Open Research Online

The Open University's repository of research publications and other research outputs

The Role of Ethological Observation for Measuring Animal Reactions to Biotelemetry Devices

Conference or Workshop Item

How to cite:

Paci, Patrizia; Mancini, Clara and Price, Blaine (2017). The Role of Ethological Observation for Measuring Animal Reactions to Biotelemetry Devices. In: Proc. Fourth International Conference on Animal-Computer Interaction, ACI2017, 21-23 Nov 2017, Milton Keynes, United Kingdom.

For guidance on citations see [FAQs](#).

© 2017 ACM

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:
<http://dx.doi.org/doi:10.1145/3152130.3152144>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

The Role of Ethological Observation for Measuring Animal Reactions to Biotelemetry Devices

Patrizia Paci

The Open University
Milton Keynes, UK
patrizia.paci@open.ac.uk

Clara Mancini

The Open University
Milton Keynes, UK
clara.mancini@open.ac.uk

Blaine A. Price

The Open University
Milton Keynes, UK
b.a.price@open.ac.uk

ABSTRACT

This paper presents a methodological approach used to assess the wearability of biotelemetry devices in animals. A detailed protocol to gather quantitative and qualitative ethological observations was adapted and tested in an experimental study of 13 cat participants wearing two different GPS devices. The aim was twofold: firstly, to ascertain the potential interference generated by the devices on the animal body and behavior by quantifying and characterizing it; secondly, to individuate device features potentially responsible for the influence registered, and establish design requirements. This research contributes towards the development of a framework for evaluating the design of wearer-centered biotelemetry interventions for animals, consistent with values advocated by Animal-Computer Interaction researchers.

Author Keywords

Biotelemetry; wearability; wearer-centered design; animal-computer interaction

ACM Classification Keywords

H.5.2: User-centered Design

INTRODUCTION

The practice of monitoring animals in their naturalistic settings by means of biotelemetry devices such as radio tags, GPS locators, or bio-logging sensors, has been widespread for more than 60 years in areas of biological research such as animal ecology, aiming to acquire information which is usually difficult to obtain with observational techniques alone [33]. For example, elusive endangered wild species have been monitored in order to understand their home ranges and, with respect to their movements, mark protected area boundaries in which they can live undisturbed [5]. More recently, monitoring animals via biotelemetry devices has become a trend within pet caring practices and farm animal husbandry. For example, dogs are increasingly tracked by

their guardians who worry about losing them when the dogs are allowed to walk off lead [18]. Similarly, farm animals are increasingly monitored by farmers who want to locate them when the animals graze outside their paddocks [27].

Biotelemetry has advanced the way in which animals are studied and understood. For example, it has increased the range of physiological and behavioral parameters gathered while decreasing the researchers' intrusion in the animals' habitat [33]. However, there is evidence that carrying biotelemetry tags may influence the very physiology and behavior that is being investigated [31]. Since biotelemetry devices are typically attached to the animal's body, they can constitute a physical intrusion that impinges on the animal's welfare ([13], chapter 6) and consequently affect the reliability of collected data ([22], pp. 15-16). This issue has led a number of biologists to propose design guidelines that could help reduce device-induced effects on individual wearers [3; 21; 34]. However, these guidelines do not systematically reflect a wearer-centered perspective, which has prompted ACI researchers to propose the development of a design framework to systematically inform the design of *wearer-centered* animal biotelemetry [23].

One aim of Animal-Computer Interaction (ACI) is to design animal-centered interactive technologies that can improve animal wellbeing, including minimizing the negative effects that technology used or worn by animals might have on them [15]. However, understanding if an animal-machine interaction is detrimental for an animal can be a difficult process when the effects do not result in obvious signs on the body of the animal (e.g., skin irritation on the site of attachment of a wearable device). The behavioral effects derived by an interaction with technology may be subtle, whereby reactions pro or against a (worn) device may be difficult to identify and use for assessing the animal's experience (e.g., whether the animal is stressed by the presence of the device, or whether they prefer one device or another). Nevertheless, albeit subtle, certain behavioral manifestations may indicate a significant impact on the experience and welfare of the animal, so it is paramount that such indicators are recognized during the interaction design process. There is therefore a need in ACI to assemble a tool box of methodologies for eliciting design requirements and evaluating technology that can enable researchers to detect and assess the subtleties of animals' reactions to technological interactions, especially *implicit* interactions.

Paste the appropriate copyright/license statement here. ACM now supports three different publication options:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

To this end, the paper presents an approach for assessing animals' responses to biotelemetry wearable devices, or other technological interventions, to either establish requirements for, or evaluate the interaction of, such technologies. The proposed approach makes use of protocols derived from the field of ethology and envisages the integration of existing quantitative and qualitative methods for observing and measuring animal behavior, and for analyzing the resulting data. While the use of such integrated approaches has been widely advocated by ACI researchers [10; 15; 35], these have yet to come to fruition within ACI. Here such integration is illustrated through an in-the-wild study of wearable biotelemetry with 13 cats. The aim of the study was to explore the efficacy of the ethological approach to: 1) evaluating the cats' experience when wearing off-the-shelf biotelemetry devices marketed for use on domestic cats; 2) ultimately informing the design of animal-centered technology.

BACKGROUND

Understanding Animals' Interaction with Biotelemetry: Methodological Challenges

ACI is a field of research within computing and interaction design, currently mostly represented by HCI scholars [35] who are well familiar with qualitative methodologies within *User-Centered design* (UCD) and other Interaction Design (ID) approaches. It is therefore no surprise that, in an effort to design animal-centered technologies, ACI researchers have often borrowed and adapted methodologies and techniques from Human-Computer Interaction (HCI) (e.g., [8; 25; 32; 36]). Successful examples of this adaptation have mostly, albeit not only, concerned qualitative methods to design or study the effects of technology for domesticated animals, such as dogs, with whom humans have cohabitated and cooperated for millennia [14]. For example, *iterative physical prototyping* has been used to elicit requirements for technology aimed at medical alert dogs, giving them the opportunity to express preferences for alternative designs [17]; and *multispecies ethnography* has been used to evaluate the effect of tracking technology on the interaction between companion dogs and humans, based on their guardians' accounts and observations of the animals' behavior [18]. Quantitative methods derived from HCI have also been used to evaluate the usability of canine technology designed to support specific tasks. For example, *Fitts' Law* has been used to assess the efficiency of touch screen interfaces for assistance dogs [36]; and other forms of task analysis have been used to evaluate the learnability of tangible interfaces for search and rescue dogs [11].

But what if the animals one designs for do not share the same history of cohabitation and cooperation with humans as dogs do? What if the technology in question is not there to enable animals to perform specific tasks, but is simply physically attached to them? How can such a physical interaction be properly evaluated and requirements for the animal-centered design of such a technology be identified?

When studying the bodily effects that wearable devices may have on animals, being able to observe meaningful signs of reactivity from the wearers is key. However, this is hard to do because the reaction may be cryptically expressed. For example, individual animals may express their irritation to the device they wear through postures or body part movements (e.g., tail, ears, eyes), without necessarily directing their irritation to the device itself (e.g., by trying to remove the tag from their body). On the other hand, more obvious behaviors may appear to be a reaction to the device, when in fact they may be part of the animal's normal behavioral repertoire (e.g., scratching the location where the device is positioned may be a behavior that the animal would perform regardless of the presence of the device). Therefore, there is a risk that any apparent interaction between the animal and the wearable technology may be misleading, if the data are obtained solely through qualitative observations. This is especially so because, when working with animals, researchers do not have the benefit of self-reported validation as they may have when working with humans. Indeed, this is fundamentally a problem of interspecies differences and resulting communication barriers between human researchers and animals, which make interpreting animals' reactions to devices a complex, challenging and uncertain endeavor [10]. This is especially problematic where researchers have insufficient knowledge of species-specific behavior to allow them to appreciate potentially significant nuances and apply due caution when interpreting these [1].

Disciplinary Differences in Observational Methods

Indeed, the abovementioned interspecies communication barriers underpin key differences in how field observation methods are approached in HCI and Ethology. In HCI, *direct observation in the field* is a descriptive qualitative method usually employed to understand the details of what users do in naturalistic settings ([24], pp. 252-254). Observational field studies significantly differ from quantitative methods such as laboratory observations, where participants operate under controlled conditions. Either way, a key aspect is that in HCI observation methods can be complemented by self-reporting methods to somewhat validate observational data. This is of course not possible when working with non-human participants.

Field observation is also the technique used by animal biologists for studying animal behavior in the non-interventionist experimental ethological methodology. However, in ethology, observation is approached somewhat differently from the way in which it tends to be approached in HCI. Ethologists choose and define behavioral categories to observe and quantitatively measure behavioral parameters (e.g., frequency, duration, latency, intensity) in naturalistic settings ([20] pp. 62-66). Observational data are then usually (but not necessarily) treated statistically to verify hypotheses on the meaning of observed behaviors, which thus emerge from quantitative data (e.g., measuring the roaring rate in red deer stags and correlate it with fight occurrences to test whether the roar is an indicator of fitness [4]). This enables

ethologists to interpret the meaning of animal behavior in a relatively objective way, thus reducing the risk of anthropomorphic interpretations ([20], p. 18; [30]). When single occurrences demonstrate the meaning of the observed behavior, ethologists simply describe it (e.g., if a great tit pierces the cap of a milk bottle and then drinks the content, the behavior's meaning is self-explanatory: the bird recognizes milk as food and is motivated to reach for it). However, if ethologists wish to understand how the behavior developed, they still need to observe multiple occurrences across different individuals either in natural settings, and then treat their observations statistically, or in manipulative experiments (e.g., quantification across trained birds demonstrated that individual great tits first learnt to drink from opened bottles, and then learnt to open closed bottles [26]).

Ethological Observation Protocol

As a deductive-inductive approach, *ethological observation* aims to answer questions about adaptation, causation, evolution and development of behavior, enabling researchers to objectively test research hypotheses ([6], chapter 1; [30]), while preserving experimental integrity and safeguarding animal welfare. In order to cause the least disturbance possible, in classical ethology, animals are observed in the wild and their behavior is measured and described as rigorously as possible using non-manipulative techniques. These *natural experiments* essentially differ from *manipulative experiments* where experimental conditions are artificially controlled [30]. The principles for designing ethological observations are the same as manipulative experiments, the difference being in the settings and in the non-manipulative nature of the observation. However, when studying the behavior of domestic animals, applied ethology is the leading branch of science, where manipulative studies tend to prevail. Nevertheless, the principle of minimizing disturbance for animals remains important, along with the abovementioned four fundamental questions ([12], p. 10).

A key challenge in ethological observation is that researchers cannot control the environment and actions of the studied individuals as they could in a laboratory setting (e.g., a wild animal may disappear from the sight of the observer thus interrupting the recording). More importantly, when the behavior measured is not task related (as in laboratory studies where animals are required to perform tasks) but driven by environmental non-controllable external stimuli (e.g., time of day, interactions with conspecifics), interpreting its meaning is even more difficult. Both issues are addressed by ethologists through the development of an observational technique that focuses on a) controlling the observer instead of the animals and their environments; and b) applying the three principles of experimental design ([6], chapter 4; [9], pp. 76-85). Observations are thus planned according to a framework, which comprises four observer's choices and three experimental design principles. The selection of choices and implementation of principles

depends on the specific question that is to be addressed in the study and on the related experimental hypotheses [6] [20].

The four observation choices are [6]:

1. The *granularity of observation* – if animals are to be observed as groups, individuals or body parts (e.g., when predicting that herds protect the off-springs by keeping them surrounded by the adults during migration, the observational units are independent groups, such as familial units; while when studying if off-springs move always when their parents move, the level is the individual young);
2. The *unit of behavior* – namely, the exact behavioral pattern to be observed (e.g., scratching);
3. The *type of sampling* – which depends on what is available to observe, and what question one is trying to answer (e.g., when studying a rare behavior, this should be recorded every time it is observed; when studying recurring behaviors, it might be more convenient to establish fixed periods of time and record them only when they are happening within those periods);
4. The *type of record* - whether the behavior chosen has to be registered continuously (exact start and end each time it is performed) or into short sample intervals (time sampling).

The three principles are [9]:

- a. *Independent replications* – replicating the observations independently is essential if data are to be treated statistically, as this is the way to detect whether an event occurred by chance.
- b. *Not confounding variables* – recognizing when an effect is not caused by the variable the researcher thinks is causing it; this can be avoided in natural settings by randomizing the independent variable (e.g., color of the species coat, time of the day when observation is carried out), or balancing it when the randomization turn out to be skewed (i.e., by chance all the individuals with the same color coat are observed at the same time of the day, for example at night).
- c. *Removing variation* – where variation between sampled individuals cannot be avoided, it can be compensated for by grouping together different individuals (i.e., *blocking*) or comparing the behavior of the same individual between control (i.e., when the independent variable is not present) and experimental conditions (so called *repeated measures design* ([20], p. 29)).

The principal strength of the ethological approach is that it allows researchers to measure animals' behavior, and interpret it, in an objective and reliable way, while minimally disrupting the animals and thus respecting their welfare. It allows researchers to study phenomena that cannot be controlled, or whose control is not desirable (e.g., for ethical reasons, or because an intervention can disrupt the very behavior to be studied). In this respect, the methodology can both replace manipulative experiments and precede them. In

the latter case, observations allow researchers to first explore the behavior of the observed animals, and then plan subsequent experiments, having formulated clearer hypotheses based on early findings, which increases their chance of successfully identifying the meaning of the behavior of interest ([6], p. 6; [20], p. 56).

However, the use of such methodology also presents a number of limitations. Namely, although ethological observations can effectively separate correlation from causation ([6], p. 8), discriminating between the two requires greater caution compared to manipulative experiments, and a robust analysis conducted to high statistical standards, which is not always possible. This might be problematic if an adequate sample cannot be found (that is, enough independent sample units are not available in the context studied). Moreover, individuals under observation may disappear, or leave their habitat, or they may respond to the observer's presence in a way that invalidates the data (e.g., if they express curiosity, escape or otherwise alter their behavior). In these cases, it is important to adjust the research question in such a way that it can be investigated under the existing conditions. This is highly dependent on the choice of the species to be studied, in terms of what it is known about its behavior, and spatial-temporal constraints on observations, which require a compromise between ideal and practicable experimental conditions.

Given its methodological characteristics, the quantitative ethological observation protocol described above is a potentially effective tool for identifying and measuring meaningful reactions that may be caused by the presence of a biotelemetry device on the body, especially when such reactions are subtle or ambiguous. However, ACI researchers are interested in animal's behavior from a design rather than biological point of view. Thus, here the protocol is proposed as a useful way of applying the ethological methodology in ACI, without utilizing the four-question framework of ethology. More importantly, when it comes to evaluating the responses of individual animals to technological interventions, the quantitative methodological limitations described above may underestimate singular but potentially meaningful behaviors. Namely, when the ethological protocol relies on statistical analysis, behavior tends to be considered meaningful if its occurrence is statistically significant. Given that animals under observation may express an individual response that, in terms of type of behavior performed and its occurrence, is inconsistent with the observation choices made by the observer (e.g., in terms of unit of behavior and type of sampling), these animals would be considered outliers and their behavior dismissed as non-significant [7], even though their individual responses might directly derive from the presence of the device. For example, if an animal was to chew off components of the device, but was the only individual of the sample population to do so and only did it once, this important behavior might be omitted from the overall analysis, or treated as an anomaly. However, for the

purposes of designing animal-centered technology, that anomaly may indicate a very noteworthy design flaw. Hence, integrating the quantitative observational protocol with descriptive observations strengthens the methodology, especially when this is applied to the design context.

The study presented in this paper illustrates the *application* of the ethological protocol described above, exploring its usefulness in the context of an ACI project, with a particular focus on investigating cats' responses to commercial tracking collars and eliciting requirements for improving the wearability of such biotelemetry devices. More specifically, ethological observations in the field were conducted to gather data that can be analyzed both quantitatively and qualitatively, thus achieving greater descriptive and inferential power, while reducing interpretational bias, when designing for non-human wearers.

THE STUDY METHODOLOGY

The methodology proposed here is grounded in both manipulative and natural experimental protocols, whereby animal participants were fitted with biotelemetry devices but observed in their living environment to describe and quantify their behavior. Participants were not required to carry out any particular task. Instead, they were observed as they went about their habitual daily activities, while wearing one or another biotelemetry device, and without being restricted in any way, in order to avoid stress induced by habit changes.

Model Species

Domestic cats (*Sp. Felis catus*) were chosen as model species particularly suitable for carrying out a wearability test on biotelemetry devices. Firstly, they are domestic animals accustomed to the presence of and interaction with humans (as opposed to unrestrained wild fauna). This reduces the effects of the disturbance caused by the human observer and the potential bias derived from it ([20], p. 31). Secondly, cats are common pets thus widely available for observation, which made it easier to find participants for the study. Thirdly, due to their domesticity and availability, cats have been extensively studied and a well-documented *ethogram* (i.e., list of behaviors exhibited by an animal) for the species already exists that can be used as a reference [29]. Fourthly, cats are a particularly suitable target group for studying the issue of biotelemetry wearability, since many wearable products available on the market are specifically targeted to them, including the ones tested here.

Experimental Design

The study involved 13 indoor cat participants. It aimed at characterizing and quantifying seemingly unpleasant reactions to (and actions towards) electronic collars fitted around their neck. Indoor cats were chosen to facilitate time standardization of observations, allowing the observer to keep the cat constantly in view, and for ethical reasons (see the dedicated ethics section in this paper). Two different models of GPS devices sold on the market as products specifically designed for pets were tested (Figure 1). For the purposes of the study, the functionality of the devices was

irrelevant and was therefore disregarded. The devices in question were:

- a) PawTrax® Halo (Figure 2a). It is a tracking GPS developed specifically for cats. It weighs 21.7 grams and it is 8.3mm thick. It is flexible at its center thanks to a joint, which allows it to be bent following the cat's neck. The device is designed to slide along a padded 13.5mm-wide collar (5 g) made of synthetic rubber and Velcro® material, which serves to fasten it; it was provided with the device.
- b) Tractive® (Figure 2b). It is a standalone tracking unit that can be fitted to a standard collar. It is generically advertised for be worn by pets, cats included. It is a 41.2 grams device sizing 51x41x15mm. The collar can be fixed to the tag box through a clip, but it is not provided with the device. For this study, a 9.4mm-wide leather collar (8.8 g) was chosen to attach the device onto the cat's neck, which was fastened through a buckle. The choice was made to comply with the manufacturer's suggestion to use standard collars. An advertised 'standard collar' was therefore purchased.

The collars were used as advertised to maintain the integrity of the devices, since the aim was to assess the wearability of the commercial products as sold.

Figure 1. PawTrax (left) and Tractive (right) fitted on the same cat.

Figure 2. PawTrax (a) and Tractive (b) devices compared with a pound coin.

A control condition and two experimental conditions were tested: without wearing anything; wearing the PawTrax collar; wearing the Tractive unit. Cats were monitored for three non-consecutive days, each day with a different condition. All three conditions (control, PawTrax, and Tractive) were assigned randomly to avoid order effects ([20], p. 71). The sampling technique involved focusing on

each individual and recording predetermined behaviors for six 20-minute repetitions each day, for a total recording of 2 hours per day per cat (with each cat wearing each of the two trackers for a total of 6 hours per day). This systematic measurement over fixed lapses was done to increase reliability and as a precaution to the bias of selective recording upon the occurrences of salient events. Tested behaviors are described in detail in a subsequent section. Frequencies (i.e., number of occurrences), and durations (i.e., length of time for which each occurrence lasts expressed in seconds) were measured and compared among the three conditions, within cats.

Data were video-recorded by means of an action camera mounted on the head of the observer, which was discreetly following the cats as they moved around the house (thus minimizing the observer's interference). When the cats were hiding (e.g., under a bed), the camera was positioned in such a way as to allow the recording, so that monitoring could continue. Video aid was utilized in order to record more than one behavior at once in complex sequences where the participants were performing multiple behaviors in a row ([6], p. 85, 142). Video aid was also fundamental for the capture of singular behaviors which could then be analyzed qualitatively. Ten cats were familiar with the observer, who had visited their homes in various previous occasions before the observations were carried out. The other three cats were habituated for half an hour to the presence of the observer, and they were monitored with the presence of their human companions, as this was deemed reassuring for the cats. All the cats were observed (and fitted with the trackers) once they showed signs of sufficient confidence (e.g., rubbing their head on the observer). The experimental design was assessed and approved by the research ethics review body of the Open University and conformed with its ACI research ethics protocol.

Ethical Considerations

As well as receiving the approval of The Open University's Animal Welfare Ethical Review Body, the study presented here was conducted in accordance with recently proposed animal-centered ethical frameworks for ACI research [16]. This argues for the need to preserve the welfare and respect the autonomy of animals involved in research at all times, while recognizing the need to engage with real-world situation in which the autonomy and welfare of the animals may already be compromised ([16], p. 229). Unlike other felines, domestic cats (especially those living indoors) have a special relationship with their carers whereby their daily activities are limited or controlled. For example, carers decide about their roaming freedom, eating habits, reproductive abilities, etc., as well as deciding whether to fit their cats with bell, ID tag, or tracking collars. In fact, the practice of fitting animals with identification tags or biotelemetry devices through collars is accepted by many pet carers, wildlife researchers, and farmers. On the pet market in particular, many commercial products are available for cat carers to purchase and therefore wearing biotelemetry

devices is becoming a part of the ordinary life experience of many domestic cats. These devices are sold as ‘cat-friendly’, presumably considered non-detrimental to the wearers’ welfare. One aim of the study presented here was to test the hypothesis that even what may be considered ‘cat-friendly’ devices may in fact not be so and that better consideration has to be given to their design from the wearer’s perspective ([16], art. 3, par. 1). Therefore those used here were two among such many devices.

Mediated consent ([16], art. 2, par. 1) for cats’ participation was obtained from their carers. Prior to this, the relationship between prospective humans and cat participants was investigated by visiting the households and interviewing the humans to ensure that they had the best interest of their animals at heart. Invariably, the animal companions were regarded as members of the family and fitting GPS collars on their pets was deemed by the carers to be part of the influence they normally exert on the autonomy of their cats in order to appropriately care for them.

Contingent consent ([16], art. 2, par. 2) from cats was also sought in the study. Albeit under the carers’ supervision, cats were fitted with the devices by the researcher, which was only possible when the cat participants allowed the researcher to get close, handle them and instrument them with a foreign body. During the instrumentation process, none of the cats showed signs of fear, pain, or distress. Signs of dissent during the observations were monitored and discussed with the carers when occurring. Any cat that showed fearful or tense behavior once instrumented was withdrawn from the observational sessions and released from the device. These cats were the participants with the highest score in negative experience, which was accounted for in the qualitative data analysis.

Although biotelemetry devices are habitually used on outdoor cats, indoor participants were chosen in this study, in order to ensure the highest standard of safety ([16], art. 4, par. 1). In fact, outdoor-roaming animals fitted with collars may incur in some risks. Although considered negligible by the device manufacturers, there is a risk that a tracking collar might get stuck in vegetation or other medium when the cat roams outdoor. Hence, choosing participants that lived confined in the household property (including enclosed gardens, balconies and courtyards) provided maximum safeguard of the participants’ integrity. This was deemed the best methodological trade-off between the need to understand cats’ requirements and the participants’ safety. At the same time, the household constituted the participants’ habitual environment, which minimized the impact of the study on them ([16], art. 3, par. 3).

Fundamentally, this research is concerned with assessing the experience that animals may have when interacting with wearable technology and with informing animal-centered designs of such wearables. This is in line with the ACI’s aim to improve the quality of life of animals coming into contact with interactive technology. Even though the cats that

participated in this study were not normally subjected to biotelemetry practices, the study was deemed relevant and even beneficial for them ([16], art. 3, par. 1): in the short term, their carers learned about the effect that tracking devices could have on their cats and might thus avoid their use until better designs are available; in the longer term, it is hoped that this research will inform such improved designs which those cats might one day be required to wear.

Measured Behavior

Overall, the study aimed to investigate whether wearing a given tracking device would induce behavioral responses that could be interpreted as the result of an unpleasant experience (e.g., irritation). To this end, the specific methodological question addressed was *how to identify reactions that may indicate annoyance relatable to the device*.

Although the devices in question are marketed as suitable for cats and available to purchase by any cat carer, the general hypothesis was that such a device attached to the cat’s body increases the occurrence and/or duration of certain reactive behaviors. Reaction is defined by The Oxford English Dictionary as the “*neural, neuromuscular, or behavioral response to a stimulus, or more generally: any response to an event; something done, felt, or thought in response to a situation or event*”. However, a working definition for this research is that of perceptual awareness by Sommerville and Broom (1998) [28]. They describe it as “*a perceived stimulus [that] results in an automatic response which may be modifiable, e.g. scratching to relieve irritation, or not modifiable, e.g. blinking when an object passes close to the eye*” ([2], p. 91).

To answer the question, three behaviors that more likely could be related to the presence of a foreign object on the body of the animal were initially chosen from the cat ethogram for measuring. For example, cats may insist in grooming themselves on a particular spot in response to soiled fur or a tick, or they may scratch because of flea infestation, or shake their body to get water off the fur. Thus, the three reactions hypothesized were: 1) scratching the site of attachment to attain relief, 2) shaking the head and/or body to release an object or substance from the body, or to release a cumulative neuronal stimulus ([19], p. 6), 3) grooming a spot to discard a foreign body. During data collection, responses that differed from scratching, head/body shaking and grooming but that the observer deemed related to the presence of the device were also noted for qualitative analysis. These responses included obvious direct interactions (e.g., actual biting of the tag case) and attempts to interact with the device (e.g., movements of the head to reach the case with open mouth even if biting was not performed). Behaviors directed towards the devices (excluding scratching, head/body shaking and grooming) were grouped together into a composite category described here as *active interaction*. Behaviors noted by the observer were licking, biting, and cuffing the case or attachment of the

device, or rolling the head and raising a paw in an attempt of licking, biting or cuffing the device. The category of active interaction accounted for behavioral individuality, whereby each cat can perform a direct action in a different way [7].

All the base behaviors listed above are defined in Stanton's ethogram [29]:

- Scratching: cats scratch their bodies using the claws of their hind feet.
- Head/body shake: cats rotate their head/abdomen from side to side.
- Groom: cats clean themselves by licking, scratching, biting or chewing the fur on their body. May also include the licking of a front paw and wiping it over one's head.
- Lick: cat's tongue protrudes from mouth and strokes (*the device*).
- Bite: cats snap teeth at and are successful in biting (*the device*).
- Cuff: cats strike at (*the device*) with forepaw and contact is made. Claws are usually extended.
- Head-rolling: cats toss their heads in a circular motion (*trying to reach the device*).
- Raising paw: cats lift their forepaw as if to cuff, paw or strike at (*the device*) but do not follow through with the action.

Data Analysis

Frequencies - and durations when relevant (see 'definition of bout' section) - of scratching, head/body shaking, and grooming were quantified. In parallel, when a behavior was directed or seemingly aimed at the device, it was added to the active interaction category, and treated as a quantifiable category. The aim of this quantification was to measure systematically the cats' reactions in order to individuate a behavioral change through categories of behaviors that may function as indicators of unpleasantness.

Subsequently, all the behaviors listed above were described in detail to analyze the way in which they were performed. In this respect, the aim was to identify potential design flaws in the devices; to this end, relevant sequences of behavior that the observer deemed to be signs of annoyance and seemingly directed against the device were analyzed qualitatively to identify features of the device that might have produced such a strong reaction.

In this way, the participants most reactive to the devices were individuated. Focusing on reactive participants is key for improving the design of biotelemetry devices, the rationale being that, if a device is acceptable to a sensitive individual, it is likely to be acceptable also to the least reactive individuals (under comparable conditions).

Definition of bouts

The four behaviors of interest (i.e., grooming, scratching, head/body shaking, and active interaction) were performed

by cats as single occurrences, or in sequences of multiple behaviors (e.g., grooming, scratching, shaking the head and grooming again). Such sequences can be interrupted for few seconds by distractive events (e.g., while a cat is grooming, he/she might be distracted by a sudden noise, thus stopping and restarting after a while). To systematically record the frequency, sequences were separated into bouts.

A bout is a repeated sequence of the same act ([6], p.85; [20], p.67). In this study, for each category of behavior, bouts were defined as follow:

- A scratching bout was when participants scratched *their neck or throat* using the claws of their hind feet continuously without stopping. The bout ended when cat put the scratching feet on the ground. If the cat restarted after this, it was considered as a new occurrence.
- Head and body shaking were deemed related behaviors that measure the same reaction. Thus, they were considered a single category ([20], p. 58). Duration was not registered for these. Instead, they were treated as events because clocking the rapid rotation of the head and/or body was deemed a peripheral measure for the purposes of this study, and therefore approximated as points in time ([20], p.66). Head and body shaking were not counted when clearly performed because other stimuli occurred (e.g., when some long-haired participants shook their heads during grooming to tear off hair knots more efficiently).
- A simple grooming bout was when the cat groomed their fur continuously without interruptions (in other words, when the typical movement of the head as the cat protrudes the tongue for licking was not interrupted). However, grooming sequences can be more complex. In such cases, the cat can stop grooming to change position, or to just pause, or to pay attention to a distraction, to then come back to grooming after a while. In this analysis, any change in behavior was considered as an interruption of the bout, except for changes in position made for grooming purposes, or when staring at something that caught the cat's attention (e.g., a sudden noise). A pause of less than 30 seconds was not considered a change in behavior, but an interruption of 30 seconds or more defined the end of the bout. This lapse was deemed by the observer long enough to mark a change in behavior rather than a momentary interruption due to environmental distractions. The duration of grooming was recorded net from the pauses or sudden distractions, that is the chronometer used for recording durations was stopped during momentary interruptions, but restarted immediately as the cat resumed grooming (thus considering the action as one occurrence). For example, if during grooming the cat stopped to stare at something but resumed grooming within 30 seconds, the behavior was considered as a single occurrence and clocked for the effective time during which the stroking movement of the tongue was performed.

- An active interaction bout was when participants licked, cuffed, bit, rolled their head and/or raised their paw at the device continuously without interruption. The bout ended when the cat stopped and changed activity. If the cat then resumed, this was considered as a new occurrence (unlike with grooming, which is normally a protracted activity prone to momentary interruptions during its course).

Quantification and qualification criteria

1. Data were firstly treated quantitatively and statistically to find a difference among the three conditions and identify categories of behavior as reliable indicators of unpleasantness to evaluate the impact of the device.
2. Measuring and quantifying the behaviors served also as basis for identifying more reactive cats to the presence of the device. The behaviors of these participants were then analyzed more in detail.
3. Bouts that the observer classified as overreactions were extracted from videos and further analyzed. Episodes of the measured behaviors with the PawTrax and Tractive devices underwent qualitative analysis to identify design features of the device that might have generated the overreaction.

Statistical analysis

Parametric statistical test such as repeated-measured ANOVA were performed for duration measurements when the assumptions of normality and sphericity of data were met.

Since frequencies are counts of events (i.e., behaviors) rather than numbers (e.g., duration of scratching in seconds), they do not follow the normal distribution required to carry out parametric statistical tests, but follow a Poisson distribution instead ([9], p. 258). Hence, data for frequencies were statistically treated with the nonparametric test of Friedman's two-way ANOVA. Statistical analysis was carried out by means of the Software IBM SPSS Statistics, Version 21.0.

An Example of Result from the Empirical Study

Results were extrapolated from a total of 2 effective hours of observation each day, for three total days (i.e., 6 hours per cat in total). This paper proposes the use of a methodology borrowed from the biological field to widen the tool box available to ACI researchers, as deemed necessary at the first international workshop on research methods in ACI, held during the ACI2016 conference [35]. Hence, not all the findings are reported here. Rather, examples are reported to illustrate the kind of information that can be obtained through the application of this methodology, and a way of integrating quantitative and qualitative analysis to maximize the insights obtained from the data from an interaction design perspective. In this respect, outcomes for head/body shaking and scratching frequencies provide good insights.

Means for head/body shaking frequencies in the three experimental conditions were 3.54 (s.d. = 2.62) during the control, 9.88 (s.d. = 7.37) with the PawTrax, and 8.50 (s.d. =

6.60) with the Tractive. Medians were 3.50 for control, 11.00 with the PawTrax, and 8.00 with the Tractive.

Friedman's test showed that there was a significant difference in head and body shaking frequency across the three different conditions, $X^2(2) = 6.533$, $p = 0.038$ ($p < 0.05$).

Means for scratching frequencies in the three experimental conditions were 0.54 (s.d. = 0.51) during the control, 4.92 (s.d. = 6.44) with the PawTrax, and 4.00 (s.d. = 5.91) with the Tractive. Medians were 1.00 for control, 1.00 with the PawTrax, and 2.00 with the Tractive.

Friedman's test results showed that there was no statistically significant difference in scratching frequency across the three different conditions, $X^2(2) = 2.837$, $p = 0.242$ ($p > 0.05$).

Frequencies for scratching for each participant are additionally reported in Table 1 to look at the extent of the increment registered in some cats.

The data show that not all the cats reacted to either or both devices. The comparison among the control and each device shows that four cats showed an increment in scratching frequency of more than 3 occurrences with both devices (C4, C7, C9, and C10). Other two participants increased scratching with only one of the two devices: namely the PawTrax, (but not the Tractive) for C3, and the Tractive (but not the PawTrax) for C6. Seven cats did not show a difference bigger than 3 occurrences between conditions.

cat ID	P		C		T
C1	0	≈	1	≈	2
C2	0	≈	1	=	1
C3	12	>	1	≈	0
C4	9	>	1	<	21
C5	0	=	0	=	0
C6	0	=	0	<	4
C7	8	>	0	<	6
C8	2	≈	1	≈	0
C9	14	>	0	<	10
C10	18	>	1	<	5
C11	0	≈	1	=	1
C12	0	=	0	≈	2
C13	1	≈	0	=	0

Table 1. Comparison of scratching frequencies among PawTrax (P), Control (C) and Tractive (T) conditions.

However, for those cats that reacted in a meaningful way, qualitative descriptions of each scratching bout show that the behavior is specifically directed at the device(s).

For example, C7 showed an increment in scratching with both PawTrax and Tractive devices in a way that strongly suggests annoyance. One of the qualitative descriptions extrapolated from videos showed that, with the PawTrax

device, the participant scratched on the same spot of the neck 6 times during a sequence of 1 minute and 20 seconds. From the video's audio, it was evident that the cat was scratching the collar instead of the skin underneath, as evident from the sound generated by the claws hitting the collar. This suggests that, due to the substantial width of the collar, the cat could not relieve the irritation that seemingly caused him to want to scratch his neck, therefore he scratched repeatedly seemingly in an attempt to find relief. On the contrary, with the Tractive device, scratching was each time performed as a single occurrence suggesting that the cat could at least momentarily alleviate the irritating stimulus caused by the device. This suggests that, in this case, the narrower collar allowed the claws to reach the irritated spot. This kind of qualitative results suggests specific implications for design.

Another sequence in C7 shows the cat rubbing his neck on the tile floor of the garden while wearing the Tractive device. This produced a rubbing noise of the device on the floor. Suddenly, C7 tossed his body lying on his belly and rigidly tilted his head toward his throat freezing for an instant. Then followed jerking and agitated movements that culminated with the cat rearing on hind limbs while cuffing and biting the case as his forepaws clutched the case. This sequence shows an unambiguous reaction of annoyance. This cannot be accounted for statistically as it was exhibited by only two of the cats from the study sample. However, it points to at least two potential design problems:

1. The overall reaction was more likely triggered by the noise produced by the hard plastic material of the case when it hit the ground. In fact, the cat reacted immediately after the device rubbed against the floor. This suggests that, for future designs, the hardness of the case material should be carefully considered to avoid an influence of the case when cats rub their head or body against objects, surfaces and other individuals, as they commonly do.
2. The reaction of rearing on hind limbs and cuffing suggests that the protrusion of the device is too conspicuous. It is possible that the cat was able to see the device. Definitely, he was able to engage in a fight with it. The cuffing reaction was never performed with the PawTrax device, which is less protruding than the Tractive (8.3 versus 15 millimeters). However, with the PawTrax, the participant attempted to reach the device by rolling his head and licking the collar on several occasions. This suggests that the cat perceived the presence of the PawTrax device on the body but possibly could not see it or feel any protrusion when leaning his head forward (as was seemingly happening with the Tractive). While it can be hypothesized that both behaviors had the same purpose - removing the irritating foreign body - each behavior seemingly signals an unwanted effect resulting from a specific design feature of the device.

DISCUSSION

The study findings show that there is not a statistically significant difference in the scratching behavior of the cats

between the control and the two devices. Statistically, this would suggest that scratching is not a good indicator for annoyance, or that cats scratch their neck and throat when they wear a device just as they normally do when they do not wear it. However, looking at the data from each cat, in some of them, scratching increased in a way that is qualitatively meaningful. Namely, several instances of scratching concentrated in a same brief sequence, and instances of active interaction with the device cannot be discounted as irrelevant. Overall, the data suggest that the presence of the device produces an impact on the wearer, in spite of the absence of statistical significance. Hence, when evaluating the wearability of a device (with cats), scratching could be still used as an indicator of a negative impact (given that scratching in cats is a behavior whose function is to relieve irritation or eliminate detrimental external agents, such as parasites). On the other hand, statistical analysis attributes a significant difference to head and body shaking between experimental conditions. For this category of behavior this is a valuable result. In fact, head/body shaking is a fast and automatic reaction that does not lend itself to qualitative description. Thus, behavior quantification is a better way of investigating it as a potential indicator. Hence, in this case a quantitative and statistical analysis provides a more compelling case for the meaning of shaking as an indicator of annoyance.

Furthermore, the findings show that, with respect to scratching behavior, the PawTrax device may generate a bigger impact compared to the Tractive device; this raises the question as to which features of the device are responsible for generating a stronger reaction. To answer this question, the cases of individuals that showed a stronger reaction were examined qualitatively. For C7 - the example presented here - bouts were checked qualitatively to assess the behavior of the animal. The findings suggest that the width, thickness and material of the PawTrax collar may have prevented the cat's scratching behavior from effectively relieving an itch, thus causing him to repeat the behavior. On the other hand, with the Tractive device, the findings suggest that the bigger protrusion and case material of the unit may have stimulated the cuffing behavior. This suggests that these features of the devices should be carefully considered for improvement during the design process. To verify this, in future tests, collars might be swapped between the two devices, using the PawTrax collar with the Tractive unit and vice versa, on the same more reactive cats, to examine any behavioral variations.

Overall, the study findings illustrate the complementarity functions of the quantitative and qualitative treatment of the experimental data. On the one hand, the quantitative analysis of data from all the cats can help identify behaviors that may indicate a reaction associated to the presence of the device (particularly if those behaviors are already described in the animals' ethograms). These responses could then be taken as a reference when evaluating different designs and corresponding prototypes of a wearable device, which could

help expedite and target the evaluation process. On the other hand, the qualitative analysis of data related to the more reactive individuals (particularly occurrences of direct interaction with the device) can yield insights into which device features may provoke or increment the intensity of a reaction (e.g., scratching repeatedly because the irritation is not alleviated). This could in turn help establish specific requirements for the design or redesign of the wearable device in question and thus improve the experience of the wearer.

The experience of the individual is central to UCD and the same ethos seems to inform much of ACI work, at least in terms of what the field seems to aspire to. However, many of the research methods developed within UCD rely on the use of shared languages, or need to be used in combination with methods that in turn rely on the use of shared language. This makes their application problematic for working with non-human research participants. As mentioned above, ACI researchers have, in various occasions, adapted research methods developed within HCI, but these have mostly required the mediation of human participants, exposing ACI research to the risk of anthropomorphism. If ACI is to develop into healthy discipline, there is a need to find ways of mitigating the risk of human interpretational bias while still maintaining a strong focus on the animal user or, as in the case of the work presented here, wearer.

As a methodological approach, ethological observation was developed by ethologists to answer questions about adaptation, causation or development of behavior [26]. Ethologists wanted to explain complex evolutionary, causative and adaptive questions, for example: why a cat scratches when he has fleas (in other words, which is the causative mechanism underpinning this behavior); why scratching evolved as the response to having relief; whether scratching is innate or learnt and thus improves with practice when off-springs become adult. In other words, ethologists are interested in the biological meaning of a behavior, which is why they tend to study large numbers of individuals, rather than single individuals, focusing on hypotheses that can be tested statistically. However, behavioral singularities may play a key role in formulating appropriate hypotheses that can be tested subsequently and is therefore integral part of the ethological approach. When it comes to designing interactive systems for animals, the qualitative description of singular behaviors is essential, because even individual responses can provide critical insights. Thus, the observational protocol utilized in the study presented here allowed for the collection of data that could be analyzed both quantitatively and qualitatively, whereby the two forms of analysis yield complementary findings.

As mentioned earlier, some ACI researchers have advocated the application of research methods grounded in biological science to improve the rigor of research practice and the reliability of research findings [9]. The study presented here demonstrates the application of one such method to evaluate

the effects of off-the-shelf biotelemetry devices on cats, and the insights that can be derived. Arguably, this is highly relevant when the animals in question are not intended users of the technology, but still interact with it in more implicit ways. In this case, a quantitative assessment can enable researchers to measure and interpret responses whose meaning is potentially ambiguous. More specifically, it can help discriminate whether species-specific behaviors that are part of the animals' normal repertoire (e.g., scratching) occur in response to the presence of the device they wear (e.g., irritation caused by it) or to other external stimuli (e.g., a parasite). On the other hand, it can help examine singular behaviors directed to the device itself (e.g., grasping or chewing the device), which more explicitly derive from the presence of the device (in that they could not occur if the device was not there). However, it is proposed here that the application and possibly adaptation of the ethological observation method in ACI research more broadly could improve the objectivity of researchers' observations and reduce their interpretational bias, while allowing them to account for singularities that may be highly significant when designing animal-centered technology.

CONCLUSION

Wearing devices may have effects on the wellbeing of animals. Understanding animals' needs and experiences is fundamental for delivering animal-centered products, but is a challenging endeavor. However, biologists have found ways of interpreting animal behavior by comparing it between or within individuals. This paper adapts the ethological experimental approach to ACI design research and presents a new ACI methodology for assessing wearability issues in biotelemetry devices worn by animals. The methodology was used in an initial field study involving 13 cat participants to identify design issues in two commercial off-the-shelf collar-based GPS trackers. The study shows how this methodology can be used to investigate experiences in cats, and how to apply it to improve wearability. This methodology could allow the understanding of both experiences and design requirements to be used in the iterative process of designing animal wearables. To this end, prototyping a novel device following the results of this study will be the next step in the overall research of developing a wearer-centered framework for animal-centered devices.

ACKNOWLEDGEMENTS

The authors are grateful to all cat carers for giving access to their properties and for giving their time to the observer. Special thanks are due to the cat participants for wearing the devices and for tolerating the observer's presence. Gratitude is owed to Dr Jekaterina Rogaten and Dr Germain Van Bever for statistical advice. This research did not receive any financial support from any pet wearable companies.

REFERENCES

- [1] BASKIN, S., ANAVI-GOFFER, S., and ZAMANSKY, A., 2015. Serious Games: Is Your User Playing or Hunting? In *International*

- Conference on Entertainment Computing Springer, 475-481.
- [2] BROOM, D.M., 2003. *The evolution of morality and religion*. Cambridge University Press.
 - [3] CASPER, R.M., 2009. Guidelines for the instrumentation of wild birds and mammals. *Animal Behaviour* 78, 6, 1477-1483.
 - [4] CLUTTON-BROCK, T.H., ALBON, S., GIBSON, R., and GUINNESS, F.E., 1979. The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Animal Behaviour* 27, 211-225.
 - [5] COOKE, S.J., 2008. Biotelemetry and biologging in endangered species research and animal conservation: relevance to regional, national, and IUCN Red List threat assessments. *Endangered species research* 4, 1-2, 165-185.
 - [6] DAWKINS, M.S., 2007. *Observing animal behaviour: design and analysis of quantitative data*. OUP Oxford.
 - [7] FEATHER, J., MENDEL, M., and BATESON, P., 1986. A method for rating the individual distinctiveness of domestic cats. *Animal Behaviour* 34, 4, 1016-1025.
 - [8] FRENCH, F., MANCINI, C., and SHARP, H., 2016. Exploring methods for interaction design with animals: a case-study with Valli. In *Proceedings of the Third International Conference on Animal-Computer Interaction* ACM, 3.
 - [9] GRAFEN, A., HAILS, R., HAILS, R., and HAILS, R., 2002. *Modern statistics for the life sciences*. Oxford University Press Oxford.
 - [10] HALL, C. and ROSHIER, A., 2016. Getting the measure of behavior … is seeing believing? *interactions* 23, 4, 42-46. DOI=<http://dx.doi.org/10.1145/2944164>.
 - [11] JACKSON, M.M., VALENTIN, G., FREIL, L., BURKEEN, L., ZEAGLER, C., GILLILAND, S., CURRIER, B., and STARNER, T., 2015. FIDO—Facilitating interactions for dogs with occupations: wearable communication interfaces for working dogs. *Personal and Ubiquitous Computing* 19, 1, 155-173.
 - [12] JENSEN, P., 2002. *The ethology of domestic animals: an introductory text*. CABI.
 - [13] KENWARD, R.E., 2000. *A manual for wildlife radio tagging*. Academic Press, London.
 - [14] LARSON, G. and BRADLEY, D.G., 2014. How much is that in dog years? The advent of canine population genomics. *PLoS genetics* 10, 1, e1004093.
 - [15] MANCINI, C., 2011. Animal-computer interaction: a manifesto. *interactions* 18, 4, 69-73.
 - [16] MANCINI, C., 2017. Towards an animal-centred ethics for Animal-Computer Interaction. *International Journal of Human-Computer Studies* 98, 221-233.
 - [17] MANCINI, C., HARRIS, R., AENGENHEISTER, B., and GUEST, C., 2015. Re-centering multispecies practices: a canine interface for cancer detection dogs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* ACM, 2673-2682.
 - [18] MANCINI, C., VAN DER LINDEN, J., BRYAN, J., and STUART, A., 2012. Exploring interspecies sensemaking: dog tracking semiotics and multispecies ethnography. In *Proceedings of the 2012 ACM conference on ubiquitous computing* ACM, 143-152.
 - [19] MANNING, A. and DAWKINS, M.S., 1992. *An introduction to animal behaviour*. Cambridge University Press.
 - [20] MARTIN, P., BATESON, P.P.G., and BATESON, P., 1993. *Measuring behaviour: an introductory guide*. Cambridge University Press.
 - [21] MORTON, D.B., HAWKINS, P., BEVAN, R., HEATH, K., KIRKWOOD, J., PEARCE, P., SCOTT, L., WHELAN, G., and WEBB, A., 2003. Refinements in telemetry procedures. *Seventh report of the BVAAWF/FRAME/RSPCA/UFAW Joint working group on refinement, part A. Lab Anim* 37, 4, 261-299.
 - [22] MURRAY, D. and FULLER, M., 2000. A critical review of the effects of marking on the biology of vertebrates. In *Research techniques in animal ecology: controversies and consequences*, B.L.A.F. TK Ed. Columbia University Press, New York, 15-64.
 - [23] PACI, P., MANCINI, C., and PRICE, B.A., 2016. Towards a wearer-centred framework for animal biotelemetry. In *Measuring Behaviour 2016*, A.J. SPINK Ed., Dublin, 465-469.
 - [24] PREECE, J., SHARP, H., and ROGERS, Y., 2015. *Interaction design: beyond Human-Computer Interaction (fourth edition)*. John Wiley & Sons.
 - [25] ROBINSON, C., MANCINI, C., VAN DER LINDEN, J., GUEST, C., and HARRIS, R., 2014a. Canine-centered interface design: supporting the work of diabetes alert dogs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 3757-3766.
 - [26] SHERRY, D.F. and GALEF, B.G., 1984. Cultural transmission without imitation: milk bottle opening by birds. *Animal Behaviour* 32, 3, 937-938.
 - [27] SIKKA, P., CORKE, P., VALENCIA, P., CROSSMAN, C., SWAIN, D., and BISHOP-HURLEY, G., 2006. Wireless adhoc sensor and actuator networks on the farm. In *Proceedings of the 5th international conference on Information processing in sensor networks* ACM, 492-499.
 - [28] SOMMERVILLE, B. and BROOM, D., 1998. Olfactory awareness. *Applied Animal Behaviour Science* 57, 3, 269-286.

- [29] STANTON, L.A., SULLIVAN, M.S., and FAZIO, J.M., 2015. A standardized ethogram for the felidae: A tool for behavioral researchers. *Applied Animal Behaviour Science* 173, 3-16.
- [30] TINBERGEN, N., 1963. On aims and methods of ethology. *Ethology* 20, 4, 410-433.
- [31] WALKER, K.A., TRITES, A.W., HAULENA, M., and WEARY, D.M., 2012. A review of the effects of different marking and tagging techniques on marine mammals. *Wildlife Research* 39, 1, 15-30.
- [32] WESTERLAKEN, M. and GUALENI, S., 2014. Felino: The philosophical practice of making an interspecies videogame. In *The Philosophy of Computer Games Conference*, 1-12.
- [33] WILMERS, C.C., NICKEL, B., BRYCE, C.M., SMITH, J.A., WHEAT, R.E., and YOVOVICH, V., 2015. The golden age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology. *Ecology* 96, 7, 1741-1753.
- [34] WILSON, R.P. and MCMAHON, C.R., 2006. Measuring devices on wild animals: what constitutes acceptable practice? *Frontiers in Ecology and the Environment* 4, 3, 147-154.
- [35] ZAMANSKY, A., ROSHIER, A., MANCINI, C., COLLINS, E.C., HALL, C., GRILLAERT, K., MORRISON, A., NORTH, S., and WIRMAN, H., 2017. A Report on the First International Workshop on Research Methods in Animal-Computer Interaction.
- [36] ZEAGLER, C., ZUERNDORFER, J., LAU, A., FREIL, L., GILLILAND, S., STARNER, T., and JACKSON, M.M., 2016. Canine computer interaction: towards designing a touchscreen interface for working dogs. In *Proceedings of the Third International Conference on Animal-Computer Interaction* ACM, 2.