Does the treadmill support valid energetics estimates of field locomotion?

O. R. Bidder[1,2*], C. Goulding[3], I. A. Toledo[3], T. A. van Walsum[3], U. Siebert[1], L. G. Halsey[3]

[1]Institute for Terrestrial and Aquatic Wildlife Research (ITAW), University of Veterinary Medicine Hannover, Werftstr. 6, Büsum, Germany 25761
[2]Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, CA 94720
[3]Derpartment of Life Sciences, University of Roehampton, London SW15 4JD
*corresponding author: o.r.bidder@gmail.com

Keywords: metabolic rate, energy expenditure, accelerometry, heart rate, walking, running, treadmill, calibration, meta-analysis

Short title: treadmill proxy for field locomotion

Abstract

Quantifying animal energy expenditure during locomotion in the field is generally based either on treadmill measurements or on estimates derived from a measured proxy. Two common proxies are heart rate (*f*H) and dynamic body acceleration (accelerometry). Both *f*H and accelerometry have been calibrated extensively under laboratory conditions, which typically involves prompting the animal to locomote on a treadmill at different speeds whilst simultaneously recording its rate of oxygen uptake ($\dot{V}O_2$) and the proxy. Field estimates of $\dot{V}O_2$ during locomotion obtained directly from treadmill running or from treadmill-calibrated proxies make assumptions about similarities between running in the field and in the laboratory. The present study investigated these assumptions, focussing on humans as a tractable species. First we investigated experimentally if and how the rate of energy expenditure during treadmill locomotion differs to that during field locomotion at the same speeds, with participants walking and running on a treadmill, on tarmac and on grass, while wearing a mobile respirometry system. $\dot{V}O_2$ was substantially higher during locomotion in both of the field conditions compared to on a level treadmill: 9.1% on tarmac and 17.7% on grass. Second, we included these data in a meta-analysis of previous, related studies. The results were influenced by the studies excluded due to particulars of the experiment design, suggesting that participant age, the surface type and the degree of turning during field locomotion may influence by how much treadmill and field locomotion $\dot{V}O_2$ differ. Third, based on our experiments described earlier, we investigated the accurancy of treadmill-calibrated accelerometry and f_{H} for estimating $\dot{V}O_2$ in the field. The mean algebraic estimate errors varied between 10 and 35%, with the f_{H} associated errors being larger than those derived from accelerometry. The mean algebraic errors were all underestimates of field $\dot{V}O_2$, by around 10% for fH and varying between 0 and 15% for accelerometry. Researchers should question and consider how accurately a treadmill-derived proxy calibration of $\dot{V}O_2$ will estimate $\dot{V}O_2$ during terrestrial locomotion in free-living animals.

Introduction

Many animals spend a lot of energy moving (Rezende et al., 2009, Williams et al., 2014, Scantlebury et al., 2014, Halsey et al., 2015b, Gefen, 2011). Locomotion costs are an important component of their finite energy budgets, from which they must also pay for all other behaviours. A good insight into the energy that animals invest in locomotion is therefore essential to understanding their ecology (Halsey, 2016). However, quantifying the energy expenditure of wild animals is difficult. Researchers cannot usually measure energy expenditure in the field directly but must instead estimate it. One option is to estimate a free-ranging animal's energy expenditure during locomotion from time-energy budgets combined with measurements of rate of energy expenditure during treadmill running (e.g. Hoyt and Kenagy, 1988, Kenagy and Hoyt, 1989). This of course assumes that rate of energy expenditure on the treadmill is equivalent to that in the field. An alternative is to estimate energy expenditure from measurements of a correlated proxy that can be recorded in free-ranging animals (Halsey et al., 2011b, Green, 2011). Two such proxies of energy expenditure in free living animals are heart rate (f_{H}) and body acceleration (Green et al., 2009, Wilson et al., 2006). In both cases, their measurement requires instrumentation of the subject animal with an electronic device that measures and then records or transmits the data (Green et al., 2009).

Both the accelerometry and heart rate techniques have been investigated extensively under laboratory conditions, which are required for calibration of the proxy (e.g. Butler et al., 2004, Wilson et al., 2006, Green et al., 2009, Halsey et al., 2011b). For terrestrial species, calibration typically involves encouraging the animal to walk or run at various speeds on a treadmill within a respirometry chamber (e.g. Halsey et al., 2009b). This is intended to induce systematic variations both in rate of energy expenditure (measured as rate of oxygen consumption; $\dot{V}O_2$) and in the proxy, to enable a calibration curve to be fitted (Halsey et al., 2008, Nolet et al., 1992, Bevan et al., 1995, Hawkins et al., 2000, Froget et al., 2001, Green et al., 2001, Brage et al., 2006). An implicit assumption of this process is that the calibration relationship between rate of energy expenditure and the proxy holds in the field. However, for a variety of reasons pedestrian locomotion on a treadmill may not be a suitable surrogate for the same activity outside (Van Ingen Schenau, 1980), and thus the suitability of the treadmill as a basis for estimating the locomotion costs of free-ranging aninmals is questionable. Yet to date, differences in energy expenditure between locomotion on a treadmill and other terrains has not been investigated in animals.

In the human literature, however, several published studies provide data on rate of energy expenditure during walking or running both on a treadmill and on surfaces outside of the lab. One outcome of this work is that, within the sports science discipline, it is common for researchers and coaches to conduct treadmill training at a 1% gradient because this is reported to most accurately emulate the energetic cost of field running in ideal conditions (Jones and Doust, 1996). However, much of the human literature is conflicting, with some concluding that locomotion on a treadmill is more energetically demanding (Parvataneni et al., 2009, Berryman et al., 2012, Barnett et al., 2015) and others the contrary (Daniels et al., 1953, Wyndham et al., 1971, Pearce et al., 1983). Other studies found no evidence for a difference in energy expenditure between treadmill and field locomotion (Ralston, 1960, Jankowski et al., 1972, Murray et al., 1985). The experimental procedures in each of these aforementioned studies somewhat differ. For example, the surface type varies substantially outdoors, which has been shown elsewhere to influence the cost of transport and the proxies of energy expenditure in both humans and other animals (Bidder et al., 2012, Pandolf et al., 1976, Knapik et al., 2004, Crête and Larivière, 2003, Fancy and White, 1987).

The issue of whether calibrated proxies of energy expenditure will accurately estimate energy costs in the field could be more nuanced than simply whether walking/running on the treadmill accurately simulates the energy costs of moving at the same speeds in the field. Not only might the energetic costs to move at a given speed differ between the treadmill and the field but the relationship between rate of energy expenditure and the proxy might differ between the two conditions (Figure 1). If this is the case, then the energy costs of locomotion at a given speed could be the same between the treadmill and the field yet the treadmill-calibrated proxy return an inaccurate estimate of $\dot{V}O_2$ in the field. Conversely, if the relationships between $\dot{V}O_2$ and the proxy are the same on the treadmill and in the field then even if the energetic costs of locomotion at a given speed are different in these two conditions, there will be a concomitant difference in the magnitude of the measured proxy and thus the proxy will produce an accurate estimate of the cost of field locomotion.

In our study, first we investigated if and how the rate of energy expenditure to walk/run on a treadmill differs to that to walk/run on tarmac or on cut grass, employing humans because they are the most tractable species for such a study. We recorded $\dot{V}O_2$ as a measure of rate of energy expenditure during aerobic activity using a mobile respiratory gas analyser. Second, we then included these data in a meta-analysis of previous studies to understand the magnitude of the difference in rate of energy expenditure between treadmill and field locomotion in general. Third, we investigated the accurancy of treadmill-calibrated proxies (body acceleration and f_H) for application to the field. To achieve this we compared the measured values of $\dot{V}O_2$ for pedestrian locomotion in the field with the estimates of $\dot{V}O_2$ obtained from the calibrated proxies.

Methods

Participants and experiments

The experimental protocols of this study were approved by the Ethics Committee at the University of Roehampton. Between December 2015 and March 2016, seven men and thirteen women of mean age 24.8 ± 1 standard deviation (1SD) 8.1 years, encompassing a range of statures and medium to high fitness levels, participated. Most participants undertook all experimental conditions on the same day. Before the start of the experiments, the participants' weights (mean ± 1SD: 66.7 ± 9.3 kg) and heights (170.1 ± 10.5 cm) were recorded. Participants did not have any cardiac or metabolic disorders and were not currently taking any medication for chronic ailments. Informed consent was obtained from all participants and the physical activity readiness questionnaire (PAR-Q) completed before the experiments began. The PAR-Q was used to assess whether the individual should seek medical advice prior to participation in the study. Six of the participants self-identified as regular users of treadmills, six stated they used treadmills irregularly, and eight stated they had no previous experience with treadmills. All participants were given a period at the beginning of the experiment to familiarise themselves with walking and running on the treadmill before measurements were taken.

To compare the cost of locomotion on a treadmill and in the field, three experimental surface conditions were used: a speed-calibrated laboratory treadmill (Woodway Ergo ELG 70), flat hard tarmac and flat soft cut grass. The experiments consisted of walking at two speeds and jogging at one speed, on each surface. During the experiments, $\dot{V}O_2$ (ml min^{-1}) and f_H (beats min^{-1}) were recorded along with the acceleration (*g*) of an accelerometer data logger instrumented to the participant. On the treadmill, all three locomotion speeds were undertaken both on the flat and at 1° incline. Both in the laboratory and over tarmac and grass, the speeds were undertaken in a randomised order, however the first speed was always a walk. Around half the participants undertook the treadmill conditions first. Soil penetration resistance along the grass verge was also recorded before each experimental session (mean±1SD: 0.945 ± 0.36 kg cm^{-3}). Participants wore sports clothing, and were free to add or remove appareil as the experiments progressed. Ambient temperature outside ranged from 5 to 14□ (determined retrospecively using public archive data at data.gov.uk) with no wind chill effect (wind speed was negligible during all experiments outside), and 18 to 20□ in the laboratory. None of the participants reported shivering or being cold during the experiments, indicating that temperature induced metabolic penalties were avoided through exercise induced heating (Jacobs et al., 1985, McArdle et al., 1976)).

The participants undertook each speed/surface combination for a minimum of four minutes and rested for five minutes after each jogging period. Participants were asked to walk at 3 km h^{-1} and 5 km h^{-1}, and to jog at 8 km h^{-1} (0.83, 1.39 and 2.22 ms^{-1} , respectively). To ensure a constant speed over grass and tarmac, participants were required to remain alongside an experimenter who set the pace with the queue from accurately-spaced cones along the walking route in conjunction with an auditory metronome (see Wilson et al., 2013b). The respiratory exchange quotient remained below 1.1 in all participants at all times indicating that they were always exercising within the bounds of their levels of aerobic fitness

Physiological measurements

Participants' $\dot{V}O_2$, rate of carbon dioxide output ($\dot{V}CO_2$), and respiratory exchange ratio (RER) were measured using a portable gas analyser (Oxycon Mobile, Jaeger), incorporating an oxygen paramagnetic analyser and a carbon dioxide infrared analyser. Values were initially measured at 'barometric pressure and temperature, saturated' (BPTS) and converted to 'standard temperature and pressure, dried' (STPD) using the Haldane transformation. This system involved continuous monitoring of breath-by-breath measurements utilising a lightweight and low-resistance facemask from which samples of the expired air were drawn through tubing at a constant rate. This breath-by-breath data provided confirmation that the participant reached steady state during each condition and RER did not reach 1 (indicating that anaerobic metabolism was negligible). Heart rate (f_H) was measured using a monitor (Polar CS100 wearlink and transmitter, Polar Electro) recording at 0.2 Hz. The f_H monitor was attached to the chest of the participant just above the sternum with a strap. Acceleration was measured using a Model X6-2 tri-axial data logger (Gulf Coast Data Concepts) recording 0 to ± 6 *g* at 12 Hz and 12-bit resolution. This logger was attached to the centre of the lumber region of the participant's back, using a Silastic® harness (Dow Corning Corporation, Midland, MI). The acceleration logger did not noticeably move relative to the body during exercise, ensuring that the logger recorded only acceleration attributable to body movement.

After the experiments, means of $\dot{V}O_2$ were calculated from the final minute or 30 s of each condition based on visualisation of the breath by breath $\dot{V}O_2$ data indicating when physiological steady state had been reached (Meijer et al., 1989, Terrier et al., 2001, Achten et al., 2002). This is usually after around 2 to 3 min, but can be less in reasonably fit individuals (Chilibeck et al., 1996, Whipp and Wasserman, 1972, ACSM, 2013), particularly when rest between conditions is relatively short. Subsequently, means of heart rate and acceleration data were calculated for the same periods.

Analysis of laboratory data

Movement of body parts results in movement of the body's centre of mass, and the latter has been shown to correlate with $\dot{V}O_2$ (Halsey et al., 2009a). Raw acceleration data from an instrumented acceleration data logger encapsulate two gravitational components: static acceleration due to gravity and dynamic acceleration due to body movement (Gleiss et al., 2011). Thus the acceleration of the body's centre of mass due to the movement of its body parts can be determined by recording acceleration experienced by the data logger attached to a fixed point on the body, such as the torso, and then from those data extracting an approximation of absolute *g* ($1\ g = 9.81\ \mathrm{m\ s^{-2}}$) due only to dynamic acceleration of the body in each of the three dimensions (Gleiss et al., 2011, Halsey et al., 2011a). This extraction was achieved in the present study by removing an approximation of the static acceleration calculated via a running mean, which spanned 30 data points, i.e. ~ 3 s; a suitable smoothing duration for many exercise scenarios (Halsey et al., 2009a, Green et al., 2009, Shepard et al., 2008). The resulting absolute dynamic values were then combined to produce two derivations of ‘dynamic body acceleration’. The absolute summation of the three axes of dynamic body acceleration is referred to as overall dynamic body acceleration (ODBA; (see Wilson et al., 2006, for more details), while the vectorial summation of the absolute dynamic values is termed vectorial dynamic body acceleration (VeDBA; (Qasem et al., 2012).

Statistical analyses

To provide a broad comparison between $\dot{V}O_2$ values observed in each experimental condition, means for each condition were calculated across speeds and participants. The percentage difference was then calculated as the absolute difference between each of the means and that observed on the treadmill at 0°. To investigate the differences in $\dot{V}O_2$ during pedestrian locomotion for each surface type (treadmill at 0°, treadmill at 1°, level grass, level tarmac) at each speed, mean $\dot{V}O_2$ values for each participant were submitted to a repeated measures general linear model (GLM) using the R package *nlme* (Pinheiro et al., 2012), with both surface type and speed included as factors: $\dot{V}O_2$ ~ surface_type + speed + surface_type:speed + individual[random]. The same procedure was used to compare *f*H, ODBA and VeDBA between surface types. A technical difficulty produced erroneous *f*H data for a single participant on grass and tarmac, so this participant’s *f*H data were removed from analysis.

To investigate the relationships between $\dot{V}O_2$ and each proxy and surface type, separate models were produced for each proxy (*f*H, ODBA and VeDBA): $\dot{V}O_2$ ~ surface_type + proxy + surface_type:proxy + individual[random]. Model fit was compared between each of the models using Akaike information criterion (AIC) calculated in R. Following this, the models were rerun with the interaction term removed to investigate how this affected the model fit.

The values of mean *f*H, ODBA and VeDBA measured on the treadmill at 0° incline were independently regressed against measured $\dot{V}O_2$ for each individual separately to generate individual-specific calibrations. These calibrations were then used to estimate $\dot{V}O_2$ from values recorded for each proxy during locomotion on tarmac and on grass. The absolute difference between the estimated value and the measured value recorded with the portable respirometer was calculated, and these values were compared to investigate whether the extent of the difference was influenced by surface type and speed. To make this comparison, we calculated the mean algebraic error, i.e., the mean of all positive and negative errors, to show the estimate error average across all participants, and the mean absolute error, which better reflects the error on an individual participant basis. This procedure was repeated for calibrations based on the 1° incline treadmill data.

Meta-analysis

To investigate the general trend in the published literature in terms of $\dot{V}O_2$ during locomotion on a treadmill compared to locomotion in the field, a meta-analysis following the principles set out in Cumming (2012) was undertaken using the accompanying ESCI meta-analysis software.

A review of the literature uncovered 12 published articles, dating from 1953 – 2015, that compared the metabolic costs for human locomotion on variable speed treadmills and field running. Studies were found using search terms such as ‘outdoor versus treadmill running’, ‘energy expenditure on treadmill’ and variations thereof. Some studies were uncovered as they referenced earlier studies. For inclusion in the meta-analysis, studies were required to compare treadmill locomotion to comparable conditions on a firm surface, provide values for $\dot{V}O_2$ via respirometry and report travel speeds to ensure this factor was broadly comparable between studies. Despite some differences in experimental protocol, 8 suitable studies were included in the meta-analysis (Table 1). Many of these studies tested different speeds and so values for $\dot{V}O_2$ obtained at speeds closest to 1.5 m s^{-1} (a fast walk) were used. The mean locomotion speed associated with the values used in the meta-analysis was 1.54 ±0.16 (1SD) m s^{-1}, and was usually the median speed tested. The mean $\dot{V}O_2$ and SD for the 0° treadmill and tarmac experiments calculated for all participants at 1.4 ms^{-1} (5 km hr^{-1}) during the present study were also included in the meta-analysis.

The meta-analysis was conducted using the mean mass-specific $\dot{V}O_2$ and SDs for treadmill and field locomotion in each study. A random effects model was used to somewhat account for differences in the experimental procedures between studies (as opposed to a fixed effects model, which assumes that all studies have exactly the same aim). This model estimated the mean effect size (the difference in mean mass-specific $\dot{V}O_2$ between the two conditions) and 95% confidence interval across the studies included. For a detailed account of this model, see Chapter 8 of Cumming (2012).

Results

In the current study, as would be expected there was a statistically significant increase in $\dot{V}O_2$ at higher speeds (F = 1844,0, df = 2, p <0.001). Surface type (treadmill, tarmac or grass) also affected $\dot{V}O_2$ (F = 26.4, df = 3, p <0.001). Mean $\dot{V}O_2$ was higher when participants locomoted on either of the outdoor surfaces than on the treadmill (Figure 2). Mean absolute $\dot{V}O_2$ across all speeds was 9.1% higher on tarmac and 17.7% higher on grass compared to a treadmill at 0° incline. *f*H, ODBA and VeDBA also increased with increasing speed of locomotion, again as expected. However, while *f*H and ODBA varied statistically significantly with surface type, surprisingly VeDBA did not (Table 2). This indicates that the proxies *f*H and ODBA were more sensitive to the surface underfoot than was VeDBA.

Accuracy of the treadmill derived calibrations

Analysis of the $\dot{V}O_2$–proxy relationships uncovered a differing effect of surface type. Where *f*H was the proxy there was no significant interaction with surface type, whereas both the ODBA and VeDBA models included a significant interaction effect (Table 3). This can be clearly seen upon inspection of the best fit regression lines between $\dot{V}O_2$ and each of the proxies for each surface type (Figure 4). For *f*H (Figure 4, panel a) the slopes are parallel, with the two outdoor surfaces consistently returning slightly higher $\dot{V}O_2$ values compared to the 0° or 1° treadmill. In contrast, the slopes for the surface conditions in the ODBA and VeDBA models differ (Figure 4, panels b and c); again, $\dot{V}O_2$ is higher at any given accelerometry value on the two outdoor surfaces but this difference becomes greater as $\dot{V}O_2$ increases. Comparison between the models indicated that ODBA provided marginally the best model given the data (AIC scores: 3065, 3093, 3211 for ODBA, VeDBA and fH repectively).

The relationships between $\dot{V}O_2$ and each proxy caused statistically significant differences in the magnitude of the algebraic estimate errors between conditions within each of the three proxies (Table 4; Figures 5 and 6). The aforementioned differences in how the relationships between $\dot{V}O_2$ and each proxy diverged across surface types explains the magnitudes of the algebraic errors in the estimates of $\dot{V}O_2$ in the field conditions based on the treadmill proxy calibrations. When the *f*H data were calibrated against $\dot{V}O_2$ data on the treadmill at 0°, for the tarmac and grass conditions the mean algebraic error for $\dot{V}O_2$ was underestimated by a fairly consistent amount of between 7.5 and 11.9%, depending upon suface type and speed (Figure 5). For both ODBA and VeDBA, $\dot{V}O_2$ treadmill calibrations at 0° produced relatively small (ODBA: between 2.1 and 8.1% mean algebraic error; VeDBA: between 1.5 and 7.8% mean algebraic error) underestimations in $\dot{V}O_2$ during slow locomotion (0.83 ms^{-1}) on both outdoor surfaces, and relatively high underestimations at the fastest speeds (ODBA: between 9.6 and 11% mean algebraic error; VeDBA: between 11.2 and 13.1% mean algebraic error at 2.2 ms^{-1}, Figure 5). ODBA was the proxy for which the mean algebraic error was lowest across all speeds and surfaces, followed by VeDBA and *f*H (5.7 ±13.8%, 6.5 ±13.9% and 8.8 ±24.4%, respectively).

fH data calibrated against $\dot{V}O_2$ data on the treadmill at 1° produced relatively large underestimations for locomotion on tarmac and grass (Figure 6). Mean algebraic errors for $\dot{V}O_2$ estimates on grass and tarmac were between 13.3 and 18.9% for all participants. 1° treadmill calibrations for ODBA and VeDBA also underestimated measured $\dot{V}O_2$ during locomotion on tarmac and grass, and again these underestimations increased with locomotion speed for both ODBA and VeDBA (ODBA: between 2.3 and 12.3% mean algebraic error; VeDBA: between 0.3 and 6.8% mean algebraic error at 0.83 ms^{-1}; ODBA: between 4.8 and 13.9% mean algebraic error; VeDBA: between 13.3 and 15.3% mean algebraic error at 2.22 ms^{-1}). Again, ODBA was the proxy that produced the lowest mean algebraic error across all speeds and surfaces in $\dot{V}O_2$ estimation when calibrated with the 1° treadmill data, followed by VeDBA and *f*H (6.2 ±10.9%, 6.7 ±11.1% and 12.1 ±26.3%, respectively).

Comparing 0° and 1° treadmill calibrations shows that using the 0° calibration in general produces less error when estimating $\dot{V}O_2$ from all three proxies for locomotion on tarmac and on grass.

Figure 7 illustrates the mean absolute error for calibrations produced on 0° and 1° treadmills. Estimates for $\dot{V}O_2$ from *f*H incurred consistently higher error than accelerometry and scaled negatively with speed, with the largest mean absolute error observed on tarmac at 3 km h^{-1} when calibrating from a treadmill at 1° (36.2% at 2.2 ms^{-1}). Mean absolute errors for accelerometry were lower (ODBA: between 11.3 and 14.9%; VeDBA: between 11.2 and 15.5% from 0° treadmill calibrations and ODBA: between 8.5 and 14.2%; VeDBA: between 8.2 and 15.2% from 1° treadmill calibrations).

Meta-analysis

The meta-analysis included 8 prior studies and the results of the present experiment. The standardised mean difference, calculated as the mean field $\dot{V}O_2$ minus the mean treadmill $\dot{V}O_2$, was $\mu = -0.662$ ml min^{-1} kg^{-1} (df = 9). The standard deviation of the population for the studies included in the meta-analysis was $\tau = 1.52$ ml min^{-1} kg^{-1}. In a meta-analysis, heterogeneity is the term given to the variability in outcomes between studies, for example due to differences in experimental protocol (Cumming, 2012). The weighted sum of squares between studies in the meta-analysis, a measure of heterogeneity, was Q = 114.86. The proportion of the total variance, which reflects the variation in the true effect size, was $I^2 = 93.03\%$, indicating that there is considerable heterogeneity between studies included in the meta-analysis. These figures reflect both the varied protocols adopted by the studies and/or the high variation in $\dot{V}O_2$ between individual participants. While the literature suggests that over 100 individuals are required to account for between-individual variability in metabolic rate (Hox, 2002), our meta-analysis contains data obtained for 191 individuals.

Discussion

In our study, $\dot{V}O_2$ measured by a mobile respiratory gas analyser during pedestrian locomotion showed a statistically significant difference between the surfaces investigated, with both tarmac and grass incurring a greater $\dot{V}O_2$ than the treadmill. However, somewhat in contrast, our meta-analysis highlighted that, at least at a fast walking speed, treadmill locomotion is energetically more expensive, although the effect size is small and influenced by the details of the experimental design. Our analysis of the accuracy of treadmill-calibrated proxies of $\dot{V}O_2$ (*f*H and accelerometry) to estimate $\dot{V}O_2$ on tarmac and grass indicates that under many situations a considerable measurement error is generated. The mean absolute error for all proxies was typically at least 10%, with errors from *f*H-based estimates often 20-40%. Mean algebraic errors were smaller; consistently around 10% for *ƒ*H and positively related to speed for accelerometry, ranging between around 0 and 15%.

Differences in cost of transport between field and treadmill locomotion

The results of the present experiments indicate that, at the three speeds tested, $\dot{V}O_2$ for locomotion 'in the field', whether on hard (tarmac) or soft (grass), is markedly higher than that for treadmill locomotion, even when the latter is on a 1° incline (which is often considered to account for the slight increases in resistance inherent in locomotion outdoors in ideal conditions; Jones and Doust, 1996). It would be reasonable to expect the energy expenditure of animals moving over grass to be greater than that for locomotion on a treadmill; a soft surface is deformed as an animal moves over it, demanding additional mechanical work (Coward and Halsey, 2014) and such a surface may also incur increased muscle-tendon work (while walking) or decreased muscle-tendon efficiency (while running ,Coward and Halsey, 2014, Lejeune et al., 1998). Indeed, for humans, measurements have shown that travel over less firm substrates incurs greater energy costs (Pinnington and Dawson, 2001) and reindeer must expend greater energy to travel over less firm tundra as opposed to densely packed substrate (White and Yousef, 1978). Harder surfaces such as tarmac allow more efficient locomotion by supporting more energy rebound than is typically experienced from a deforming surface (Kerdok et al., 2002, Hardin et al., 2004).

Van Ingen Schenau (1980) suggests that decreased air resistance during treadmill locomotion may explain the difference in $\dot{V}O_2$ between locomotion on a treadmill and on tarmac, although the contribution of air resistance in calm air at typical walking and jogging speeds is thought to be negligible (Pugh, 1970, Pugh, 1971, Davies, 1980). Another possibility is that there may be more energy return experienced during treadmill locomotion than any typical outdoor surfaces since the tread moves around a plank of wood. Otherwise, the energy savings experienced during treadmill locomotion in the present study might be due to differences in locomotion gait or kinematics between treadmill and 'free' walking and running. Pearce et al., (1983) suggest that less mechanical lift work is performed during treadmill locomotion due to kinematic changes in running gait, e.g. longer stance phases, shorter strides and higher cadence (Stolze et al., 1997, Alton et al., 1998, Warabi et al., 2005, Watt et al., 2010, Wearing et al., 2013) or possibly that part of this mechanical lift work is compensated for by the treadmill motor (Ralston, 1960). The effects on $\dot{V}O_2$ of the differences in locomotion kinetics between the treadmill and field running could be investigated using a treadmill with variable bed compliance set to match field conditions, such as the treadmill used in Hardin et al., (2004).

However, somewhat contrary to the empirical results of the present study, the meta-analysis for fast walking $\dot{V}O_2$ based on multiple previous studies along with the present data produced a standardised mean difference of 0.66 ml min^{-1} kg^{-1} (4.3% of the mean $\dot{V}O_2$ recorded across all studies). This indicates that, at least for the walking gait at around 1.5 ms^{-1}, treadmill locomotion requires a marginally greater $\dot{V}O_2$ than does locomotion on tarmac or an athletics track (Figure 3). This increase is estimated to be around 50 ml min^{-1} for an average-sized adult, which equates to about 1 kJ min^{-1} or 0.2 kcal min^{-1}; for most contexts this can be considered negligible.

Massaad et al. (Massaad et al., 2007) offer some support to the conclusion of the meta-analysis. They found the kinematic differences observed on the treadmill (i.e. higher cadence and shorter stride lengths when running on the treadmill, Stolze et al., 1997, Alton et al., 1998, Warabi et al., 2005, Watt et al., 2010, Wearing et al., 2013) resulted in decreased vertical mass displacement (i.e. a flatter trajectory), which in fact results in increased energy expenditure _by requiring greater mechanical work be performed at the hip, knee and ankle joints (Gordon et al., 2009). Nonetheless, the fact that kinematic differences were implicated in both decreased (Pearce et al., 1983) and increased energey expenditure on the treadmill (Massaad et al., 2007) suggests that other factors may be involved to produce the disparity in the studies on this topic (including in our present experimental results).

The most likely explanation for why the overall conclusion of the meta-analysis (similar energy costs to walk on the treadmill or on a firm surface ‘in the field’) somewhat contradicts the results of the experiment performed in the present study, is provided by considering the differing protocols under which field locomotion was tested between studies. Some of the studies required participants to walk around an athletics track (Barnett et al., 2015), and in each case $\dot{V}O_2$ was higher on the treadmill, which ultimately led to overestimation of both speed and $\dot{V}O_2$ over-ground. Unfortunately, many of the studies provided little information on the exact type of substrate of these tracks (e.g. Berryman et al., 2012, Yngve et al., 2003).

However, the rubber surfaces commonly found on athletics tracks may provide energy return that improves energy economies of locomotion (Kerdok et al., 2002). If studies that did or may have used a rubberised athletics track (i.e. reported route distances that might well have been traversed on an athletics track: 105 m, 140 m and 400 m; Parvataneni et al., 2009, Berryman et al., 2012, Barnett et al., 2015) are excluded from the meta-analysis, the standardised mean difference between treadmill and field locomotion decreases to just 0.03 ml min^{-1} kg^{-1} (0.21% of mean $\dot{V}O_2$ across all studies) with an I^2 of 75.03% (Figure 3).

Studies that explicitly compare the rate of energy expenditure on treadmill and athletic track surfaces are scarce. Wee et al., (2016) found no significant difference in $\dot{V}O_2$ between locomotion on a Mondo athletics track and a treadmill, although athletes reported a higher rate of perceived exertion on the motorized treadmill. However, that study involved participants travelling at speeds far higher (3.3 to 4.4 ms^{-1}) than that tested in the meta-analysis.

During field trials in the present study, participants were required to travel along a straight-line course approximately 60 m long. Once participants reached the end of the course, they made a 180° turn during which they were required to maintain pace. It is likely that the increased energy expenditure to perform these turns (Wilson et al., 2013b) increased the measured $\dot{V}O_2$ for the field conditions. Pearce et al. (1983) used a similar linear outdoor course (see Table 1) and their conclusions concur with those obtained in the present study. The studies that found treadmill locomotion to be more energetically expensive incorporated elliptical or large circular tracks with less abrupt turns (e.g. Murray et al., 1985, Barnett et al., 2015); gentler turns require less energy (Wilson et al., 2013b). Thus, the difference in the frequency and extent of turns in the protocols between the studies may offer some explanation for the varied conclusions found in the literature and the heterogeneity observed in the meta-analysis. If studies that incorporated abrupt turns are excluded from the meta-analysis (the results of the present experiment and Pearce et al., 1983), the standardised mean difference changes to 1.2 ml min^{-1} kg^{-1} more during treadmill locomotion (7.6% of mean $\dot{V}O_2$ measured across all studies) with an I^2 of 90.75% (Figure 3). The direction of effect is the same as that reported by the original meta-analyis but with a greater magnitude, thus offering some supporting evidence that locomotion on a treadmill incurs a slightly higher $\dot{V}O_2$ than does locomotion on outdoor surfaces.

Two of the studies that concluded locomotion on a treadmill incurs higher metabolic costs utilised participants over the age of 70 years (Parvataneni et al., 2009, Berryman et al., 2012). Older individuals tend to exhibit gait disorders (Waters and Mulroy, 1999), and recruit a greater proportion of their motor units at a given walking speed, utilising a higher percentage of fast twitch fibres (Martin et al., 1992). They may also have reduced gait stability and balance (Hausdorff et al., 1997, Woledge et al., 2005, Mian et al., 2006). During locomotion on a treadmill, the environment is static (Lavcanska et al., 2005) but proprioceptive information is received from moving muscles whilst optic flow is constant (Dal et al., 2010). Such a mismatch between sensory inputs may affect walking speed and motor output (Mulavara et al., 2005). Indeed, Dal et al. (2010) observed such an effect, as they found that the preferred walking speed determined on a treadmill was significantly lower to that observed during field locomotion. These studies suggest that treadmill locomotion may neccessitate greater balance and coordination. Given that both balance and coordination tend to deterioriate with age, older participants may incurr greater energetic penalties on a treadmill by adopting a less efficient gait. This may explain the observation of increased energy costs of treadmill locomotion in studies that involved participants above 70 years of age (Parvataneni et al., 2009, Berryman et al., 2012).

The results of Greig et al (1993) support this hypothesis. During a test comparing the energy costs of treadmill locomotion with that during locomotion down a corridor that involved groups of elderly participants (71-80 years) and young, healthy volunteers (21-37 years), only the elderly group showed increased heart rate and step rate on the treadmill. This suggests that the pattern in the literature towards the conclusion that treadmill locomotion incurs greater metabolic costs is at least partly influenced by the synergistic effect of age. If the studies that used older participants are removed from the meta-analysis, the standardised mean difference is -0.36 ml min^{-1} kg^{-1} more energy used on the treadmill (6.4% of mean $\dot{V}O_2$ measured across all studies).

In conclusion, the difference in $\dot{V}O_2$ between the treadmill and firm surfaces ‘in the field’ is typically small; 4.3% of mean $\dot{V}O_2$ across all the studies analysed (Figure 3). Our meta-analyses suggest that key details of the protocols underlying measurements of pedestrian locomotion on the treadmill and firm surfaces in the field can influence which, if either, of these conditions is the marginally more energetically expensive.

How accurate are treadmill-calibrated proxies of energy expenditure in the field?

As shown in the present study and previous studies, $\dot{V}O_2$ during locomotion in the field may be different to that on a treadmill. This raises the question as to whether the proxies *f*H, ODBA and VeDBA, having been calibrated with $\dot{V}O_2$ on the treadmill, are able to provide accurate estimates for $\dot{V}O_2$ in the field despite differences in the substrate underfoot (and perhaps other differences such as gait kinematics). Researchers have investigated whether calibrations of *f*H with $\dot{V}O_2$ are accurate across different environments in various animal species, but predominantly where the $\dot{V}O_2$-fH relationships are moderated by stress levels rather than surface type (Bisson et al., 2009, Cyr et al., 2009, Groscolas et al., 2010). Barnett et al., (2015) found that treadmill calibrations of $\dot{V}O_2$ against Actigraph counts (derived from measures of acceleration over a specified epoch) produced appreciable over-estimations when applied to participants travelling outdoors at a range of speeds (see Table 1).

The models describing the variability in $\dot{V}O_2$ according to both the proxies and the surface conditions showed that in all cases the relationship between $\dot{V}O_2$ and the proxy was very similar for the treadmill at 0° and 1° incline, and very similar for the two outdoor surfaces (cut grass and tarmac), but these two couplets of relationships differed. How they differed contrasted for accelerometry and for *f*H. Specifically, there was an interaction effect between ODBA or VeDBA (converging lines of best fit for each surface type; Figure 4, panels b and c) and surface condition, while in contrast no interaction effect was present for *f*H (parallel lines of best fit; Figure 4, panel a). For *f*H, this manifests as a divergence in the relationship between $\dot{V}O_2$ and accelerometry as the values of accelerometry, and hence locomotion speed, increase, with a steeper slope gradient for the two outdoor surfaces. For *f*H there are, in absolute terms, consistently higher values of $\dot{V}O_2$ for any given accelerometry value (and hence any given locomotion speed) for the outdoor surfaces.

Consequently, across all the proxies tested, the treadmill calibrations of $\dot{V}O_2$, at both 0° and 1° incline, on average produce underestimations of $\dot{V}O_2$ during field locomotion (Figures 5 and 6). Figure 5 shows that the $\dot{V}O_2$ estimated from ODBA and VeDBA calibrated with a treadmill at 0° is on average accurate for slow walking (3 km hr^{-1}) on tarmac but underestimates $\dot{V}O_2$ by ~5 and 13% for fast walking (5 km h^{-1}) and slow jogging (8 km h^{-1}), respectively. Underestimates of $\dot{V}O_2$ are marginally greater on grass, and on this surface type even for walking, these calibrations underestimate $\dot{V}O_2$ by more than 5% on average.

On average *f*H typically underestimated $\dot{V}O_2$ more than did ODBA or VeDBA, at all speeds and surface conditions tested, but generally this difference was greatest at slower speeds. The underestimates are over 10% in nearly all conditions, and nearly 20% for slow walking on tarmac. The fact that ODBA and VeDBA usually produced more accurate estimates of $\dot{V}O_2$ than did *f*H might be surprising because the accuracy of accelerometry data is potentially affected by movement of the logger relative to the animal's body (i.e. over and above movement due to the animal's body) (Preston et al., 2012), which could differ between surfaces. However, a similar finding to the present study was reported for chickens; accelerometry outperformed *f*H as a proxy for energy expenditure when the animals were active (Green et al., 2009).

In terms of absolute error, *f*H was again typically less accurate than accelerometry (Figure 7). This suggests that on an individual person basis, on average *f*H returned less accurate estimates of $\dot{V}O_2$ than did accelerometry. The mean absolute errors associated with *f*H were always greater than 10% and often greater than 20%. In contrast, the mean absolute errors for accelerometry were sometimes less than 10% and never greater than 20%. A further difference between the proxies is that the magnitude of the absolute error markedly decreased with speed for *f*H for every surface type, but tended to mildly increase with speed for accelerometry. This can be explained from inspection of Figure 4. The scatter around the lines of best fit between $\dot{V}O_2$ and accelerometry is greater at the higher values of ODBA or VeDBA, i.e. at the high locomotion speeds, resulting in greater mean absolute errors. For *f*H, the scatter is fairly consistent across the range of *f*H values and thus at the higher locomotion speeds when *f*H is therefore also higher, the mean absolute error as a percentage of the true value diminishes.

That treadmill-based calibrations can include such errors when applied to estimating $\dot{V}O_2$ in the field may be problematic for studies focussing on estimating $\dot{V}O_2$ during intense activity, such as during prey capture (e.g. Wilson et al., 2013a, Viviant et al., 2010, Williams et al., 2014). However, the majority of animals move through their environments at relatively low speeds most of the time, in order to conserve energy or remain hidden (Moen, 1976, Kenagy and Hoyt, 1989, Wickler et al., 2000), and the present findings indicate that in this context estimate errors may be small, particularly when accelerometry is the proxy. Our study suggests that in instances where animals are expected to employ a range of higher speeds or activity types, gait-specific and activity-specific calibrations for $\dot{V}O_2$ should be used to minimise error (Jeanniard-du-Dot et al., 2016, Volpov et al., 2015), and might be particularly valuable when subject animals are encountering complex, heterogeneous environments (Kareiva, 1990, Wiens et al., 1993, Morales and Ellner, 2002).

Finally, we must flag up the surprising result that one of the derivations of accelerometry data – VeDBA – did not differ statistically significantly between surface types, in contrast to ODBA, and also *ƒ*H. P values should be interpreted with great caution (Halsey et al., 2015a), however, and VeDBA estimated similar values of $\dot{V}O_2$ to those estimated by ODBA (Figures 5 and 6). Nonetheless, this statistical result suggests some evidence that in the scenario of the present study at least, VeDBA is less sensitive to changes in substrate type than are the other proxies. Using VeDBA as an uncalibrated proxy of gait kinematics, underfoot substrate or (qualified) locomotion energetics may be less effective than employing ODBA or *ƒ*H.

Final thoughts

The results of the present study should make researchers question and consider how accurately a laboratory-derived proxy calibration of $\dot{V}O_2$ will estimate $\dot{V}O_2$ during terrestrial locomotion of a human or other animal in the field, including when field conditions appear comensurate with the treadmill; flat firm ground in a wind-free environment. Our data suggest that at relatively low speeds the errors may tend to be smaller than at relatively higher speeds, and thus treadmill calibrations may perform better for animals that mostly locomote at their lower speeds. On the other hand, where the substrate underfoot is more different to the treadmill (such as snow or sand; Crête and Larivière, 2003, Pandolf et al., 1976, Lejeune et al., 1998, Pinnington and Dawson, 2001), or is not relatively flat (Halsey et al., 2008, Halsey and White, 2017), it is possible that estimate errors of $\dot{V}O_2$ will be greater. We stress, however, that this is far from certain because it depends on how changes in $\dot{V}O_2$ due to the substrate are recognised by changes in the measured proxy.

Numerous studies have derived laboratory-based energetics calibrations for aquatic and volant animals using shallow dive tanks and wind tunnels (Green, 2011, Ward et al., 2001, Halsey et al., 2007). Similarly to the limitations in accuracy of using lab-based terrestiral locomotion protocols to estimate field-based terrestrial locomotion, the same is likely for other forms of locomotion, and indeed may be greater given the particular difficulties in simulating free-ranging swimming, diving and flying in the laboratory (Elliot et al., 2013, Hansen and Ricklefs, 2004).

Where more accurate estimates of field energy expenditure are desired, we suggest that researchers consider combining proxies to record more data types related to metabolic rate, most obviously fH and accelerometry (Elliot, 2016). In a study of sockeye salmon, *f*H and accelerometry in combination proved a considerably better proxy for energy expenditure than did fH or accelerometry alone (Clark et al., 2010). This finding was mirrored in an early treadmill calibration study comparing accelerometry with *f*H (Halsey et al., 2008). Data logger designs can now incorporate both *f*H and accelerometry, and although intermittent sampling may be required to preserve battery life, the combined data sets have proved insightful (Bishop et al., 2015). Furthermore, doubly-labelled water may be used as a potential calibrator, using time-specific activity budgets to calculate the energy expenditure of shorter-lived behaviours (Elliot et al., 2013).

Acknowledgements

Tom Reeve provided technical support during data collection. Useful chats with Craig White helped to conceive the project idea. ORB was funded by a Post-Doctoral Fellowship from the Alexander von Humboldt Foundation. This paper is dedicated to Taren and Emrys.

References

ACHTEN, J., GLEESON, M. & JEUKENDRUP, A. 2002. Determination of the exercise intensity that elicits maximal fat oxidation. *Medicine and Science in Sports and Exercise,* 34**,** 92-97.

ACSM 2013. *American College of Sports Medicine's guidelines for exercise testing and prescription*, Lippincott Williams & Wilkins.

ALTON, F., BALDEY, L., CAPLAN, S. & MORRISSEY, M. C. 1998. A kinematic comparison of overground and treadmill walking. *Clinical Biomechanics,* 13**,** 434-440.

BARNETT, A., CERIN, E., VANDELANOTTE, C., MATSUMOTO, A. & JENKINS, D. 2015. Validity of treadmill- and track-based individual calibration methods for estimating free-living walking speed and VO_2 using the Actigraph accelerometer. *BMC Sports Science, Medicine and Rehabilitation,* 7**,** 29.

BERRYMAN, N., GAYDA, M., NIGAM, A., JUNEAU, M., BHERER, L. & BOSQUET, L. 2012. Comparison of the metabolic energy cost of overground and treadmill walking in older adults. *European Journal of Applied Physiology,* 112**,** 1613-1620.

BEVAN, R. M., WOAKES, A. J., BUTLER, P. J. & CROXALL, J. P. 1995. Heart rate and oxygen consumption of exercising gentoo penguins. *Physiological Zoology,* 68**,** 855-877.

BIDDER, O. R., QASEM, L. A. & WILSON, R. P. 2012. On Higher Ground: How Well Can Dynamic Body Acceleration Derermine Speed in Variable Terrain? *PLoS One,* DOI: 10.1371/journal.pone.0050556.

BISHOP, C. M., SPIVEY, R. J., HAWKES, L. A., BATBAYAR, N., CHUA, B., FRAPPELL, P. B., MILSOM, W. K., NATSAGDORJ, T., NEWMAN, S. H., SCOTT, G. R., TAKEKAWA, J. Y., WIKELSKI, M. & BUTLER, P. J. 2015. The roller coaster flight strategy of bar-headed geese conserves energy during Himalayan migrations. *Science,* 347**,** 250-254.

BISSON, I., BUTLER, L. K., HAYDEN, T. J., ROMERO, L. M. & WIKELSKI, M. 2009. No energetic costs of anthropogenic disturbance in a songbird. *Proceeding of The Royal Society: Biological Sciences,* 276**,** 961-969.

BRAGE, S., BRAGE, N., EKELUND, U., LUAN, J. A., FRANKS, P. W., FROBERG, K. & WAREHAM, N. J. 2006. Effect of combined movement and heart rate monitor placement on physical activity estimates during treadmill locomotion and free-living. *European Journal of Applied Physiology,* 96**,** 517-524.

BUTLER, P. J., GREENWOOD, P. J., BOYD, I. L. & SPEAKMAN, R. 2004. Measuring metabolic rate in the field: the pros and cons of the doubly labelled water and heart rate methods. *Functional Ecology,* 18**,** 168-183.

CHILIBECK, P. D., PATERSON, D. H., PETRELLA, R. J. & CUNNINGHAM, D. A. 1996. The Influence of Age and Cardiorespiratory Fitness on Kinetics of Oxygen Uptake. *Canadian Journal of Applied Physiology,* 21**,** 185-196.

CLARK, T. D., SANDBLOM, E., HINCH, S. G., PATTERSON, B. R., FRAPPELL, P. B. & FARRELL, A. P. 2010. Simultaneous biologging of heart rate and acceleration, and their relationship with energy expenditure in free-swimming sockeye salmon (Oncorhynchus nerka). *Journal of Comparitive Physiology B,* 180**,** 673-684.

COWARD, S. R. L. & HALSEY, L. G. 2014. Energy expended during horizontal jumping: investigating the effects of surface compliance. *Biology Open,* 3**,** 815-820.

CRÊTE, M. & LARIVIÊRE, S. 2003. Estimating the costs of locomotion in snow for coyotes. *Canadian Journal of Zoology,* 81**,** 1808-1814.

CUMMING, G. 2012. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis,* New York, Routledge.

CYR, N. E., DICKENS, M. J. & ROMERO, L. M. 2009. Heart rate and heart-rate variability responses to acute and chronic stress in a wild-caught passerine bird. *Physiological and Biochemical Zoology,* 82**,** 332-344.

DAL, U., ERDOGAN, T., RESITOGLU, B. & BEYDAGI, H. 2010. Determination of preferred walking speed on treadmill may lead to high oxygen cost on treadmill walking. *Gait and Posture,* 31**,** 366-369.

DANIELS, F., VANDERBIE, J. H. & WINSMAN, F. R. 1953. Energy cost of treadmill walking compared to road walking. Report No. 220 of Environmental Protection Division, Q. M., Research and Development Laboratory, Natick, MA.

DAVIES, C. T. M. 1980. Effects of wind assistance and resistance on the forward motion of a runner. *Journal of Applied Physiology,* 48**,** 702-709.

ELLIOT, K. H. 2016. Measurement of flying and diving metabolic rate in wild animals: Review and recommendations. *Comparative Biochemistry and Physiology Part A: Molecular & Intergrative Physiology,* 202**,** 63-77.

ELLIOT, K. H., RICKELEFS, R. E., GASTON, A. J., HATCH, S. A., SPEAKMAN, J. R. & DAVOREN, G. K. 2013. High flight costs, but low dive costs, in auks support the biomechanical hypothesis for flightlessness in penguins. *Proceedings of the National Academy of Sciences,* 110**,** 9380-9384.

FANCY, S. G. & WHITE, R. G. 1987. Energy expenditures for locomotion by barren-ground caribou. *Canadian Journal of Zoology,* 65**,** 122-128.

FROGET, G., BUTLER, P. J., HANDRICH, Y. & WOAKES, A. J. 2001. Heart rate as an indicator of oxygen consumption: influence of body condition in the king penguin. *Journal of Experimental Biology,* 204**,** 2133-2144.

GEFEN, E. 2011. The relative importance of respiratory water loss in scorpions is correlated with species habitat type and activity pattern. *Physiological and Biochemical Zoology,* 84**,** 68-76.

GLEISS, A. C., WILSON, R. P. & SHEPARD, E. L. C. 2011. Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution,* 2**,** 23-33.

GORDON, K. E., FERRIS, D. P. & KUO, A. D. 2009. Metabolic and mechanical energy costs of reducing vertical center of mass movement during gait. *Arch Phys Med Rehabil,* 90.

GREEN, J. A. 2011. The heart rate method for estimating metabolic rate: Review and recommendations. *Comparative Biochemistry and Physiology Part A: Molecular & Intergrative Physiology,* 158**,** 287-304.

GREEN, J. A., BUTLER, P. J., WOAKES, A. J., BOYD, I. L. & HOLDER, R. L. 2001. Heart rate and rate of oxygen consumption of exercising macaroni penguins. *Journal of Experimental Biology,* 204**,** 673-684.

GREEN, J. A., HALSEY, L. G., WILSON, R. P. & FRAPPELL, P. B. 2009. Estimating energy expenditure of animals using the accelerometry technique: activity, inactivity and comparison with the heart-rate technique. *Journal of Experimental Biology,* 212**,** 471-482.

GREIG, C., BUTLER, F., SKELTON, D., MAHMUD, S. & YOUNG, A. 1993. Treadmill walking in old age may not reproduce the real life situation. *Journal of the American Geriatrics Society,* 41**,** 15-18.

GROSCOLAS, R., VIERA, V., GUERIN, N., HANDRICH, Y. & COTE, S. D. 2010. Heart rate as a predictor of energy expenditure in undisturbed fasting and incubating penguins. *Journal of Experimental Biology,* 213**,** 153-160.

HALSEY, L. G. 2016. Terrestrial movement energetics: current knowledge and its application to the optimising animal. *Journal of Experimental Biology,* 219**,** 1424-1431.

HALSEY, L. G., CURRAN-EVERETT, D., VOWLER, S. L. & DRUMMOND, G. B. 2015a. The fickle P value generates irreproducible results. *Nature Methods,* 12**,** 179-185.

HALSEY, L. G., FAHLMAN, A., HANDRICH, Y., SCHMIDT, A., WOAKES, A. J. & BUTLER, P. J. 2007. How accurately can we estimate energetic costs in a marine top predator, the king penguin? *Zoology,* 110**,** 81-92.

HALSEY, L. G., GREEN, A. J., WILSON, R. & FRAPPELL, P. B. 2009a. Accelerometry to estimate energy expenditure during activity: best practice with data loggers. *Physiological and Biochemical Zoology,* 82**,** 396-404.

HALSEY, L. G., MATTHEWS, P. G. D., REZENDE, E. L., CHAUVAUD, L. & ROBSON, A. A. 2015b. The interactions between temperature and activity levels in driving metabolic rate: theory, with empirical validation from contrasting ectotherms. *Oecologia,* 177**,** 1117-1129.

HALSEY, L. G., SHEPARD, E. L. C., HULSTON, C. J., VENABLES, M. C., WHITE, C. R., JEUKENDRUP, A. E. & WILSON, R. P. 2008. Acceleration versus heart rate for estimating energy expenditure and speed during locomotion in animals: Tests with an easy model species, *Homo sapiens*. *Zoology,* 111**,** 231-241.

HALSEY, L. G., SHEPARD, E. L. C., QUINTANA, F., LAICH, A. G., GREEN, J. A. & WILSON, R. P. 2009b. The relationship between oxygen consumption and body acceleration in a range of species. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology,* 152**,** 197-202.

HALSEY, L. G., SHEPARD, E. L. C. & WILSON, R. P. 2011a. Assessing the development and application of the accelerometry technique for estimating energy expenditure. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology,* 158**,** 305-314.

HALSEY, L. G. & WHITE, C. R. 2017. A different angle: comparative analyses of whole-animal transport costs running uphill. *Journal of Experimental Biology,* 220**,** 161-166.

HALSEY, L. G., WHITE, C. R., ENSTIPP, M. R., WILSON, R. P., BUTLER, P. J., MARTIN, G. R., GRÉMILLET, D. & JONES, D. R. 2011b. Assessing the validity of the Accelerometry Technique for Estimating the Energy Expenditure of Diving Double-Crested Cormorants *Phalacrocorax auritus*. *Physiological and Biochemical Zoology,* 84**,** 230-237.

HANSEN, E. S. & RICKLEFS, R. E. 2004. Foraging by deep-diving birds is not constrained by an aerobic diving limit: a model of avian depth-dependent diving metabolic rate. *The American Naturalist,* 163**,** 358-374.

HARDIN, E. C., VAN DEN BOGERT, A. J. & HAMILL, J. 2004. Kinematic adaptations during running: effects of footwear, surface, and duration. *Medicine & Science in Sports & Exercise,* 36**,** 838-844.

HAUSDORFF, J. M., EDELBERG, H. K., MITCHELL, S. L., GOLDBERGER, A. L. & WEI, J. Y. 1997. Increased gait unsteadiness in community-dwelling elderly fallers. *Archives of Physical Medicine and Rehabilitation,* 78**,** 278-283.

HAWKINS, P. A., BUTLER, P. J., WOAKES, A. J. & SPEAKMAN, J. R. 2000. Estimation of the rate of oxygen consumption of the common eider duck (Somateria mollissima), with some measurements of heart rate during voluntary dives. *Journal of Experimental Biology,* 203**,** 2819-2832.

HOX, J. 2002. *Multilevel analysis: techniques and applications,* Mahwah, NJ, Lawrence Erlbaum Associates.

HOYT, D. F. & KENAGY, G. J. 1988. Energy Costs of Walking and Running Gaits and Their Aerobic Limits in Golden-Mantled Ground Squirrels. *Physiological Zoology,* 61**,** 34-40.

JACOBS, I., ROMET, T. T. & KERRIGAN-BROWN, D. E. 1985. Muscle glycogen depletion during exercise at 9C and 21C. *European Journal of Applied Physiology,* 54**,** 35-39.

JANKOWSKI, L. W., FERGUSON, R. J., LANGELIER, M., CHANIOLIS, L. N. & CHOQUETTE, G. 1972. Accuracy of methods of estimating oxygen cost of walking in coronary patients. *Journal of Applied Physiology,* 33**,** 672-673.

JEANNIARD-DU-DOT, T., GUINET, C., ARNOULD, J. P. Y., SPEAKMAN, J. R. & TRITES, A. W. 2016. Accelerometers can measure total and activity-specific energy expenditures in free-ranging marine mammals only if linked to time-activity budgets. *Functional Ecology,* doi:10.1111/1365-2435.12729.

JONES, A. M. & DOUST, J. H. 1996. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *Journal of Sports Sciences,* 14**,** 321-327.

KAREIVA, P. M. 1990. Population dynamics in spatially complex environments: theory and data. *Philosophical Transactions of the Royal Society, Biological Sciences,* 330**,** 175-190.

KENAGY, G. J. & HOYT, D. F. 1989. Speed and time-energy budget for locomotion in golden-mantled ground squirrels. *Ecology,* 70**,** 1834-1839.

KERDOK, A. E., BIEWENER, A. A., MCMAHON, T. A., WEYAND, P. G. & HERR, H. M. 2002. Energetics and mechanics of human running on surfaces of different stiffnesses. *Journal of Applied Physiology,* 92**,** 469-478.

KNAPIK, J. J., REYNOLDS, K. L. & HARMAN, E. 2004. Soldier load carriage: historical, physiological, biomechanical and medical aspects. *Military medicine,* 169**,** 45-56.

LAVCANSKA, V., TAYLOR, N. F. & SCHACHE, A. G. 2005. Familiarization to treadmill running in young unimpaired adults. *Human Movement Science,* 24**,** 544-557.

LEJEUNE, T. M., WILLEMS, P. A. & HEGLUND, N. C. 1998. Mechanics and energetics of human locomotion on sand. *Journal of Experimental Biology,* 201**,** 2071-80.

MARTIN, P. E., ROTHSTEIN, D. E. & LARISH, D. D. 1992. Effects of age and physical activity status on the speed-aerobic demand relationship of walking. *Journal of Applied Physiology,* 73**,** 200-206.

MASSAAD, F., LEJEUNE, T. M. & DETREMBLEUR, C. 2007. The up and down bobbing of human walking: a compromise between muscle work and efficiency. *J Physiol,* 582.

MCARDLE, W. D., MAGEL, J. R., LESMES, G. R. & PECHAR, G. S. 1976. Metabolic and cardiovascular adjustment to work in air and water at 18, 25 and 33 °C. *Journal of Applied Physiology,* 40**,** 85-90.

MEIJER, G. A., WESTERTERP, K. R., KOPER, H. & TEN HOOR, F. 1989. Assessment of energy expenditure by recording heart rate and body acceleration. *Medicine & Science in Sports & Exercise,* 21**,** 343.

MIAN, O. S., THOM, J. M., ARDIGO, L. P., NARICI, M. V. & MINETTI, A. E. 2006. Metabolic cost, mechanical work, and efficiency during walking in young and older men. *Acta Physiologica,* 186**,** 127-139.

MOEN, A. N. 1976. Energy conservation by white-tailed deet in the winter. *Ecology,* 57**,** 192-198.

MORALES, J. M. & ELLNER, S. P. 2002. Scaling up animal movements in heterogeneous landscapes: the importance of behavior. *Ecology,* 83**,** 2240-2247.

MULAVARA, A. P., RICHARDS, J. T., RUTTLEY, T., MARSHBURN, A., NOMURA, Y. & BLOOMBERG, J. J. 2005. Exposure to a rotating virtual environment during treadmill locomotion causes adaptation in heading direction. *Experimental Brain Research,* 166**,** 210-219.

MURRAY, M. P., SPURR, G. B., SEPIC, S. B., GARDNER, G. M. & MOLLINGER, L. A. 1985. Treadmill vs. floor walking: kinematics, electromyogram, and heart rate. *Journal of Applied Physiology,* 59**,** 87-91.

NICHOLS, J. F., MORGAN, C. G., CHABOT, L. E., SALLIS, J. F. & CALFAS, K. J. 2000. Assessment of Physical Activity with the Computer Science and Applications, Inc., Accelerometer: Laboratory versus Field Validation. *Research Quarterly for Exercise and Sport,* 71**,** 36-43.

NOLET, B. A., BUTLER, P. J., MASMAN, D. & WOAKES, A. J. 1992. Estimation of Daily Energy Expenditure from Heart Rate and Doubly Labeled Water in Exercising Geese. *Physiological Zoology,* 65**,** 1188-1216.

PANDOLF, K. B., HAISMAN, M. F. & GOLDMAN, R. F. 1976. Metabolic Energy Expenditure and Terrain Coefficients for Walking on Snow. *Ergonomics,* 19**,** 683-690.

PARVATANENI, K., PLOEG, L., OLNEY, S. J. & BROUWER, B. 2009. Kinematic, kinetic and metabolic parameters of treadmill versus overground walking in healthy older adults. *Clinical Biomechanics,* 24**,** 95-100.

PEARCE, M. E., CUNNINGHAM, D. A., DONNER, A. P., RECHNITZER, P. A., FULLERTON, G. M. & HOWARD, J. H. 1983. Energy Cost of Treadmill and Floor Walking at Self-selected Paces. *European Journal of Applied Physiology,* 52**,** 115-119.

PINHEIRO, J., BATES, D., DEBROY, S., SARKER, D. & TEAM, R. C. 2012. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-103. *Available at: http://cran.r-project. org/web/packages/nlme/nlme.pdf.*

PINNINGTON, H. C. & DAWSON, B. 2001. The energy cost of running on grass compared to soft dry beach sand. *Journal of Science and Medicine in Sport,* 4**,** 416-430.

PRESTON, T., BALTZER, W. & TROST, S. 2012. Accelerometer validity and placement for detection of changes in physical activity in dogs under controlled conditions on a treadmill. *Research in Veterinary Science,* 93**,** 412-416.

PUGH, L. G. C. E. 1970. Oxygen intake in track and treadmill running with observations on the effect of air resistance. *The Journal of Physiology,* 207**,** 823-835.

PUGH, L. G. C. E. 1971. The influence of wind resistance in running and walking and the mechanical efficiency of work against horizontal and vertical forces. *The Journal of Physiology,* 213**,** 255-276.

QASEM, L., CARDEW, A., WILSON, A., GRIFFITHS, I., HALSEY, L. G., SHEPARD, E. L. C., GLEISS, A. C. & WILSON, R. 2012. Tri-Axial Dynamic Acceleration as a Proxy for Animal Energy Expenditure; Should We Be Summing Values or Calculating the Vector? *PLoS ONE,* 7**,** e31187.

RALSTON, H. J. 1960. Comparison of energy expenditure during treadmill walking and floor walking. *Journal of Applied Physiology,* 15**,** 1156.

REZENDE, E. L., GOMES, F., CHAPPELL, M. A. & GARLAND JR., T. 2009. Running Behavior and Its Energy Cost in Mice Selectively Bred for High Voluntary Locomotor Activity. *Physiological and Biochemical Zoology,* 82**,** 662-679.

SCANTLEBURY, D. M., MILLS, M. G. L., WILSON, R. P., WILSON, J. W., MILLS, M. E. J., DURANT, S. M., BENNETT, N. C., BRADFORD, P., MARKS, N. J. & SPEAKMAN, J. R. 2014. Flexible energetics of cheetah hunting strategies provide resistance against kleptoparasitism. *Science,* 346**,** 79-81.

SHEPARD, E. L. C., WILSON, R. P., HALSEY, L. G., QUINTANA, F., GOMEZ LAICH, A., GLEISS, A. C., LIEBSCH, N., MYERS, A. & NORMAN, B. 2008. Derivation of body motion via appropriate smoothing of acceleration data. *Aquatic Biology,* 4**,** 235-241.

STOLZE, H., KUHTZ-BUSCHBECH, J. P., MONDWURF, C., BOCZEK FUNCKE, A., JOHNK, K., DEUSCHL, G. & ILLERT, M. 1997. Gait analysis during treadmill and overground locomotion in children and adults. *Electroencephalography and Clinical Neurophysiology,* 105**,** 490-497.

TERRIER, P., AMINIAN, K. & SCHUTZ, Y. 2001. Can accelerometry accurately predict the energy cost of uphill/downhill walking? *Ergonomics,* 44**,** 48-62.

VAN INGEN SCHENAU, G. J. 1980. Some fundamental aspects of the biomechanics of overground versus treadmill locomotion. *Medicine and Science in Sports and Exercise,* 12**,** 257-261.

VIVIANT, M., TRITES, A. W., ROSEN, D. A. S., MONESTIEZ, P. & GUINET, C. 2010. Prey capture attempts cam be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biology,* 33**,** 713-719.

VOLPOV, B. L., ROSEN, D. A. S., TRITES, A. W. & ARNOULD, J. P. Y. 2015. Validating the relationship between 3-dimensional body accleration and oxygen consumption in trained Steller sea lions. *Journal of Comparitive Physiology B,* 185**,** 695-708.

WARABI, T., KATO, M., KIRIYAMA, K., YOSHIDA, T. & KOBAYASHI, N. 2005. Treadmill walking and overground walking of human subjects compared by recording sole-floor reaction force. *Neuroscience Research,* 53**,** 343-348.

WARD, S., MÖLLER, U., RAYNER, J. M. V., JACKSON, D. M., BILO, D., NACHTIGALL, W. & SPEAKMAN, J. R. 2001. Metabolic power, mechanical power and efficiency during wind tunnel flight by the European starling *Sturnus vulgaris*. *Journal of Experimental Biology,* 204**,** 3311-3322.

WATERS, R. L. & MULROY, S. 1999. The energy expenditure of normal and pathologic gait. *Gait and Posture,* 9**,** 207-231.

WATT, J. R., FRANZ, J. R., JACKSON, K., DICHARRY, J., RILEY, P. O. & KERRIGAN, D. C. 2010. A three-dimensional kinematic and kinetic comparison of overground and treadmill walking in healthy elderly subjects. *Clin Biomech (Bristol, Avon),* 25.

WEARING, S. C., REED, L. F. & URRY, S. R. 2013. Agreement between temporal and spatial gait parameters from an instrumented walkway and treadmill system at matched walking speed. *Gait Posture,* 38.

WEE, V. M., VON HEIMBURG, E. & VAN DEN TILLAAR, R. 2016. Comparison of perceptual and physiological variables of running on a track, motorized treadmill, and non-motorized curved treadmill at increasing velocity. *Acta Kinesiologiae Universitatis Tartuensis,* 22**,** 20-35.

WHIPP, B. & WASSERMAN, W. 1972. Oxygen uptake kinetics for various intensities of constant-load work. *Journal of Applied Physiology,* 33**,** 351-356.

WHITE, R. G. & YOUSEF, M. K. 1978. Energy expenditure in reindeer walking on roads and on tundra. *Canadian Journal of Zoology,* 56**,** 215-223.

WICKLER, S. J., HOYT, D. F., COGGER, E. A. & HIRSCHBEIN, M. H. 2000. Preferred speed and cost of transport: the effect of incline. *Journal of Experimental Biology,* 203**,** 2195.

WIENS, J. A., STENSETH, N. C., VAN HORNE, B. & IMS, R. A. 1993. Ecological mechanisms and landscape ecology. *Oikos,* 66**,** 369-380.

WILLIAMS, T. M., WOLFE, L., DAVIS, T., KENDALL, T., RICHTER, B., WANG, Y., BRYCE, C., ELKAIM, G. H. & WILMERS, C. C. 2014. Instantaneous energetics of puma kills reveal advantage of felid sneak attacks. *Science,* 346**,** 81-85.

WILSON, A. M., LOWE, J. C., ROSKILLY, K., HUDSON, P. E., GOLABEK, K. A. & MCNUTT, J. W. 2013a. Locomotion dynamics of hunting in wild cheetahs. *Nature,* 498**,** 185-189.

WILSON, R. P., GRIFFITHS, I. W., LEGG, P. A., FRISWELL, M. I., BIDDER, O. R., HALSEY, L. G., LAMBERTUCCI, S. A. & SHEPARD, E. L. C. 2013b. Turn costs change the value of animal search paths. *Ecology Letters,* 16**,** 1145-1150.

WILSON, R. P., WHITE, C. R., QUINTANA, F., HALSEY, L. G., LIEBSCH, N., MARTIN, G. R. & BUTLER, P. J. 2006. Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *Journal of Animal Ecology,* 75**,** 1081-1090.

WOLEDGE, R. C., BIRTLES, D. B. & NEWHAM, D. J. 2005. The variable component of lateral body sway during walking in young and older humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences,* 60**,** 1463-1468.

WYNDHAM, C. H., VAN DER WALT, W. H., VAN RENSBURG, A. J., ROGERS, G. G. & STRYDOM, N. B. 1971. The influence of body weight on energy expenditure during walking on a road and on a treadmill. *Internationale Zeitschrift fur Angewandte Physiologie,* 29**,** 285-292.

YNGVE, A., NILSSON, A., SJÖNSTRÖM, M. & EKELUND, U. 2003. Effect of monitor placement and of activity setting on the MTI accelerometer output. *Med Sci Sports Exerc,* 35.