INTRODUCTION

Models that predict human total energy expenditure (TEE, kcal day$^{-1}$) are used to develop energy and nutrition standards as well as estimate TEE among industrialized and nonindustrialized populations (Aiello and Wheeler, 2003, Dufour and Piperata, 2008; FAO/WHO/UNU, 2001; Katzmarzyk et al., 1996; Leonard et al., 1995, 1997; Spurr et al., 1996). They have also been applied to produce energy expenditure estimates for past populations (Froehle and Churchill, 2009; Leonard and Robertson, 1997, Steudel-Numbers and Tilkens, 2004). The currently recommended, and most frequently used, model for predicting TEE without physiological measurements is the Factorial Method (FAO/WHO/UNU, 1985, 2001). However, the Factorial Method consistently underestimates TEE (Durnin, 1990; Haggarty et al., 1994, Leonard et al., 1995, 1997; Roberts et al., 1991; Spurr et al., 1996).

The Factorial Method estimates TEE by summing the energetic cost of basal metabolic rate (BMR) and activity throughout the day. Each activity cost is estimated as a multiple of BMR based on activity intensity (FAO/WHO/UNU, 1985, 2001). Comparisons with TEE measurements using the doubly labeled water (DLW) and Flex-Heart Rate (Flex-HR) methods have found that the Factorial Method underestimates TEE by 16–22% (Leonard et al., 1995, 1997; Roberts et al., 1991). These underestimations can be as great as 30% among highly active populations. It has been suggested that discrepancies in BMR and physical activity cost estimations are the root of this underestimation (Leonard et al., 1997). Furthermore, the Factorial Method does not include cost estimates for thermoregulation nor the thermic effect of food (TEF), both of which can comprise a significant proportion of TEE. Thermoregulatory demands are known to increase BMR among indigenous cold populations by as much as 20% (Leonard et al., 2005; Snodgrass et al., 2005, 2006, 2008). TEF compromises roughly 10% of the overall TEE budget of a person in energy balance (Kinabo and Durnin, 1990).

Although these are well-established concerns, little effort has gone into producing a new model that better represents TEE and its multiple interacting components. The primary goal of the work presented here is to produce an accurate model for predicting human TEE across a range of climates and physical activity levels (PALS). The new model presented here, the Allocation and Interaction Model (AIM), improves upon current methods by including interacting cost terms for BMR, physical activity, thermoregulation, and the TEF. The general form of this model is:

$$\text{TEE} = \text{BMR} + E_{\text{activity}} + E_{\text{therm}} + \text{TEF}$$

where BMR is basal metabolic rate, $E_{\text{activity}}$ is the metabolic cost of physical activity, $E_{\text{therm}}$ is the metabolic cost of thermoregulation, and TEF is the thermic effect of food. Although AIM is a factorial type of model, it allows for interactions among its components, which have been shown to significantly impact TEE in extreme conditions such as cold climates (Steegmann, 2007).

Here, I present AIM. I then test it using a population of highly active adults living outdoors in temperate, hot, and cold climates taking part in National Outdoor Leadership School semester-long courses. Data from a month-long pilot study ($N = 6$) were also included.
School (NOLS) courses. I used the Flex-HR and DLW methods to measure TEE. First, I compare TEE of the NOLS population to Western and indigenous populations to ensure that they are more physically active, to better test the accuracy of AIM at high levels of TEE where the Factorial Method traditionally fails. Second, as the Flex-HR method has not been validated at high levels of TEE, I compare Flex-HR measurements to DLW measurements. Finally, I compare AIM and the Factorial Method TEE predictions to DLW and Flex-HR method TEE measurements. I demonstrate that the Flex-HR method needs further evaluation at high levels of activity, and that AIM is an accurate new tool for predicting human TEE.

**SUBJECTS AND METHODS**

**Subjects**

Participants included 59 healthy volunteers (40 males, 19 females, aged 18–44 years) from four 12- to 16-week courses and one 5-week outdoor education course in the western United States in 2010 and 2011. The Institutional Review Board of Washington University, St. Louis, (IRB protocol 201104106) approved this study and subjects gave informed consent prior to participating.

NOLS operated these courses and provided logistical support for field data collection. Two of the courses (N = 25) were in the spring and summer, course names NS1 and NS2, that lasted 12 weeks, and two (N = 28) were in the fall and winter that lasted 16 weeks, course names FS5 and FS8. An additional six subjects took part in a shorter (5 weeks) pilot study that I conducted during the summer of 2010. This resulted in five separate courses taking place in three different conditions. The NS1 and NS2 course participants experienced temperate (mean 14.6°C) and hot (mean 23.4°C) climates, FS5 and FS8 participants experienced temperate (mean 14.0°C) and cold (mean −7.2°C) climates, and the Pilot study took place solely in a temperate climate (mean 12.8°C).

**Data collection**

All subjects (N = 59) took part in two types of data collection. The first consisted of BMR, heart rate calibration, anthropometric, and bioelectrical impedance measurements; this is referred to as Calibration. I collected these data three times throughout the semester long course: before the course began (Calibration 1), in between temperate and extreme climate (hot or cold) regimes (Calibration 2), and at the end of the course (Calibration 3). The second type of data collection consisted of in-field heart rate, DLW, food diary, activity diary, and daily temperature data collection. I collected these data twice during each semester course, once during the temperate regime and once during the extreme, either hot or cold, regime. This data collection is referred to as the Energy and Activity Assessment (EAA) (Table 1).

**Data collection settings**

The 12- to 16-week courses, Pilot study excluded, followed a similar schedule. Subjects arrived at the NOLS headquarters in Lander, WY, and spent several days there to meet their course mates and instructors as well as prepare gear and rations for the first of the three sections of their semester. NS1 and NS2 (N = 25) arrived in Lander in late May 2011. During this time, I performed Calibration 1 with NS1 and NS2. They embarked on the temperate climate portion of their course in early June in the Absaroka Mountain Range, WY, which consisted of hiking. After subjects had been in the field for 2 weeks, I met them to conduct the EAA for 11 days. After another 2 weeks, the NOLS students finished the hiking portion of their semester and returned to Lander, WY to change and refurbish their gear and rations. During this time, I conducted Calibration 2.

The second portion of the semester consisted of rock climbing at the City of Rocks, ID (NS1) and Devils Tower, WY (NS2) in late July 2011, this comprised the hot climate portion of the study. Similarly, I conducted the EAA for 6 days after subjects had been at these new locations for 2 weeks. Once the remainder of the hot climate course was complete, subjects returned to Lander, WY for one final gear change and ration replenishment, during which I conducted Calibration 3. The third portion of the semester involved kayaking and river rafting based out of Vernal, UT. I did not perform the EAA during this portion given the extremely high risk of nonwater proof equipment being fully submerged in river water.

FS5 and FS8 (N = 28) followed a similar schedule, however, their courses consisted of a 2-month long temperate climate hiking section and ended with a cold climate cross-country skiing section. Both courses arrived in Lander, WY in early September 2011 during which time they prepared for their course and I performed Calibration 1. They embarked on their temperate climate, hiking section in the Wind River Mountain Range, WY. After 2 weeks in the field, I met each course and performed the EAA for 7 days with FS5 and 8 days with FS8. Once they completed this section of their course, they returned to Lander, WY in mid-November 2011, and I performed Calibration 2. The NOLS students then began the cold climate, cross-country skiing portion of their course in the Absaroka Mountain Range, WY. Once the subjects had

---

Table 1. A summary of the total number of subjects, the number of subjects taking part in the different measurements

<table>
<thead>
<tr>
<th>Course</th>
<th>N</th>
<th>Course duration</th>
<th>Climate</th>
<th>Mean temp. (°C)</th>
<th>Flex-HR participants</th>
<th>DLW participants</th>
<th>EAA location</th>
<th>Duration of EAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS1*</td>
<td>14</td>
<td>6/2/11–8/10/11</td>
<td>Temperate</td>
<td>15.6</td>
<td>14</td>
<td>1</td>
<td>Absaroka Mountain Range, WY</td>
<td>11</td>
</tr>
<tr>
<td>NS2*</td>
<td>11</td>
<td>6/4/11–8/12/11</td>
<td>Temperate</td>
<td>13.5</td>
<td>11</td>
<td>1</td>
<td>Absaroka Mountain Range, WY</td>
<td>11</td>
</tr>
<tr>
<td>FS5</td>
<td>14</td>
<td>9/4/11–12/3/11</td>
<td>Temperate</td>
<td>13.8</td>
<td>14</td>
<td>1</td>
<td>Wind River Range, WY</td>
<td>7</td>
</tr>
<tr>
<td>FS8*</td>
<td>14</td>
<td>9/8/11–12/10/11</td>
<td>Temperate</td>
<td>14.2</td>
<td>14</td>
<td>1</td>
<td>Wind River Range, WY</td>
<td>8</td>
</tr>
<tr>
<td>Pilot</td>
<td>6</td>
<td>7/1/10–8/4/10</td>
<td>Temperate</td>
<td>12.8</td>
<td>6</td>
<td>3</td>
<td>Wind River Range, WY</td>
<td>6</td>
</tr>
</tbody>
</table>

*Indicates courses for which the same subject participated in the DLW measurements for both the temperate and extreme climate.
been in the field for 2 weeks, I met them to perform the EAA. The students finished their course, and I performed Calibration 3 in early- to mid-December 2011 in Lander, WY.

The Pilot study (N = 6) consisted of a five weeklong course, which took place entirely in the Wind River Mountain Range, WY in July and August of 2010. This course consisted of hiking and moderate mountaineering in a temperate climate. Subjects from the Pilot study arrived in Lander, WY in late June for several days of course preparation during which time I performed Calibration 1. Once subjects were in the Wind River Mountain Range, WY for 2 weeks, I met them to perform the EAA for 6 days. Subjects finished the rest of their course and returned to Lander, WY where I performed Calibration 2 before subjects left for their respective homes (Table 1).

Metabolic measurements

Basal metabolic rate. BMRs were collected from each subject using a portable respirometry unit (Cosmed K4b2, Chicago, IL) following standard practice (Gayda et al., 2010). This system measures oxygen consumption and carbon dioxide production using a breath-by-breath analysis. I took BMR measurements in the morning before subjects had their first meal. Subjects were in a supine position on foam pads placed on the floor, in a temperature controlled room, and rested 15–20 min before I took measurements. I took measurements for 6–8 min with the last 4 min of the measurement averaged to determine BMR.

Flex-HR measurements. I calculated TEE for each subject from heart-rate data using the Flex-HR method (Leonard, 2003). I measured heart rate using a chest-strap monitor worn continuously for 2 weeks; data were logged using an ActiTrainer device (Actigraph, Pensacola, FL) worn on the hip. To convert heart rate to energy expenditure, I collected calibration measurements for each subject following their BMR measurement. I asked subjects to stand, walk (1, 1.5, and 2 m s⁻¹), and run (2, 2.5, and 3 m s⁻¹) for 5 min at each speed on a treadmill while I recorded heart rate (bpm) and respirometry (kcal min⁻¹) data simultaneously. I determined the Flex-HR flex-point for each subject as the mean of the highest heart rate at rest and the lowest heart rate during exercise following Leonard (2003). I then determined the relationship between heart rate and energy expenditure as the least-squares regression line for heart rate and energy expenditure. Expenditure during sleeping hours was calculated using their BMR value. I filled in missing heart rate values using averaged values (beats/min) calculated from the available data for each day to calculate 24 h Flex-HR TEE predictions. The Flex-HR calibration that took place closest in time to in-field heart rate measurements was used to calculate TEE to take into account changes in body composition and cardiovascular fitness.

Doubly labeled water method. For N = 8 subjects, I measured TEE using the DLW method for 6–11 days (Table 1). I measured three subjects twice, once in the temperate climate and once in the extreme hot or cold climate. I measured two subjects from the same course, FS5, only once, one in the temperate climate and one in the cold climate. The subject who participated in the temperate climate DLW measurement opted to not participate during the cold climate DLW measurement. Three subjects from the pilot study took part in one set of DLW measurements. This resulted in 8 total subjects, but 11 DLW measurements.

I gave subjects an oral dose of DLW (116.08–122.62 g; 10% H²¹⁸O, 6% ²H₂O). I rinsed dose bottles with bottled water twice which was also consumed by subjects to ensure the full dose was administered. I collected urine samples prior to the DLW dose, 6–8 h after the dose, and then every other day for the duration of the EAA. I collected urine in clean, dry wax coated paper cups. I filled four 2-ml cryovials (Sarstedt) at each urine sample collection and placed vials in waterproof plastic bags, kept cold in a small soft-pack cooler using pack snow, ice, or mountain river water.

The three DLW samples from the 2010 Pilot study were analyzed with gas-isotope mass spectrometry at the Baylor College of Medicine, under the direction of Dr. William Wong. I analyzed the DLW samples from the five subjects from the full 2011 study using the Cavity Ring-Down Spectrometry system (Picarro, Sunnyvale, CA) at Hunter College in New York. Standard equations for determining CO₂ production and TEE were used and are described elsewhere (International Atomic Energy Agency, 2009). Analyzing samples in two laboratories presents the possibility of variation between the results from the two laboratories. An interlaboratory variation protocol was not performed.

Activity, food, and clothing diaries

I asked subjects to keep activity and food diaries for the duration of the EAA; the duration of this assessment for each course can be found in Table 1. Subjects reported activity type, distance or duration of activity, and backpack weight. Activity logs were compared to course instructor official travel logs, official course maps, and NOLS curricula travel schedules to ensure accuracy. The typical NOLS course schedule consisted of a mix of strenuous activity days and rest days filled with wilderness education curriculum. Activities consisted of hiking and mountaineering in temperate climates; rock climbing and hiking in hot climates; and cross country skiing and snow shoveling in cold climates. Subjects also reported type and quantity of food eaten. Data from the food logs were transcribed into a Microsoft Excel spreadsheet using the NOLS Cookery (Pearson, 2004), NOLS Backcountry Cooking (Pearson and Kuntz, 2008), and NOLS Backcountry Nutrition (Howley Ryan, 2008) to breakdown typical backcountry recipes. The official USDA National Nutrient Database for Standard Reference was used to assign nutritional values and calories to the foods consumed (U.S. Department of Agriculture, Agricultural Research Service, 2012). Calories were summed for each day. I provided collapsible measuring cups to aid measuring accuracy. Subjects also documented the clothing they took with them including the brand and garment name.

Temperature data

I measured temperature using an Extech RHT10 Humidity and Temperature USB Data-logger (Extech

American Journal of Human Biology
Industries, Nashua, NH) carried by the course instructors. This device measured and recorded temperature and humidity on a minute-by-minute basis. I downloaded temperature data using the Extech software (Extech Industries, Nashua, NH). I calculated high, low, and mean temperatures for each day as well as averaged across the EAA.

**Predictive models for daily energy expenditure**

**Factorial model.** I predicted TEE for each subject in each climate using the Factorial Method (FAO/WHO/UNU, 1985). The general form of the Factorial Method is:

$$TTE = \text{BMR} + \text{Activity}$$

BMR was calculated using existing equations that incorporate age, sex, and body mass (Henry, 2005).

**For males:** 16.0M + 545

**For females:** 13.1M + 558

Where $M$ is mass (kg). These equations were chosen for their generalizability, which is ideal for application of this model to multiple populations. I calculated activity costs as PAL values (i.e., a multiple of BMR) based on the intensity of the activity, using a standardized list of activity-specific values (FAO/WHO/UNU, 2001). These values can be found in Table 2. Subjects’ activity logs were used to determine type and duration of activities.

The Allocation and Interaction model. I predicted TEE for each subject in each climate using AIM. This model takes the general form of:

$$TTE = \text{BMR} + E_{\text{activity}} + E_{\text{therm}} + \text{TEF}$$

I calculated BMR following equations from Henry (2005) listed above. $E_{\text{activity}}$ was determined by activity-specific cost equations (Table 2). I calculated $E_{\text{therm}}$ following the COMFA outdoor thermal comfort model (Kenny et al., 2009). This model is based on first principles of metabolic heat production, convection, radiation, and evaporation. Derivation of this equation and its details are described elsewhere (Kenny et al., 2009). I was able to measure all the necessary variables for using the COMFA outdoor thermal comfort model. The general form of this model is:

$$E_{\text{therm}} = M + R_{RT} - C - E - L$$

Where $M$ is the metabolic heat generated by a person calculated using BMR and the metabolic cost of activity. $R_{RT}$ is radiation absorbed by a person calculated following Kenny et al. (2008) using body temperature and ambient temperature. $C$ is the convective heat loss calculated using body temperature, ambient temperature, and clothing resistance. $E$ is the evaporative heat loss calculated using body temperature, ambient temperature, exposed skin area, clothing resistance, atmospheric pressure, and known constants for skin tissue resistance to vapor transfer (Kenny et al., 2009). $L$ is the long-wave radiation heat loss calculated using known constants for the emissivity of human skin and clothing as well as body temperature, exposed skin area, and ambient temperature. A constant body temperature of 37°C was used, and I measured ambient temperature using the Extech RHT10 Humidity and Temperature USB Data-logger. In temperate climates, I used an estimate of 25% exposed skin surface area, 10% for cold climates, and 60% for hot climates following International Standards Organization (2007) guidelines. I estimated TEF, the metabolic cost incurred from digesting food, to be 10% of caloric intake (Kinabo and Durnin, 1990).

**Anthropometrics and body composition**

I collected several external anatomical measurements including height, lower limb length, and bi-iliac breadth following standard procedures (Lohman et al., 1988). I collected these measurements using a standard cloth measuring tape and large calipers. I collected data on body mass, percent body fat, and muscle mass using a bioelectrical impedance scale, Tanita BC-558 Ironman Segmental Body Composition Monitor, the Tanita equations are unpublished (Tanita Corporation, Arlington Heights, IL).

**Statistical analysis**

I generated plots using Microsoft® Excel® for Mac 2010 and RStudio, ©RStudio, Inc. 2009-2012. I performed all statistical analyses including linear regressions, multiple regressions, and Tukey’s pairwise comparisons using IBM® SPSS® Version 21. I considered results significant at $P < 0.05$. I used multiple regressions controlling for age, sex, fat mass free, and height followed by a Tukey’s pair-wise comparisons to compare the NOLS population to

### Table 2. Activity specific equations for determining the total metabolic cost of activity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Equation</th>
<th>Unit</th>
<th>Physical activity level$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking$^c$</td>
<td>$4.1M^{-0.449}$</td>
<td>kcal km$^{-1}$ kg$^{-1}$</td>
<td>3.0</td>
</tr>
<tr>
<td>Running$^c$</td>
<td>$30.55 + 1.595(M) - 0.9(L_L)$</td>
<td>kcal min$^{-1}$</td>
<td>4.4</td>
</tr>
<tr>
<td>Climbing$^c$</td>
<td>$0.1352M^2 + 1.7853$</td>
<td>kcal min$^{-1}$</td>
<td>5.1</td>
</tr>
<tr>
<td>Hiking$^c$</td>
<td>$0.14[1.5M + (2.0M + B x BM)^{-1}]^2 + 0.1(M + B[1.5c^2 + 0.35cg])$</td>
<td>kcal min$^{-1}$</td>
<td>5.1</td>
</tr>
<tr>
<td>Cross-country skiing$^c$</td>
<td>$0.274Mt$</td>
<td>kcal min$^{-1}$</td>
<td>5.1</td>
</tr>
<tr>
<td>Downhill skiing$^c$</td>
<td>$0.162Mt$</td>
<td>kcal min$^{-1}$</td>
<td>3.8</td>
</tr>
<tr>
<td>Digging snow$^d$</td>
<td>$6.0Mt$</td>
<td>kcal min$^{-1}$</td>
<td>5.1</td>
</tr>
<tr>
<td>Light swimming$^e$</td>
<td>$0.1Mt$</td>
<td>kcal min$^{-1}$</td>
<td>3.8</td>
</tr>
<tr>
<td>Push-ups and sit-ups$^f$</td>
<td>$0.09Mt$</td>
<td>kcal min$^{-1}$</td>
<td>2.8</td>
</tr>
<tr>
<td>Yoga$^f$</td>
<td>$0.1Mt$</td>
<td>kcal min$^{-1}$</td>
<td>2.8</td>
</tr>
</tbody>
</table>

$^a$ Values determined by activity-specific cost equations (Table 2). $^c$ Ainsworth et al. (2000); $^d$ Booth et al. (1999); $^e$ Capelli et al. (1998); $^f$ Henry, 2005.

$M$ = body mass (kg), $L_L$ = lower limb length, $B$ = backpack weight (kg), $g$ = percent grade of terrain, and $t$ = time (h). Climbing speed estimated at 3.2 m min$^{-1}$ (Booth et al. 1999). Percent grade of the terrain was determined using distance and elevation traveled documented in the activity logs. Hiking speed (m s$^{-1}$) was determined following Pandolf et al. (1977), and $B$ is the backpack weight (kg). Sources are as follows: $^a$FAO/WHO/UNU (2001); $^b$Rubenson et al. (2005); $^c$Staedel-Numbers and Tilkens, (2004); $^d$Booth et al. (1999); $^e$Pandolf et al. (1977); $^f$McArdis et al. (2001); $^c$Audet (1994); $^d$Anisworth et al. (2000); $^e$Capelli et al. (1998).
TABLE 3. TEE (kcal day$^{-1}$) measurements and predictions

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Mass (kg)</th>
<th>Climate</th>
<th>DLW</th>
<th>Flex-HR</th>
<th>AIM</th>
<th>Factorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS1-12 M</td>
<td>M</td>
<td>89.9</td>
<td>Temperate</td>
<td>4,264</td>
<td>5,427</td>
<td>3,280</td>
<td>3,156</td>
</tr>
<tr>
<td>NS2-1 F</td>
<td>F</td>
<td>64.5</td>
<td>Temperate</td>
<td>2,837</td>
<td>2,814</td>
<td>3,217</td>
<td>2,591</td>
</tr>
<tr>
<td>FS5-12 F</td>
<td>F</td>
<td>65.8</td>
<td>Temperate</td>
<td>2,593</td>
<td>3,949</td>
<td>2,595</td>
<td>2,196</td>
</tr>
<tr>
<td>FS8-10 M</td>
<td>M</td>
<td>72.7</td>
<td>Temperate</td>
<td>3,597</td>
<td>3,138</td>
<td>3,158</td>
<td>2,839</td>
</tr>
<tr>
<td>Pilot 1 F</td>
<td>F</td>
<td>68.7</td>
<td>Temperate</td>
<td>3,340</td>
<td>3,729</td>
<td>3,675</td>
<td>2,986</td>
</tr>
<tr>
<td>Pilot 3 M</td>
<td>M</td>
<td>70.0</td>
<td>Temperate</td>
<td>3,641</td>
<td>3,651</td>
<td>3,537</td>
<td>2,644</td>
</tr>
<tr>
<td>Pilot 4 M</td>
<td>M</td>
<td>69.7</td>
<td>Temperate</td>
<td>3,790</td>
<td>5,668</td>
<td>3,629</td>
<td>3,093</td>
</tr>
<tr>
<td>NS1-12 M</td>
<td>M</td>
<td>95.0</td>
<td>Hot</td>
<td>4,313</td>
<td>4,889</td>
<td>4,276</td>
<td>2,839</td>
</tr>
<tr>
<td>NS2-1 F</td>
<td>F</td>
<td>65.5</td>
<td>Cold</td>
<td>2,837</td>
<td>3,675</td>
<td>3,093</td>
<td>2,286</td>
</tr>
<tr>
<td>FS5-12 F</td>
<td>F</td>
<td>73.8</td>
<td>Cold</td>
<td>4,517</td>
<td>9,155</td>
<td>5,090</td>
<td>3,031</td>
</tr>
<tr>
<td>FS8-10 M</td>
<td>M</td>
<td>71.9</td>
<td>Cold</td>
<td>4,137</td>
<td>4,678</td>
<td>5,687</td>
<td>3,261</td>
</tr>
</tbody>
</table>

Fig. 1. The relationship between DLW measured TEE and Flex-HR measured TEE for the NOLS subjects (N = 11) who took part in DLW measurements.

TEE values for indigenous and nonindigenous populations collected from the literature. I used Bonferroni adjusted Student’s t-tests to compare TEE measured through different methods (DLW, Flex-HR, the Factorial Method, and AIM) within subjects. I performed Bland-Altman analyses using Microsoft© Excel© for Mac 2010 to determine bias within the Factorial Method and AIM.

RESULTS

TEE measurements

TEE, as measured by the DLW method (N = 11 measurements from eight subjects), ranged from 2,593 to 4,517 kcal day$^{-1}$ and had a mean of 3,624 ± 660 kcal day$^{-1}$. TEE, as measured by the Flex-HR Method (N = 59), ranged from 2,150 to 9,730 kcal day$^{-1}$ and had a mean of 4,586 ± 1,499 kcal day$^{-1}$. A summary of the different TEE measurement results can be found in Table 3. Comparison with other populations confirmed that the subjects in this study had relatively high TEE. A linear regression controlling for age, sex, fat free mass, and height followed by a Tukey’s pair-wise comparison (F = 11.036, P < 0.001) revealed that the NOLS sample had a significantly higher TEE than traditional Hadza hunter-gatherers (N = 30) (Pontzer et al., 2012), subsistence-agricultural Bolivians (N = 24) (Kashiwazaki et al., 2009), U.S. and European populations (N = 51) (Davidson et al., 1997; Prentice et al., 1986; Schulz et al., 1989; Seale et al., 1990; Welle et al., 1992), rural Yakut Siberians (N = 27) (Snodgrass et al., 2006), and urban Guatemalans (N = 14) (Stein et al., 1988) (P < 0.001 for all cases).

Flex-HR method vs. DLW method

Flex-HR TEE measurements tended to be lower than DLW measurements at lower levels of TEE, but this difference did not meet the criterion for significance after Bonferroni correction (Bonferroni adjusted α = 0.008, P = 0.026 paired t-test). However, at high levels of energy expenditure, the Flex-HR method tended to overestimate TEE. Figure 1 shows the relationship between DLW measured TEE and Flex-HR TEE. The percent difference between Flex-HR TEE and DLW TEE ranged from −15.6 to 102.7% with a mean of 24 ± 34.1% (Table 3). Subject FS5-1 had an exceptionally large Flex-HR measurement compared to both DLW and Flex-HR measurements. After this correction, Flex-HR TEE ranged from 2,150 to 8,076 kcal day$^{-1}$ and had a mean of 3,867 ± 1,176 kcal day$^{-1}$.

Modeled TEE compared to measured TEE

AIM produced daily TEEs with a range of 1,947–7,080 kcal day$^{-1}$ and a mean of 3,548 ± 1,090 kcal day$^{-1}$. The Factorial Method produced daily TEEs with a range of 1,894–4,156 kcal day$^{-1}$ and a mean of 2,775 ± 423 kcal day$^{-1}$. Figure 3 shows TEE as measured by the DLW method and Flex-HR method as well as TEE predicted by AIM and Factorial Method. A full summary of AIM and Factorial Method calculated mean daily TEEs for each course is found in Table 3.

The Bland–Altman method was applied to the data to determine if there was any bias in the Factorial Method and AIM (Fig. 4). The Factorial Method tended to underestimate TEE at greater levels of energy expenditure in comparison to both DLW and Flex-HR measurements. AIM did not present this bias. AIM had the tendency to produce worse predictions at higher levels of TEE;
however, the inaccuracy did not bias toward overestimation or underestimation.

**AIM compared to the factorial method**

A linear regression drawn through the origin of AIM with DLW measured TEEs produced a slope of 0.97 (95% CI: 0.84–1.07), $r^2 = 0.48$. A linear regression drawn through the origin of the Factorial Method with the DLW TEE values for daily TEE produced a slope of 1.31 (95% CI: 1.24–1.42), $r^2 = 0.70$. The slope from the Factorial Method was significantly different from a slope of one, but the slope from AIM was not (Fig. 5), and the slopes from the two models differed (AIM: $F = 328.98, P < 0.001$; Factorial Method: $F = 1,126.688, P < 0.001$). Forcing these slopes through the origin gave similar results (Fig. 6). AIM produced a slope of 1.04 (95% CI: 0.99–1.1) $r^2 = 0.37$, and the Factorial Method produced a slope of 1.40 (95% CI: 1.33–1.47) $r^2 = 0.35$ (AIM: $F = 1,364.5, P < 0.001$; Factorial Method $F = 1,626.834, P < 0.001$).

For the entire dataset, AIM overestimated TEE by 4.1%, a smaller absolute difference ($P < 0.001$, paired t-test) than the 25.3% underestimation produced by the Factorial Method. At TEEs > 3,000 kcal day$^{-1}$, the Factorial Method underestimated TEE by 31.6%, which was significantly higher than the AIM underestimation of TEE by only 10.7% ($P < 0.001$, paired samples t-test).

As a final test of the models’ effectiveness, a within-subjects analysis was performed for a subsample of 12 subjects, three from each semester course, with high-quality Flex-HR calibrations and in-field data collection. The predictions and measurements were compared on a day-to-day basis for the temperate climate. AIM produced a mean slope across subjects of 1.15 ± 0.27 with a range of 0.73–1.62 and a mean $r^2 = 0.36 ± 0.24$ with a range of 0.01–0.67. When pooled, the confidence intervals were 1.0–1.2, $z = 0.05$ ($F = 936.3, P < 0.001$) and not significantly different from a slope of one. The Factorial Method produced a mean slope of 1.33 ± 0.29 with a range of 0.87–1.8 and a mean $r^2 = 0.29 ± 0.19$ with a range of 0.02–0.71. When pooled, the Factorial Method confidence intervals were 1.1–1.3, $z = 0.05$ ($F = 762.8, P < 0.001$), and significantly different from a slope of one.

**DISCUSSION**

This study presented a new model, AIM, for predicting human TEE. AIM includes specific terms for BMR, thermoregulation, activity, and TEF, and allows for interactions...
among these variables. AIM and Factorial Method TEE estimates were compared to DLW and Flex-HR TEE measurements among healthy, highly active participants of NOLS courses. The Flex-HR method overestimates TEE at high levels of energy expenditure, and AIM produces more accurate TEE estimates than the Factorial Method.

Limitations to this study

First, the Cosmed k4b2 has been known to overestimate BMR (Duffield et al., 2004). This would impact the BMR measurements along with heart rate calibrations. Second, this study does not take into account psychological stress, which can increase heart rate and metabolic rate. This could account for some of the unusually high Flex-HR measurements observed. Third, two laboratories were used to analyze the DLW results from this study. No protocol was used to determine if there was significant variation between the two laboratories, introducing the possibility of DLW measured TEE error. Fourth, the NOLS subjects in this study are not representative of all populations. AIM needs to be validated among a variety of populations before it can broadly applied.

Flex-HR discrepancies

In this study, the Flex-HR method produced TEE estimates of greater than 9,000 kcal day$^{-1}$. This measurement is substantially higher than the highest DLW measured human TEE of roughly 7,000 kcal day$^{-1}$ among Tour de France cyclists (Westerterp et al., 1986). However, this is not uncommon. Flex-HR discrepancies have been reported to range from $-22.2$ to $52.1\%$ of DLW measurements at the individual level (Livingstone et al., 1990, Leonard, 2003), and $10\%$ at the group level (Leonard, 2003).

There are a number of reasons for the divergence between DLW and Flex-HR measurements among the NOLS sample. The ActiTrainer devices used to collect HR data were used for extended periods of time without recharging, used for eight different $6-11$ days of data collection over 7 months, exposed to the elements in the backcountry, and exposed to possible interference from satellite phones and avalanche beacons. As there is currently no research of ActiTrainer data degradation over repeated use and abuse or interference from other devices, it is difficult to confirm that any of the above reasons are possible causes for the large difference between the DLW and Flex-HR results.

Recent work has also suggested that climatic extremes can impact heart rate, disrupting the traditionally held relationship between heart rate and metabolic rate upon which the Flex-HR method depends. The Frank–Starling law of the heart relates heart stroke volume to the volume of blood filling the heart, such that a change in blood pressure accompanies a change in heart rate (Wilson et al.,

![Fig. 5. Linear regression of the AIM and the Factorial Method TEE data against the observed DLW TEE data for NOLS subjects (N = 11, measurements for eight subjects) who took part in DLW measurements.](image1)

![Fig. 6. Linear regression of the AIM and the Factorial Method TEE data against the corrected Flex-HR TEE data for all N = 59 NOLS subjects.](image2)
In cold climates, humans experience cutaneous arterial vasconstriction, which increases blood pressure thereby reducing heart rate. In hot conditions, humans experience vasodilation, which increases heart rate (Wilson et al., 2009). These changes in heart rate do not correspond to complementary changes in metabolic rate. All heart rate calibrations for this study were performed under thermoneutral conditions. This suggests that under cold conditions subjects would experience depressed heart rates, and thermoneutral calibrated TEE estimates would be lower than actual TEE. The converse is true in hot climates, TEE estimates would be greater than actual energy expenditure. This environmental impact on heart rate makes the use of the Flex-HR method in extreme temperatures difficult. Making models that do not rely on heart rate, such as AIM, preferable.

The AIM outperforms the factorial method

AIM was designed to produce more accurate estimates of human TEE across a range of climates and PALs. The NOLS population was used because of its high level of physical activity, which allowed for AIM to be tested where the Factorial Method fails, at high levels of TEE (Leonard et al., 1997). The results show that AIM is more accurate at both low and high levels of physical activity, making it a superior method for predicting TEE. This was achieved by including more specific metabolic cost terms and allowing for interaction among them, which has been shown to be an important factor particularly in cold climates (Steegmann, 2007).

AIM performs particularly well at high TEEs. This is likely due to the ability of AIM to produce TEE estimates greater than 4,000 kcal day⁻¹ (Fig. 6). The Factorial Method is unable to account for possible internal tradeoffs when energy expenditures are high. AIM appears to avoid this issue. However, both models have low r²-values. Given the high level of individual variation in metabolic rate, this is not wholly unexpected.

Once more broadly validated, AIM can be used to analyze energy expenditure within and between populations. Furthermore, as AIM is a more explicit model, it can be used to assess energy allocation differences among populations inhabiting different climates. A better understanding of how humans allocate energy to costly activities such as thermoregulation, physical activity, and reproduction can help us explore the subtle, and possibly adaptive, differences in life history strategies.

This analysis demonstrates that AIM is more accurate at predicting human TEE than the Factorial Method, and possibly even the Flex-HR method, across a range of PALs and climates. Furthermore, AIM succeeds where the Factorial Method has traditionally failed—at high levels of energy expenditure. The results presented here suggest that AIM should be used in place of the Factorial Method for estimating human TEE.

ACKNOWLEDGMENTS

I would like to thank Herman Pountzer, Erik Trinkaus, Hannah Wilson, Elizabeth Rowe, Michaela Howells, and Chris Lynn for their input throughout this project. I would also like to thank John Gookin, Ryan Hutchins-Cabibi, Sheila Baynes, Lauren Crenshaw, and Kate Herden from the National Outdoor Leadership School for their support and collaboration. Finally, I would like to thank the participants of this study. There are no conflicts of interest associated with this study.

LITERATURE CITED


American Journal of Human Biology