**p**
POS-COG	0.21	1.00	0.21	0.83	POS-COG	0.87	0.58	1.50	0.13
ISO-COG	0.15	0.92	0.17	0.87	ISO-COG	0.73	0.56	1.30	0.19
ISO-POS	-0.06	1.00	-0.06	0.96	ISO-POS	-0.14	0.58	-0.24	0.81

Table S3: Random-effect variance components of the GLMER model for Accuracy.

Random effects			
Groups	**Name**	**Variance**	**Std.Dev.**
Individual	(Intercept)	0.58	0.76
Pen	(Intercept)	0.00	0.00
Site	(Intercept)	0.04	0.20
SelectionLine	(Intercept)	0.12	0.35

Table S4: LMER model for Latency to cross the fence in A trials and B trials.

Fixed Effects									
A trials					**B trials**				
	est.	**s.e.**	***z***	***p***		**est.**	***s.e.***	***z***	**p**
COG	1.90	0.26	7.40	<0.001	COG	2.86	0.26	11.13	<0.001
POS	1.83	0.26	6.91	<0.001	POS	2.47	0.26	9.32	<0.001
ISO	1.85	0.25	7.28	<0.001	ISO	2.74	0.25	10.77	<0.001
COG:I(Trial-1)	-0.10	0.09	-1.07	0.28	COG:I(Trial-1)	-0.06	0.09	-0.68	0.50
POS:I(Trial-1)	-0.14	0.10	-1.44	0.15	POS:I(Trial-1)	0.02	0.10	0.17	0.86
ISO:I(Trial-1)	0.09	0.09	0.97	0.33	ISO:I(Trial-1)	-0.08	0.09	-0.82	0.41
Contrast in…	**est.**	**s.e.**	**z**	**p**	**Contrast in…**	**est.**	**s.e.**	**z**	**p**
POS-COG	-0.08	0.31	-0.25	0.80	POS-COG	-0.40	0.31	-1.31	0.19
ISO-COG	-0.06	0.30	-0.20	0.84	ISO-COG	-0.13	0.30	-0.43	0.67
ISO-POS	0.02	0.30	0.05	0.96	ISO-POS	0.27	0.30	0.89	0.37

Table S5: Random-effect variance components of the LMER model for Latency to cross the fence.

Random effect			
Groups	**Name**	**Variance**	**SD**
Individual	(Intercept)	0.35	0.59
Pen	(Intercept)	0.01	0.12
Site	(Intercept)	0.00	0.00

SelectionLine	(Intercept)	0.04	0.20
Residual		1.15	1.07

Table S6: GLMER model for *Touched*.

Fixed effect				
	est.	**s.e.**	***z***	***p***
COG	0.86	1.37	0.62	0.53
POS	1.52	1.38	1.11	0.27
ISO	1.83	1.42	1.29	0.20
COG:I(Trial-1)	−0.28	0.18	−1.57	0.12
POS:I(Trial-1)	−0.70	0.20	−3.58	<0.001
ISO:I(Trial-1)	−0.63	0.21	−2.99	0.003
Contrast in…				
POS-COG	0.67	0.75	0.90	0.37
ISO-COG	0.98	0.81	1.21	0.23
ISO-POS	0.31	0.82	0.38	0.71

Table S7: Random-effect variance components of the GLMER model for *Touched*.

Random effect			
Groups	**Name**	**Variance**	**SD**
Individual	(Intercept)	1.20	1.10
Pen	(Intercept)	0.40	0.64
Site	(Intercept)	0.85	0.92
SelectionLine	(Intercept)	2.28	1.51

Table S8: GLMER model for *Opened* (dairy goats only).

Fixed effect				
	est.	**s.e.**	***z***	***p***
COG	−2.85	3.59	−0.79	0.43
POS	−3.86	3.55	−1.09	0.28
ISO	−1.51	3.52	−0.43	0.67
COG:I(Trial-1)	0.00	0.37	0.00	1.00
POS:I(Trial-1)	0.64	0.36	1.77	0.08
ISO:I(Trial-1)	0.51	0.38	1.33	0.18
Contrast in…				
POS-COG	−1.01	2.23	−0.45	0.65
ISO-COG	1.34	2.45	0.55	0.58
ISO-POS	2.35	2.39	0.98	0.33

Table S9: Random-effect variance components of the GLMER model for *Opened* (dairy goats only).

Random effect			

Groups	Name	Variance	SD
Individual	(Intercept)	15.79	3.97
Pen	(Intercept)	1.61	1.27
Site	(Intercept)	18.35	4.28

Table S10: LMER model for *LatencyT*.

Fixed effect				
	est.	**s.e.**	***z***	***p***
COG	2.90	0.62	4.71	<0.001
POS	2.56	0.61	4.19	<0.001
ISO	2.88	0.61	4.69	<0.001
COG:I(Trial-1)	0.12	0.06	1.95	0.05
POS:I(Trial-1)	0.21	0.06	3.25	0.001
ISO:I(Trial-1)	0.16	0.06	2.47	0.01
Contrast in…				
POS-COG	−0.34	0.25	−1.34	0.18
ISO-COG	−0.02	0.26	−0.08	0.93
ISO-POS	0.31	0.25	1.26	0.21

Table S11: Random-effect variance components of the LMER model for *LatencyT*.

Random effect			
Groups	**Name**	**Variance**	**SD**
Individual	(Intercept)	0.23	0.48
Pen	(Intercept)	0.01	0.11
Site	(Intercept)	0.13	0.36
SelectionLine	(Intercept)	0.55	0.74
Residual		0.56	0.75

Table S12: LMER model for *LatencyO* (dairy goats only).

Fixed effect				
	est.	**s.e.**	***z***	***p***
COG	3.79	0.49	7.77	<0.001
POS	4.18	0.50	8.34	<0.001
ISO	3.34	0.47	7.03	<0.001
COG:I(Trial-1)	−0.35	0.10	−3.49	<0.001
POS:I(Trial-1)	−0.51	0.11	−4.72	<0.001
ISO:I(Trial-1)	−0.24	0.10	−2.37	0.02
Contrast in…				
POS-COG	0.38	0.39	0.99	0.32
ISO-COG	−0.46	0.37	−1.22	0.22
ISO-POS	−0.84	0.39	−2.13	0.03

Table S13: Random-effect variance components of the LMER model for *LatencyO* (dairy goats only).

Random effect			
Groups	**Name**	**Variance**	**SD**
Individual	(Intercept)	0.04	0.19
Pen	(Intercept)	0	0
Site	(Intercept)	0.29	0.55
Residual		0.71	0.84

Appendix IV

“EasieRR: An open-source software for non-invasive heart rate variability assessment”

Jeppe Have Rasmussen[1,2*], Katrina Rosenberger[3*], and Jan Langbein[2]

[*]These authors contributed equally to this paper.

[1]Texas A&M University Galveston, Department of Marine Biology, Galveston, Texas, USA

[2]Leibniz Institut for Farm Animal Biology, Institute of Behavioural Physiology, Dummerstorf, Germany

[3]Center for Proper Housing of Ruminants and Pigs, Agroscope Tänikon, Switzerland

Published in Methods in Ecology and Evolution (2020), https://doi.org/10.1111/2041-210X.1339

Author contributions

JHR has written the EasieRR software code in Matlab, reviewed the literature, proposed an article outline, wrote the initial draft, and revised KR and JL’s later article drafts. KR wrote mainly the methods and GUI overview sections of the manuscript, made the tables and figures, revised drafts following feedback from JHR and JL and alpha-tested the software for bugs. JL reviewed the literature, wrote the introduction, and revised KR and JR’s drafts of the manuscript.

Summary

1. The assessment of heart rate (HR) and heart rate variability (HRV) based on electrocardiograms (ECG) is considered a good proxy for stress in a wide range of animal species. However, problems can occur e.g., when measuring ECG in ambulatory settings such as recording of unrestrained animals using non-invasive devices. Artefacts caused by technical (i.e. bad electrode contact) or physiological (i.e. ectopic beats, non-cardiac muscle potentials) sources are common and can disturb the ECG signal. As HRV analysis is highly sensitive to artefacts in the interbeat interval (RR-interval) time series the process of visual inspection of the raw signal to detect and correct these is essential. Most of the commercially available software requires intensive training and extensive manual work to accomplish this task and/or is often not available to access for free.
2. EasieRR is an open-source, stand-alone software optimized for analysing ECG in non-restrained animals. The program allows a species-specific analysis and calculation of recommended standard HRV parameters in both, the time- and the non-linear domain (RMSSD, SDNN, SD1, and SD2).
3. Visualization of data by using Poincaré plots and tachograms of RR-intervals eases the validation of correct heart cycle interval detection and minimizes manual work for the user. Automatically detected peaks can be manually corrected via deletion, correction of spurious detections or marking of undetected peaks.
4. The HRV analysis can be exported using common formats (TXT, MAT). Figures can be plotted and exported in various formats (PDF, SVG, PNG, JPG, TIFF, EMF, EPS).
5. Included in EasieRR is the possibility for synchronization of ECG data with video to link cardiac responses to specific behavioural responses.

Introduction

Measures of heart rate (HR) and heart rate variability (HRV) have a long history as indicators of cardiac health, and stress in humans (Fleisher, 1996, Delaney and Brodie 2000, Hall et al. 2004, Carney, Freedland & Veith, 2005), and have gained considerable interest in animal behavioural studies during the last 30 years (Hopster, Werf & Blokhuis., 1998, von Borell, 2000, Mohr et al., 2002, Kovács et al., 2014). HRV is based on the analysis of normal fluctuations in the time intervals of consecutive heartbeats = interbeat intervals (IBIs) or RR-intervals. It reflects the interplay between sympathetic (SNS) and parasympathetic nervous system (PNS) and thus also provides information about cardiac vagal tone (Porges, 1995, von Borell et al., 2007). Analysing HRV is regarded as a suitable approach to determine the activity of the autonomic nervous system (ANS) in the context of stress, affect and emotion (Appelhans & Leucken 2006, Boissy et al., 2007, Thayer et al. 2012), and is progressively emerging as a suitable indicator of welfare states in farm animal research.

Further, animals of the same species cope differently with environmental challenges. As the ANS has major impact on regulation of fundamental physiological functions related to coping and stress resilience, many of these phenotypic differences in stress response are mediated by different activation of the ANS. ANS activity in different autonomic phenotypes of free-living streaked shearwater (*Calonectris leucomelas*) was linked to stress level (Muller et al. 2018), consistent across time and context. Evans et al. (2016) used HR and HRV to investigate the impact of ecological mechanisms on the hibernation process in free-ranging brown bears. These studies have speculated ANS may be a key mechanism driving phenotypic variation in animal populations and is therefore a potentially important mediator in the evolution of life history (Evans et al., 2016, Muller et al., 2018).

The process of artefact correction of the ECG signal has been shown to be essential for the appropriate analysis of HRV data (Berntson & Stowell, 1998; Shaffer & Combatalade, 2013). This makes HRV analysis especially challenging for data derived from unrestrained animals where, due to technical or physiological interference, a high occurrence of artefacts is likely. However, even when recording ECG during resting states, visual inspection is important to detect artefacts like atrioventricular blocks (AV blocks type 1-3), due to disturbances in impulse conduction at the heart. These types of artefacts can be relatively commonly found in horses, but also in other animals and if not excluded, can result in very high HRV. In veterinary and behavioural animal research, human heart rate monitors, such as POLAR ® devices (POLAR Electro Oy, Helsinki, Finland) has been widely used in unrestrained animals. The major disadvantage of many of these devices has been that only RR-intervals were extracted but the ECG itself was not accessible, allowing no visual inspection of peaks detected. The implemented routines for the correction of artefacts are often not comprehensible or transparent.

Recently, advanced technologies used in medical science capable of assessing the ECG from electrical differences on the body surface (e.g., with a sensor attached via a belt) have found application in animal science (e.g., BioHarness™, Zephyr Technology Corporation, Annapolis, MD, U.S.A.). This progression in technology has not only made raw ECG data accessible, but also confronts researchers working on physiology and behaviour in unrestrained animals with the challenge of correct pre-processing of data before HRV analysis.

Commercial HRV analysis programs such as AcqKnowledge 4.4, (BIOPAC System Inc.), Nevrokard® (Nevrokard, Izola/Slovenia), or Kubios HRV Premium (University of Eastern Finland, Kuopio/Finland; Niskanen, Tarvainen, Rantaaho, & Karjalainen, 2004), but also Open-Source software such as Artiifact (Kaufmann, Sütterlin, Schulz & Vögele, 2011) are available for HRV analysis (see Table1). However, these programs require experience in the field and can be time-consuming to learn to use correctly. In this article, we describe an open-source software with an intuitive graphical user interface (GUI) that imports ECG from any csv file or txt files exported from AcqKnowledge software, detects the prominent peaks, makes correction of spurious detections easy and transparent, and calculates a variety of HRV parameters in the time- and nonlinear domain.

Table 1: Common software available for ECG processing and HRV analysis compared to EasieRR.

Software name	Supported domains of HRV analysis	Supported data formats	Peak detection algorithm	Artefact detection and processing	Analysis reports and export of results	Availability
Kubios HRV (Premium)	Time, frequency, non-linear domain	Inter-beat-interval data from HR monitors or text file, ECG data*	QRS detection based on Pan-Tompkins	Automatic (threshold-based) artefact detection algorithm, Cubic spline interpolation, manual correction	Text file (import into MS Excel or SPSS), MAT-file, PDF report, SPSS-friendly batch process	Standard version: open access/ premium version: purchase via Kubios website

AcqKnowledge	Time, frequency, non-linear domain	NA (software only reads its own data)	Modified Pan-Tompkins QRS detector; ECG QRS Peak event	Cubic spline interpolation, manual correction using waveform transformations	HRV statistics as MS Excel file, Poincaré plot as Acq file	Purchase via BIOPAC website
ARTiiFact	Time, frequency domain	Inter-beat-interval data, ECG data	Global or local threshold detection	Distribution based artefact detection algorithm, manual correction: deletion, linear interpolation, cubic spline interpolation; batch processing,	MS Excel file, text file, PDF report, batch processing possible	Open-access, download via ARTiiFact website
EasieRR	Time, non-linear domain	ECG data	Peak prominence	Manual correction: deletion, insertion or movement of peaks	Text file and MAT-file, Figures in all commonly used formats	Open-source, download via figshare or Github

*only for Kubios HRV Premium version

EasieRR: Computational background and theory

EasieRR is an open-source software developed to assist researchers in the use of heart rate parameters and their processing and analysis. Special emphasis was put on an intuitive GUI to ease detection and manual correction of artefacts.

EasieRR has been programmed using MATLAB 2018b (Mathworks, Natick, MA, USA) and compiled as a stand-alone application with the MATLAB compiler 7.0 for Microsoft Windows operating systems (Version 7 and upwards). EasieRR is hosted at https://figshare.com/projects/EasieRR/68831 and distributed under the terms of the GNU General Public License.

Peak detection

EasieRR applies by default a band pass filter (4. Order 0-degree Butterworth filter with low cut-off at 1 Hz and high cut-off at 20 Hz), hence removing high frequency peaks in the signal as well as removing any voltage offset. These two cut-off frequencies have been shown to be applicable for HRV analysis in goats but can be modified by the user for optimal use on other species. The extraction of time intervals between heart beats is traditionally done by detecting R-peaks (RR-intervals), because they are usually the most distinct peaks (Lippmann et al. 1994). However, if variability within the heart rate cycles is calculated, other peaks and thus intervals i.e., S-S intervals can be used in case they are more prominent. It is essential that the type of peak used for detection is consistent within and between all individuals of an experiment (personal communication, Berntson). There are various algorithms for R-peak detection applied within different programs, i.e., global/local threshold detection (Berntson, Quigley, Jang & Boysen, 1990), Pan-Tompkins QRS detector (Pan & Tompkins, 1985) or template matching (Friesen et al., 1990). In EasieRR, the peak detection used for heart cycle interval determination is based on the peak prominence. This allows for robust R-peak detection when combined with the band pass filtering and a predetermined minimum timespan between RR-intervals (i.e., heart refractory time of the species used).

Artefact detection and processing

Currently available software uses different approaches to correct artefacts such as deletion of spurious RR-intervals, interpolation of missing or extra beats, i.e., cubic spline interpolation, linear interpolation, but also manual correction (see Table 1). To maximize transparency and reliability, EasieRR does not detect artefacts automatically. Instead, the GUI is specifically designed for easy visual examination and detection of artefacts. For this purpose, EasieRR is displaying the raw (and band pass filtered) ECG signal, a tachogram as well as a Poincaré plot simultaneously. Significant deviations between the lengths of successive interbeat intervals are often caused by spurious peak detection and artefacts. These divergent intervals are easily recognisable as distinct peaks in the tachogram and can also be outliers in the typically elliptical shaped Poincaré plot. Figure 1.1-1.3 displays in detail the procedure of artefact detection and correction using the 3 correction options: move mark, insert mark, mark outlier.

In contrast to most existing software which use interpolation algorithms for outlier correction (i.e., Kubios), EasieRR is based on a deletion algorithm. This method has been shown to be best suited for time-domain measures (Rincon Soler et al. 2015) and analysis of shorter sequences of ECG recordings (Lippmann, Stein & Lerman, 1994). The latter is often the case when data is obtained in ambulatory settings. Besides deleting outliers, EasieRR also allows selection of missed peaks and movement of misplaced marks.

a)

weigh_F4_test.txt

Marker size: 40

T min (s): 15.753
T max (s): 23.493
Mean HR (bpm): 124.321
RMSSD (ms): 236.830
Outlier percentage: 0.000

Time: ☐ Fixed range: 20
Time range (s): 7.740
SD1 (ms): 174.302
SD2 (ms): 105.210
SDNN (ms): 138.766

b)

weigh_F4_test.txt

Marker size: 40

T min (s): 15.753
T max (s): 23.493
Mean HR (bpm): 116.017
RMSSD (ms): 38.846
Outlier percentage: 0.000

Time: ☐ Fixed range: 20
Time range (s): 7.740
SD1 (ms): 28.590
SD2 (ms): 79.277
SDNN (ms): 58.336

Figure 1.1: Example of artefact correction using “mark outlier” button (Cycle detection method=negative peak: correction is displayed on red marks and lines only). a) Before correction: Circles mark signs of spuriously detected peak. b) After correction: Circles mark location of changes in ECG (deleted data point marked with “x” and tachogram (colored yellow for deleted data). Outliers have disappeared in Poincaré plot.

a)

Undetected peak

Outlier

weigh_D9_test.txt

Marker size: 60

T min (s): 2.620

T max (s): 22.620

Mean HR (bpm): 201.673

RMSSD (ms): 50.675

Outlier percentage: 1.563

Time: ☑ Fixed range: 20

Time range (s): 20.000

SD1 (ms): 36.112

SD2 (ms): 44.018

SDNN (ms): 40.172

b)

Correction: Insert mark

weigh_D9_test.txt

Marker size: 60

T min (s): 2.620

T max (s): 22.620

Mean HR (bpm): 203.481

RMSSD (ms): 13.537

Outlier percentage: 1.538

Time: ☑ Fixed range: 20

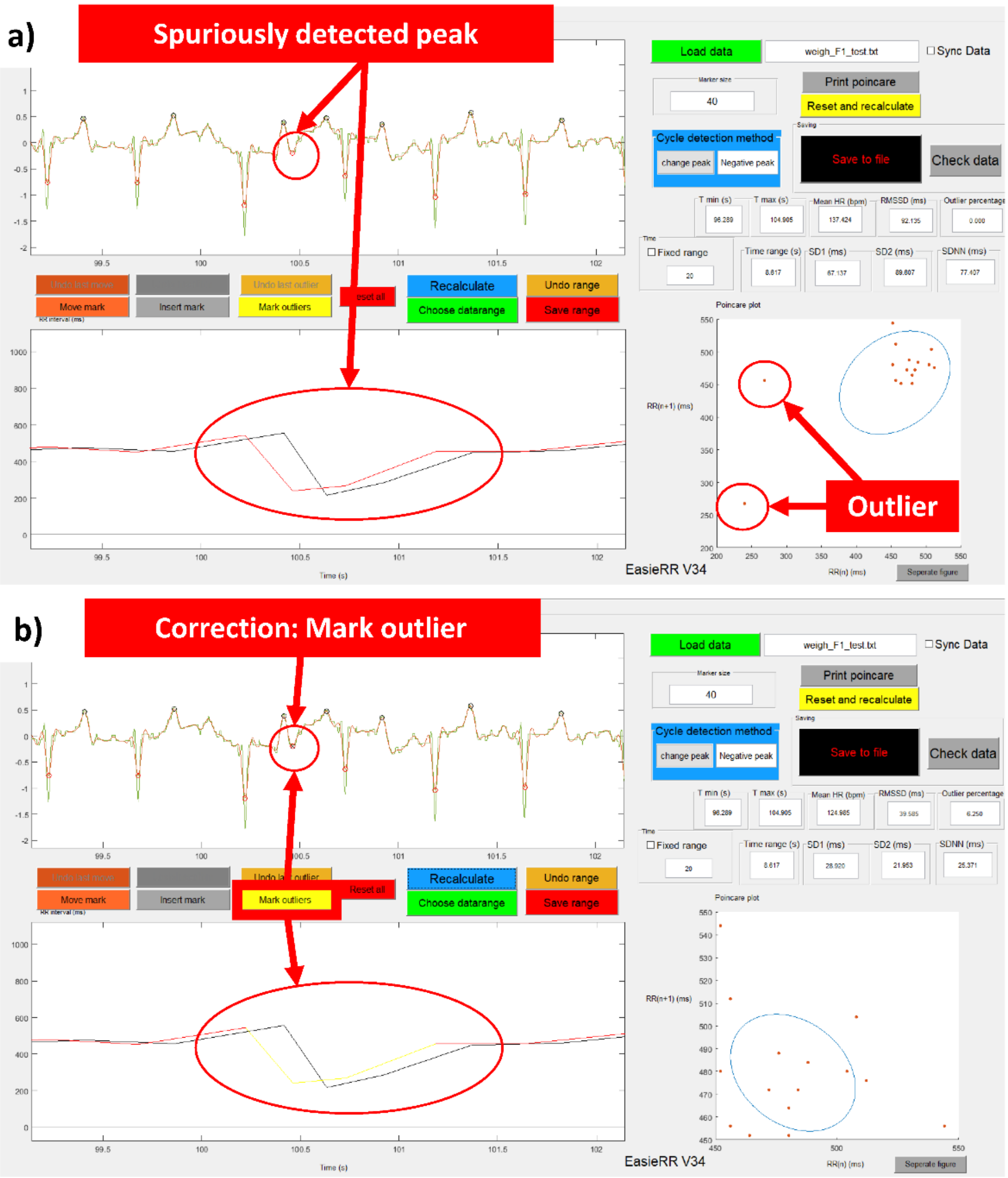
Time range (s): 20.000

SD1 (ms): 9.618

SD2 (ms): 29.661

SDNN (ms): 22.302

Figure 1.2: Example of artefact correction using “insert mark” button (Cycle detection method=negative peak: correction is displayed on red marks and lines only). a) Before correction: Circles mark signs of undetected negative peak in ECG, tachogram and Poincaré plot. b) After correction: Circles mark location of changes in ECG (“+” for new data point) and tachogram (dotted line for new data). Outliers have disappeared in Poincaré plot.



a) Spuriously detected peak

Outlier

weigh_F1_test.txt

Marker size 40

Cycle detection method: Negative peak

T min (s) 96.289
T max (s) 104.905
Mean HR (bpm) 137.424
RMSSD (ms) 92.135
Outlier percentage 0.000
Fixed range 20
Time range (s) 8.817
SD1 (ms) 67.137
SD2 (ms) 89.807
SDNN (ms) 77.407

b) Correction: Mark outlier

weigh_F1_test.txt

Marker size 40

Cycle detection method: Negative peak

T min (s) 96.289
T max (s) 104.905
Mean HR (bpm) 124.985
RMSSD (ms) 39.585
Outlier percentage 6.250
Fixed range 20
Time range (s) 8.617
SD1 (ms) 28.920
SD2 (ms) 21.953
SDNN (ms) 25.371

Figure 1.3: Example of artefact correction using “move mark” button (Cycle detection method=negative peak: correction is displayed on red marks and lines only). a) Before correction: Circles mark signs of spuriously detected negative peak. b) After correction: Circles mark location of changes in ECG (“+” for moved data point, “x” for deleted data point) and tachogram (dotted line for new data). Poincaré plot is a more defined ellipse.

Analysis in the time and non-linear domain

Assessment of HRV has been commonly done using time-domain parameters which are based on the calculation of successive RR-intervals, also called normal-to-normal (NN) intervals. EasieRR calculates the time-domain measures SDNN (Standard deviation of NN-intervals), RMSSD (Root mean square of successive differences between normal heartbeats) and mean HR (heart rate). While HR is a good indicator for overall arousal or activity, it does not allow to draw inferences on the activity of the two autonomic nervous branches, the PNS and the SNS. In contrast, RMSSD has been found to reflect PNS-mediated HRV and can quantify the instantaneous beat-to-beat variance in heart rate. SDNN rather reflects the long-term variability of beat-to-beat intervals and is usually interpreted as an indicator of the sympatho-vagal balance (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996).

In recent years, the use of non-linear methods has been gaining more interest. Non-linear methods like the Poincaré plot make it possible to include nonstationary data in the analysis (Guzik et al., 2007). A Poincaré plot is a scatter plot where each RR-interval is plotted as a function of the previous RR-interval. This typically results in an elliptical shaped scatter plot tilted 45 degrees counterclockwise relative to the x-axis (see Fig. 2). From these data points, it is possible to calculate the standard deviation in two dimensions: Along the minor axis of the ellipse (SD1) and the major axis (SD2). SD1 is describing the instantaneous variability of the RR-interval time series (see RMSSD) and reflecting parasympathetic efferent activity at the sinus node. SD2 is describing the long-term variability of the RR-interval time series. Commonly these two standard deviations are visualized within the Poincaré plot via an ellipse where the minor axis illustrates the SD1 and the major axis illustrates SD2. Thus, the major advantages of using Poincaré plots is their suitability as a quantitative visual tool allowing immediate recognition of artefacts (Myers, Workman, Birkett, Ferguson & Kienzle, 1992) as well as estimation of the activity of the PNS (Woo, Stevenson, Moser & Middlekauff, 1994; Kamen, Krum & Tonkin, 1996).

To validate EasieRR, we analysed different HRV parameters (Table 2) using identical ECG data lasting 20 sec for the programs described in Table 1. While all programs calculated almost identical values for HR and mean-RR, there are small deviations in the time-domain and non-linear parameters. This is probably due to the different underlying algorithms for artefact correction used in the different programs.

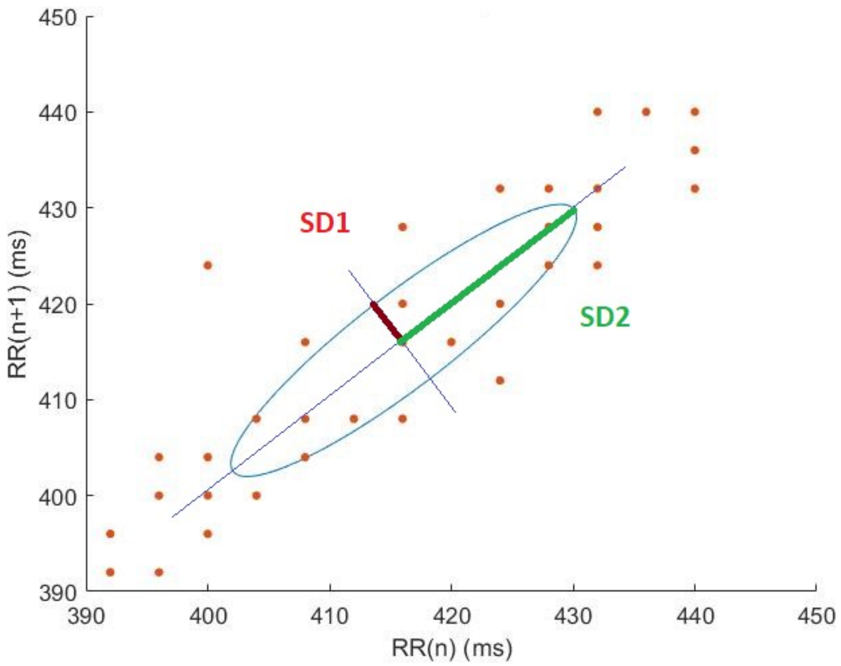


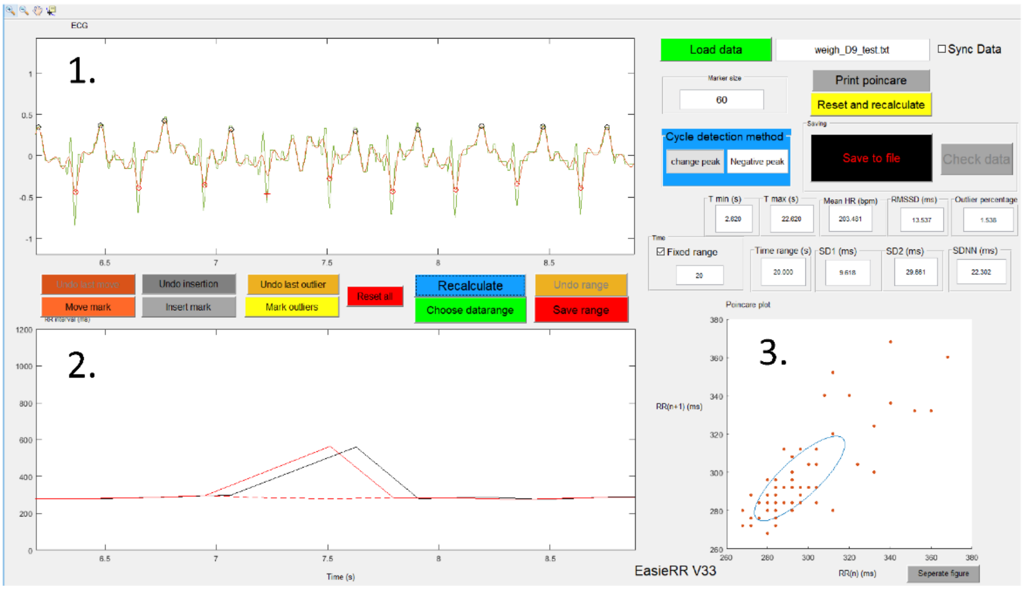
Figure 2: Poincaré plot of RR-data series with typical elliptical distribution and SD1 (red) and SD2 (green).

Table 2: Comparison of basic HRV parameters calculated using EasieRR and other commonly used HRV software. (ECG data was the same for all software; X= not available in the software analysis)

Software name	imported data format	bpm	mean RR(ms)	RMSSD	SD1	SD2	SDNN
EasieRR	ECG	188.57	318.18	7.57	5.37	23.60	17.13
Kubios HRV Premium	ECG	188.25	318.71	5.45	3.86	21.91	15.72
AcqKnowledge	ECG	188.02	319.11	7.64	X	X	17.13
ARTiiFact	ECG	188.63	318.08	10.53	X	X	17.87
Kubios	RR data series, generated with Kubios Premium	188.77	318.17	6.69	X	X	17.07

Program description

The GUI is split into three separate interactive windows (Fig. 3): an upper window displaying the ECG signal (raw and filtered) which allows to visually inspect automatically detected QRS complexes and to correct artefacts. The lower left window shows the corresponding tachogram facilitating the efficient location of artefacts and a Poincare plot at the lower right.



weigh_D9_test.txt

Marker size: 60

T min (s): 2.620

T max (s): 22.620

Mean HR (bpm): 203.481

RMSSD (ms): 13.537

Outlier percentage: 1.538

Time: ☑ Fixed range: 20

Time range (s): 20.000

SD1 (ms): 9.618

SD2 (ms): 29.661

SDNN (ms): 22.302

Figure 3: Graphical user interface of EasieRR displaying (1) raw (and filtered) ECG signal from a goat, (2) tachogram and (3) Poincaré plot.

The EasieRR GUI is operated via a range of buttons. The ‘Choose data range’ button allows the user to select a sequence to be analysed. Time-domain as well as non-linear parameters are then automatically calculated and the corresponding Poincaré plot is generated in the lower right window. For artefact correction, the GUI offers 3 buttons with the options of either manually mark (delete) artefacts, move marks of spuriously detected peaks, or insert marks for missed peaks. Their corresponding ‘Undo’ buttons allow the user to reverse their last action. After each artefact correction, the ‘Recalculate’ button will create a new tachogram as well as Poincaré plot. After artefact correction is finished the data range can be saved using the ‘save range’ button and viewed with the ‘Check data’ button. When analysis is finalized, pressing the ‘save to file’ button will save all analysed time ranges.

A more detailed user manual can be found in the documentation hosted here: (https://figshare.com/projects/EasieRR/68831).

Synchronization with behaviour

HRV data can also be synchronized with observed behaviour via import it into common software for behavioural video analysis (e.g., The Observer®). More information is provided in the user manual available online.

Conclusion and future directions

EasieRR is a free software for evaluating ECG in non-restrained animals and allows for calculation of recommended standard HRV parameters in both, the time- and the non-linear domains. The intuitive GUI facilitates the detection and correction of artefacts through the visualization of RR-intervals in tachogram and Poincaré plot.

Acknowledgements

We thank Annika Krause for sharing her knowledge on ECG acquisition, interpretation, and analysis and Katrin Siebert for help with The Observer and data import. We are grateful for suggestions from people from University of Bern and FBN Dummerstorf, and Charlotte Hursey for beta-testing previous versions. Finally, we want to thank the editor and reviewers for their constructive comments.

References

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. Review of general psychology, 10(3), 229-240. doi: 10.1037/1089-2680.10.3.229

Berntson, G. G., Quigley, K. S., Jang, J. F., & Boysen, S. T. (1990). An approach to artifact identification: Application to heart period data. Psychophysiology, 27(5), 586-598. doi: 10.1111/j.1469-8986.1990.tb01982.

Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don't miss a beat! Psychophysiology, 35(1), 127-132.

Boissy, A., Arnould, C., Chaillou, E., Désiré, L., Duvaux-Ponter, C., Greiveldinger, L., ... & Meunier-Salaün, M. C. (2007). Emotions and cognition: a new approach to animal welfare. Animal Welfare. 16(2), 37-43.

Carney, R. M., Freedland, K. E., Veith, R. C. (2005). Depression, the autonomic nervoussystem, and coronary heart disease. Psychosomatic Medicine, 67, 29–33. doi: 10.1097/01.psy.0000162254.61556.d5

Carrasco, M.J., Gaitán, R., González, O., Yánez, S. (2001). Correlation among Poincare plot indexes and time and frequency domain measures of heart rate variability. Journal of medical engineering & technology, 25(6), 240-248.

Delaney, J. P. A., & Brodie, D. A. (2000). Effects of Short-Term Psychological Stress on the Time and Frequency Domains of Heart-Rate Variability. Perceptual and Motor Skills, 91(2), 515–524. doi: 10.2466/pms.2000.91.2.515

Evans, A. L., Singh, N. J., Friebe, A., Arnemo, J. M., Laske, T. G., Fröbert, O., … Blanc, S. (2016). Drivers of hibernation in the brown bear. Frontiers in Zoology, 13, 13. doi: 10.1186/s12983-016-0140-6Fleisher, L. A. (1996). Heart rate variability as an assessment of cardiovascular status. Journal of cardiothoracic and vascular anesthesia, 10(5), 659-671. doi: 10.1016/S1053-0770(96)80146-7

Friesen, G. M., Jannett, T. C., Jadallah, M. A., Yates, S. L., Quint, S. R., & Nagle, H. T. (1990). A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Transactions on biomedical engineering, 37(1), 85-98. doi: 10.1109/10.43620

Germán-Salló, Z., & Germán-Salló, M. (2016). Non-linear methods in HRV analysis. Procedia technology, 22, 645-651. doi: 10.1016/j.protcy.2016.01.134

Guzik, P., Piskorski, J., Krauze, T., Schneider, R., Wesseling, K., Wykretowicz, A., Wysocki, H. (2007). Correlations between the Poincaré Plot and Conventional Heart Rate Variability Parameters Assessed during Paced Breathing. The journal of physiological sciences, 57, 63-71. doi:10.2170/physiolsci.RP005506

Hall, M., Vasko, R., Buysse, D., Ombao, H., Chen, Q., Cashmere, J. D., ... & Thayer, J. F. (2004). Acute stress affects heart rate variability during sleep. Psychosomatic medicine, 66(1), 56-62. doi: 10.1097/01.PSY.0000106884.58744.09.

Hopster, H., van der Werf, J. T. N., Blokhuis, H. J. (1998). Side preference of dairy cows in the milking parlour and its effects on behaviour and heart rate during milking. Applied Animal Behaviour Science, 55, 213– 29. doi: 10.1016/S0168-1591(97)00064-6

Hoshi, R. A., Pastre, C. M., Vanderlei, L. C. M., & Godoy, M. F. (2013). Poincaré plot indexes of heart rate variability: relationships with other nonlinear variables. Autonomic Neuroscience, 177(2), 271-274.

Hsu, C. H., Tsai, M. Y., Huang, G. S., Lin, T. C., Chen, K. P., Ho, S. T., ... & Li, C. Y. (2012). Poincaré plot indexes of heart rate variability detect dynamic autonomic modulation during general anesthesia induction. Acta Anaesthesiologica Taiwanica, 50(1), 12-18.

Kaufmann, T., Sütterlin, S., Schulz, S. M., & Vögele, C. (2011). ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behavior research methods, 43 (4), 1161-1170. doi: 10.3758/s13428-011-0107-7

Kamen, P. W., Krum, H., & Tonkin, A. M. (1996). Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans. Clinical science, 91(2), 201-208.

Kovács, L., Tözsér, J., Szenci, O., Póti, P., Kézér, F., Ruff, F., Gábriel-Tözsér, G., Hoffmann, D., Bakony M. & Jurkovich, V., (2014). Cardiac responses to palpation per rectum in lactating and nonlactating dairy cows. Journal of Dairy Science, 97 (11), 6955-6963. doi: 10.3168/jds.2013-7030Lippman, N. E. A. L., Stein, K. M. & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. American Journal of Physiology-Heart and Circulatory Physiology, 267(1), H411-H418. doi: 10.1152/ajpheart.1994.267.1.H411

Myers, G., Workman, M., Birkett, C., Ferguson, D., & Kienzle, M. (1992). Problems in measuring heart rate variability of patients with congestive heart failure. Journal of Electrocardiology, 25, 214-219. doi: 10.1016/0022-0736(92)90105-9

Mohr, E., Langbein, J. & Nürnberg, G. (2002). Heart rate variability: A noninvasive approach to measure stress in calves and cows. Physiology & Behavior, 75, 251–259.

Muller, M. S., Vyssotski, A. L., Yamamoto, M., & Yoda, K. (2018). Individual differences in heart rate reveal a broad range of autonomic phenotypes in a free-living seabird population. Journal of Experimental Biology, 221(19), 6. doi: 10.1242/jeb.182758

Niskanen, J. P., Tarvainen, M. P., Ranta-Aho, P. O., & Karjalainen, P. A. (2004). Software for advanced HRV analysis. Computer methods and programs in biomedicine, 76(1), 73-81. doi: 10.1016/j.cmpb.2004.03.004

Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering, 32(3), 230-236.

Porges, S. W. (1995). Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. Psychophysiology, 32(4), 301-318.

Rincon Soler, A. I., Silva, L. E. V., Fazan Jr., R., & Murta Jr., L. O. (2017). The impact of artifact correction methods of RR series on heart rate variability parameters. Journal of Applied Physiology, 124(3), 646-652. doi:10.1152/japplphysiol.00927.2016.

Shaffer, F., & Combatalade, D. C. (2013). Don't add or miss a beat: A guide to cleaner heart rate variability recordings. Biofeedback, 41(3), 121-130. doi: 10.5298/1081-5937-41.3.04

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers III, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neuroscience & Biobehavioral Reviews, 36 (2), 747-756. doi:10.1016/j.neubiorev.2011.11.009

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. Circulation, 93, 1043–1065.

Von Borell, E. (2000). Stress and coping in farm animals. Archiv für Tierzucht, 43, 144–152.

Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., ... & Valance, D. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—A review. Physiology & Behavior, 92(3), 293-316. doi:10.1016/j.physbeh.2007.01.007

Woo, M. A., Stevenson, W. G., Moser, D. K., & Middlekauff, H. R. (1994). Complex heart rate variability and serum norepinephrine levels in patients with advanced heart failure. Journal of the American College of Cardiology, 23(3), 565-569. doi: 10.1016/0735-1097(94)90737-4

Curriculum vitae

MSc, Katrina Rosenberger

Birth date	12 March, 1991
Birth place	Amstetten, Austria
Nationality	Austrian
Address	Tänikon 7, 8356 Ettenhausen
Telephone	+41 (0) 76 2981415
E mail	katrinarosenberger91@gmail.com

Core research

Animal Behaviour, Animal Welfare, Human-animal interactions, Behavioural Physiology, Applied Ethology

My main research interests are animal behaviour and cognition, human-animal interactions/conflicts, and animal welfare assessment. I am interested in the effect of human-animal interactions on human and animal welfare as well as the physiological processes shaping it. I am further interested in how we can use knowledge about behaviour and cognitive capacities of animals to not only adapt the environment they live in, but also inform the public as well as policy makers to improve animal welfare. My profile at: https://www.researchgate.net/profile/Catriona-Ro

Education

2017 – 2021	**PhD candidate** in Biomedical Sciences, Vetsuisse Faculty, Bern, Switzerland Thesis: Effect of long-term cognitive training on behaviour and welfare of goats.
2013 – 2016	**Master of Science** in Human-Animal Interactions, University of Veterinary Medicine, Vienna, Austria. Thesis: Effect of milk allowance on behaviour, growth and welfare of dairy calves.
2010 - 2013	**Bachelor of Science** in Biology (major: Zoology), University of Vienna, Austria Thesis: Locomotion behaviour in chicken *(Gallus gallus).*

Research Experience

2017 – 2021 **Research assistant** at the Centre for Proper Housing of Ruminants and Pigs, Agroscope, Tänikon, Switzerland
2018 – 2019 **Research assistant** at the Institute of Behavioural Physiology, Leibniz-Institute for Farm Animal Biology (FBN), Dummerstorf, Germany
03-11/2015 **Research assistant** at the UBC Dairy Research & Education Center, B.C., Canada
06-08/2013 **Student assistant** in the Sea Turtle Protection & Research Project, Fethiye, Turkey

Membership

Societies: International Society for Applied Ethology (ISAE), Universities Federation for Animal Welfare (UFAW), Animal Welfare Research Network (AWRN), AnimalConcepts

Funding and grants

UFAW Small Project and Travel Awards (07-19/20): Workshop "Using quantitative measures of goat behaviour from standardised stress tests to validate qualitative descriptors."

Peer-reviewed manuscripts

Rosenberger, K., Simmler, M., Nawroth, C., Langbein, J., Keil, N. (2020). Goats work for food in a contrafreeloading task. Sci Rep 10, 22336. doi: 10.1038/s41598-020-78931-w

Rasmussen JH, Rosenberger K, Langbein J. (2020). EasieRR: An open-source software for non-invasive heart rate variability assessment. Methods in Ecology and Evolution, 20(00), 1–10. doi.org/10.1111/2041-210X.13393

Rosenberger, K., Costa, J. H. C., Neave, H. W., Von Keyserlingk, M. A. G., & Weary, D. M. (2017). The effect of milk allowance on behavior and weight gains in dairy calves. Journal of dairy science, 100(1), 504-512. doi.org/10.3168/jds.2016-11195

Submitted manuscripts / in preparation

Rosenberger, K., Simmler, M., Langbein, J., Nawroth, C., Keil, N. (2021). Reactivity of domesticated goats towards different stressors following long-term cognitive test exposure. *(in preparation for Plos One)*

Rosenberger, K., Simmler, M., Langbein, J., Keil, N, Nawroth, C. (2021). Performance of goats in a detour and a problem-solving test following long-term cognitive test exposure. *(under review in Royal Society Open Science)*

Published conference abstracts

Rosenberger K., Simmler M., Nawroth C., Langbein J., Keil N. (2020). Untersuchung von Contrafreeloading bei Milch- und Zwergziegen mittels Item-Response-Tree-Analyse. In: Aktuelle Arbeiten zur artgemässen Tierhaltung 2020, KTBL-Schrift 520, 121-131. KTBL, Darmstadt

Rosenberger, K., Nawroth, C., Keil, N., Langbein, J. und Rasmussen, J.H. (2019). EasieRR: Open-access-Software zur Analyse der Herzfrequenzvariabilität. In: Aktuelle Arbeiten zur artgemässen Tierhaltung 2019, KTBL-Schrift 518, 294-297. KTBL, Darmstadt.

Rosenberger K., Nawroth C., Langbein J., Keil N. (2018). Verhaltensunterschiede zwischen Zwerg- und Milchziegen im Novel Arena und Novel Object Test. In: Aktuelle Arbeiten zur artgemässen Tierhaltung 2018, KTBL-Schrift 514, 59-66. KTBL, Darmstadt.

Talks and poster presentations

25.-26.03.2021	Symposium “Farm animal cognition and its implementation into animal husbandry and management” by FBN Dummerstorf and Agroscope Tänikon, virtual Talk Title “Cognitive testing as enrichment for farm animals. Assessment of welfare outcomes and their relevance for housing”
25.02.2021	Fachinformationskreis für ZiegenhalterInnen by Ziegen- und Schafzuchtverband Nö, virtual Talk Title “Ziegenverhalten/-fähigkeiten und Bedeutung für die Haltung“
26.-27.01.2021	Animal Behaviour Twitter conference 2021 by ASAB and ASB, virtual

	Thread Title “Contrafreeloading in goats“
08.01.2021	GCB Symposium, virtual Poster Title “Investigating contrafreeloading (CFL) in dairy and dwarf goats using Item Response Trees (IRTree)”
26.-27.11.2020	52nd International Meeting Applied Ethology, virtual Talk Title “Contrafreloading in dwarf and dairy goats using Item Response Tree Analysis”
30.-31.06.2020	Virtual UFAW conference – “Recent advances in animal welfare science VII”, Poster Title “EasieRR: an open-access software for HRV-analysis and artefact processing in non-restrained animals”
28.-30.11.2019	51st International Meeting Applied Ethology, Freiburg i. Br., Germany Poster Title “EasieRR-Open-access software for the analysis of heart rate variability”
05.-09.08.2019	53rd Congress of the International Society for Applied Ethology (ISAE), Bergen, Norway Poster Title “Dairy and dwarf goats differ in their preference for familiar and unfamiliar humans”
06.-08.02.2019	14th Annual Meeting of the Ethological Society, Hannover, Germany Talk Title “One task to test them all? Performance of goats of different selection lines in an instrumental problem-solving task”
22.-24.11.2018	50th International Meeting on Applied Ethology, Freiburg i. Br., Germany Talk Title “Behavioural differences of dwarf and dairy goats in a Novel Arena und Novel Object Test”
05.-08.09.2018	5th European student Conference on Behaviour & Cognition (ESCBC), University of Vienna, Austria Talk Title “Different selection aims affected goat behaviour in standardized behavioural tests”
11.-12.06.2018	The Joint East and West Central Europe, ISAE Regional Meeting, Warsaw, Poland; “Animal preferences and animal welfare” Talk Title “Behavioural differences between goat breeds selected for intensive and extensive use in standardized behavioural tests”
23.-25.11.2017	49th International Meeting on Applied Ethology Freiburg i. Br., Germany Flash-Presentation: Title “Effect of long-term cognitive training on livestock welfare and husbandry”

Mentoring

2018 Bachelor student, Agrarwissenschaften ETH Zürich, Philip Egli
Aktivitätsverhalten von Milch- und Zwergziegen

2019 Intern DAAD, UBC Vancouver, Canada, Charlotte Hursey
Assistance in data collection and ECG analysis (Chapter III)

2020 Intern, Vetsuisse Zürich, Franziska Kleindienst
ECG analysis for stress reactivity manuscript (Chapter III)

2020 Intern, Vetsuisse Bern, Livia Brühlhart
Assistance in data collection at Zoo Dählhölzli

Declaration of Originality

Last name, first name: Rosenberger, Katrina

Matriculation number: 17127960

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations. All data, tables, figures and text citations which have been reproduced from any other source, including the internet, have been explicitly acknowledged as such. I am aware that in case of non-compliance, the Senate is entitled to withdraw the doctorate degree awarded to me on the basis of the present thesis, in accordance with the "Statut der Universität Bern (Universitätsstatut; UniSt)", Art. 69, of 7 June 2011.

Place, date

Ettenhausen, 20.11.2021

Signature

20.11.2021

X

Signiert von: Rosenberger Katrina VM05JJ