Comparative Responses to Squats Completed with Free Weights and an Exoskeleton

Peter Neuhaus[1], Chris Jumonville[1], Rachel A. Perry[2], Roman Edwards[2], Jake L. Martin[3], Ahlam Alarbi[3], William T. Potter[3], and John F. Caruso[4]

[1]*Institute for Human and Machine Cognition, Pensacola, FL;* [2]*Exercise and Sports Science Program, University of Tulsa, Tulsa, OK;* [3]*Department of Chemistry and Biochemistry, University of Tulsa, Tulsa, OK;* [4]*Department of Health and Sport Sciences, University of Louisville, Louisville, KY*

ABSTRACT

To assess the comparative similarity of squat data collected as they wore a robotic exoskeleton, female athletes (n=14) did two exercise bouts spaced 14 days apart. Data from their exoskeleton workout was compared to a session they did with free weights. Each squat workout entailed a four-set, four-repetition paradigm with 60-second rest periods. Sets for each workout involved progressively heavier (22.5, 34, 45.5, 57 kg) loads. The same physiological, perceptual, and exercise performance dependent variables were measured and collected from both workouts. Per dependent variable, Pearson correlation coefficients, t-tests, and Cohen's d effect size compared the degree of similarity between values obtained from the exoskeleton and free weight workouts. Results show peak O_2, heart rate, and peak force data produced the least variability. In contrast, far more inter-workout variability was noted for peak velocity, peak power, and electromyography (EMG) values. Overall, an insufficient amount of comparative similarity exists for data collected from both workouts. Due to the limited data similarity, the exoskeleton does not exhibit an acceptable degree of validity. Likely the cause for the limited similarity was due to the brief amount of familiarization subjects had to the exoskeleton prior to actual data collection. A familiarization session that accustomed subjects to squats done with the exoskeleton prior to actual data collection may have considerably improved the validity of data obtained from that device.

Key words: Robotics; Flight Hardware; Countermeasures; Exercise; Strength Loss; Muscle Atrophy; Physiology; Perceptual; Validity; Familiarization

Correspondence to: John F. Caruso
Department of Health and Sport Sciences
The University of Louisville
Louisville Kentucky 40292
Telephone: 502-852-6648
E-mail: john.caruso@louisville.edu

INTRODUCTION

Long-term manned spaceflight imparts significant physiological impairments that compromise astronaut health and safety during and after exposure to μ*g*. In-flight exercise countermeasures attempt to limit those impairments. Recent advances in exercise hardware for long-term manned spaceflights include the development of robotic exoskeletons. One such device, from the Institute for Human and Machine Cognition (IHMC), weighs less than 18 kg. The exoskeleton is lighter, smaller, and consumes less power than the Advanced Resistive Exercise Device (ARED) now aboard the International Space Station (ISS) (Carpinelli, 2014; Hargens et al., 2013). Thus the portability and design of the IHMC exoskeleton may be ideally suited to the in-flight exercise hardware requirements for small space vehicles, such as Orion, which is slated to be NASA's first spacecraft to transport humans to Mars (Carpinelli, 2014). The IHMC exoskeleton attaches to a user's torso like a backpack as they stand upright. Two exoskeleton actuators, each of which is oriented parallel to the user's knee joints as they stand upright, impart high-fidelity torque control. Per actuator, one of its ends connects to a frictionless joint at the exoskeleton's base plate, while the other attaches to another frictionless joint near the user's hip. Since the IHMC exoskeleton uses motors to provide resistance, it imposes loads independent of gravity and may serve as in-flight hardware to limit lower body muscle atrophy and strength losses (Bamman et al., 1998; Hargens et al., 2013). In operation, the actuators put a compressive load between a user's feet and torso, which they resist by exerting forces. During exercise repetitions on the exoskeleton, subjects first flex and then extend their knees to mimic a squatting (deep knee bend) motion. A computer controls the actuators in order to provide different levels of resistance. The images in Figure 1 depict squats done on the IHMC exoskeleton.

Recent versions of in-flight resistive exercise hardware, such as the ARED, are typically designed to enable performance of the squat. Prime movers for the squat are the lower body extensors, which are the muscles most prone to atrophy and strength loss from actual and simulated spaceflight (Bamman et al., 1998; Stauber, 1989). It is the lower body extensors that incur the greatest degree of pre-stretch as squat depth (descent) increases, which leads to more motor unit recruitment when the aforementioned muscles shorten during the ascent phase of each exercise repetition. Thus the squat may be particularly efficacious as an in-flight counter-measure to μ*g*-induced atrophy and strength

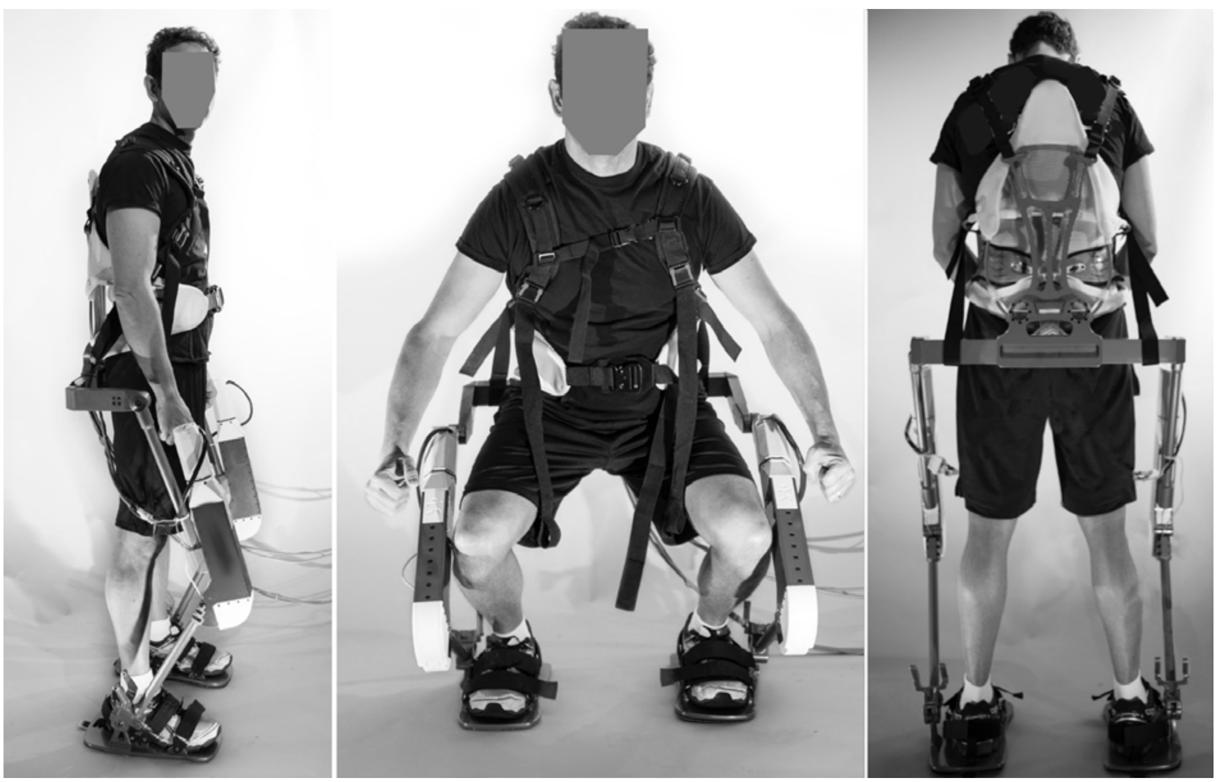

Figure 1. Side, front, and rearview images of the IHMC robotic exoskeleton. In a typical exercise configuration, it connects to the user's torso like a hiking backpack—over the shoulders and around the waist. In the linkages paralleling the user's legs are motorized, computer-controlled actuators. In operation, these actuators try to fold up, or collapse, putting a compressive load between the user's feet and torso. This force is similar to carrying a heavy backpack, or replicating a weightlifting squat exercise.

losses incurred by the lower body extensor musculature. The exoskeleton's intended operation is to function like free weight squats done with a traditional barbell, without the customary requirement of gravity to impart resistance. Exoskeleton features similar to those of free weights include its ability to impart concentric and eccentric loads to the knee extensors. This is a concern, since a lack of low level eccentric loading during manned spaceflight was deemed a stimulus for knee extensor atrophy and strength losses (Stauber, 1989). In-flight use of the exoskeleton may address this important concern, but first the degree to which its exercise data are like those for free weights must be established. Only then can the validity of the exoskeleton be assessed, and its future use as in-flight exercise hardware be determined.

For a new or novel device, validity is essential to establish before its use is accepted by the population at large (Keppel et al., 1992). Validity denotes how well new or novel devices elicit similar responses to those derived from equipment deemed a criterion or "gold standard" by industry experts (Keppel et al., 1992). As it pertains to weight training, standard free weight exercises are known to elicit muscle mass and strength gains under ambulatory conditions, as well as attenuate atrophy and strength losses in simulated µ*g* models (Bamman et al., 1998; Fleck and Kraemer, 2014). Our study's purpose was to compare and assess physiological, perceptual, and exercise performance responses derived from squats done on the IHMC exoskeleton to those from free weights. Comparatively, similar responses to those from free weights suggest the exoskeleton exhibits acceptable levels of validity, which may improve the likelihood of its use as in-flight exercise hardware. We hypothesized squats done on the exoskeleton will evoke similar responses as to those produced when the same exercise is done with free weights.

MATERIALS AND METHODS

Subjects

The University of Tulsa's Institutional Review Board (IRB) approved our protocol in advance of data collection. All current study data were collected at the University of Tulsa. For the current study, subjects performed two squat workouts spaced 14 days apart. Their first workout was done with the exoskeleton, followed by an identical exercise protocol 14 days later with free weights. Due to the late date of the current study's IRB approval, in relation to the date of the exoskeleton's arrival and length of stay in the principal investigator's laboratory, we could not counterbalance the sequence of workouts. For our study, the exoskeleton was transported, and remained in the principal investigator's laboratory for a short (~4 days) period of time before it was returned to the IHMC. The brevity of the exoskeleton's stay required the investigators to schedule and collect data from all 14 subjects over that time period. Healthy, college-age female athletes (n=14) provided written informed consent before their participation. None had injuries that compromised their involvement. They had (mean ± sd) 3.1 ± 0.8 years of experience with the squat exercise prior to their participation. As part of their regular preparation for athletic competition, they routinely performed the back squat exercise 1-2 times per week. Their 1-repetition maximum (1RM) in the back squat at the time of their current study involvement was 76.4 ± 9.7 kg. Absolute strength measurements per current study exercise device were not performed before the start of each workout so that muscle fatigue would not impact workout results. Our subjects' varsity sports participation was as follows: soccer-8, rowing-4, tennis-1, and golf-1. They were told to avoid stimulants, such as caffeine and those contained within dietary supplements, on days they performed current study workouts. They were told to come to workouts well-rested and to avoid lower body resistive exercise 24 hours prior to current study workouts. Subjects ate their pre-exercise meal 1-3 hours before workouts, arrived to our laboratory in athletic attire, and had their data collected between 1300-1700 hours to limit circadian effects. Subjects were instructed to consume their normal lunchtime meal before workouts. Per subject, they were also told to consume identical pre-exercise meals before each workout. Pre-exercise meals had an average energy intake (mean ± sd) of 520 ± 75 kcals, with a macronutrient breakdown as follows: carbohydrates 80 ± 19 g, protein 14 ± 9 g, and fat 16 ± 11 g.

Procedures and Equipment

To begin data collection, subjects submitted to a series of anthropometric measurements. Height, body mass, body fat percentage, hip width and circumference, as well as the lengths of their torso and upper and lower legs, were measured as they stood barefoot in an upright posture. Heights were measured by a stadiometer (Detecto Model 437, Webb City, MO). Body mass and composition were recorded with a calibrated bioimpedance scale (Model BF-350, Tanita Corporation, Tokyo, Japan). All hip, torso, and leg measurements were recorded in triplicate by the principal investigator (J. Caruso) with a cloth measuring tape to the nearest 0.1 cm. Hip width was measured as the lateral expanse between anterior superior iliac spines across the ventral surface of the body. Hip circumference was recorded at the level of the anterior superior iliac spines. Assessed along the left side of subjects' bodies, torso length equaled the distance between the acromioclavicular joint and the anterior superior iliac spine. Upper leg length spanned the distance from the left femur's trochanter to the lateral condyle's lower border. Lower leg length equaled the distance from the left fibula's head to its lateral malleolus.

Physiological measurements were obtained by preparing subjects' bodies for data collection. With aseptic techniques, pre-exercise saliva was obtained with oral swabs (Salimetrics, State College, PA) that were used to quantify cortisol concentrations ([C]) at a later date with enzyme-linked immunosorbent assay (ELISA) kits, and 1-2 fingertip blood drops were placed on test strips inserted within a calibrated device (Accupsort, Hawthorne, NY) to measure blood lactate concentrations ([BLa^-]). Subjects then had a torso monitor and wrist strap (Model FT4, Polar, Kempele, Finland) attached to their bodies to record heart rate (HR) values. After pre-exercise [C], [BLa^-], and HR data were obtained, they sat quietly for five minutes. Subjects' first visits continued with self-administered, passive lower body stretching, which lasted five minutes and focused on areas most heavily engaged (lower back, hips, knees, and ankles) by the squat exercise.

When stretching concluded, subjects stood next to the exoskeleton as the final preparations for data collection began. Surface EMG signals were obtained with a computer-based oscillograph and acquisition system (Model MEB-7102A, Horizon Bio-Medical, Mooresville, NC). A bipolar Ag/AgCl collection electrode, with an inter-electrode distance of 3 cm, was applied to skin marked with ink over the left vastus lateralis. A ground electrode covered the fibular head of subjects' left legs. The electrode was placed 20 cm superior to the fibular head, along the examined muscle's ventro-lateral surface, in order to monitor behavior closer to the knee, a major articulation where movement occurs during squat repetitions. Conduction paste (Elefix, Nihon Kohden, Foothill Ranch, CA) was applied to electrodes to enhance signal quality. Athletic tape was used to adhere the electrodes to the surface of subjects' skin. EMG data were amplified at a bandwidth of 10-1000 Hz and sampled at 2048 Hz. Signals were full wave rectified and low pass filtered at a cut-off frequency of 250 Hz. The gain was adjusted so that the entire signal was captured and viewed as squats were performed. Per subject, EMG procedures were standardized across workouts, which included electrode placement at the same marked location along the thigh, to provide real-time waveforms of muscle activity.

After EMG preparations concluded, subjects performed a few practice repetitions as they wore the exoskeleton. Practice repetitions were performed with no added pre-programmed load. For exoskeleton workouts, its inventor (P. Neuhaus) and the principal investigator were present to ensure repetitions were done correctly. Verified by a photoelectric sensor (Automationdirect, Cumming, GA) for each repetition, subjects descended to a depth whereby their femurs were parallel to the ground before they ascended. All squats were done in cadence with a metronome (MR-600, Matrix, South Korea) at a rate of three seconds per repetition. After practice repetitions concluded, subjects donned a neoprene mask for the collection and analysis of their respiratory gases by a metabolic cart (Model K2b4, Cosmed, Rome, Italy) to quantify their peak O_2 uptakes. Before the first set began, pre-exercise O_2 uptake rates were measured as subjects stood motionless for at least five minutes before their first set. Subjects also wore the mask during and after workouts until their O_2 consumption rates returned to pre-exercise levels.

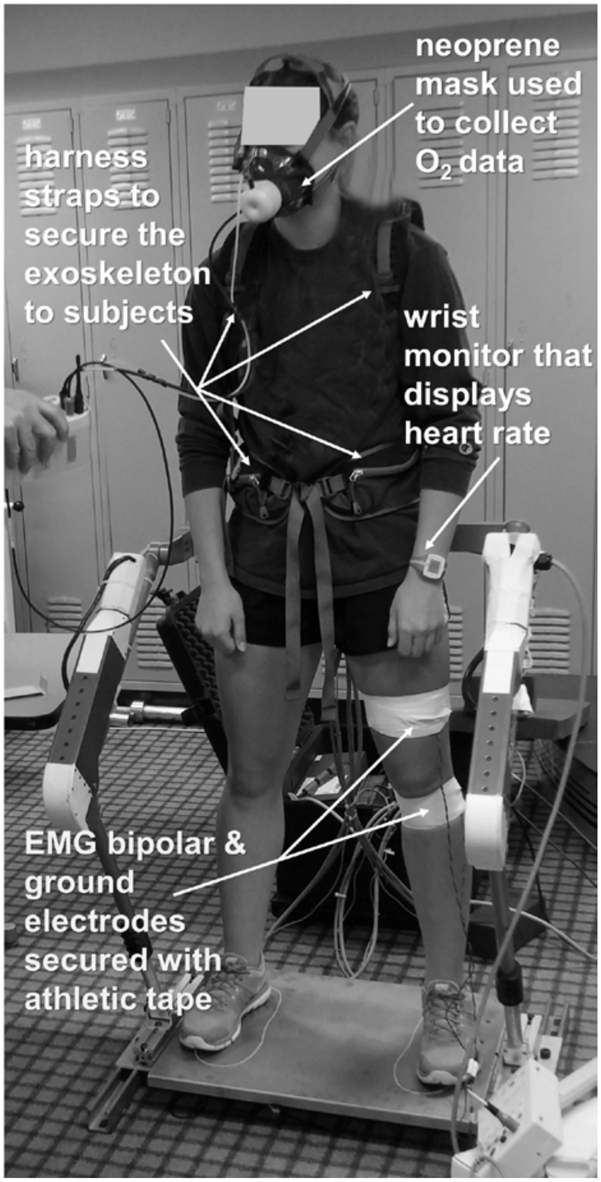

Figure 2. Subjects as they appeared at the start of exercise done on the exoskeleton. On the subject's left leg is an EMG electrode, over the nose and mouth is a mask for quantifying O_2 consumption, and on the torso (not visible) is a heart rate monitor with receiver on the left wrist.

Figure 2 depicts a subject as her data was collected from the exoskeleton workout. They performed a four-set, four-repetition squat protocol with 60-second rests between sets. Subjects did sets against progressively heavier loads in the following order: 22.5 kg, 34 kg, 45.5 kg, and 57 kg. The loads chosen were based on our sample's mean 1RM back squat value, so that repetitions could be safely performed by our subjects. Between each set, subjects stood motionless in an upright posture. Per set, an accelerometer (Myotest Inc., Royal Oak, MI) attached to the back of the exoskeleton measured peak force, peak velocity, and peak power. The accelerometer was shown previously to evoke high levels of validity for the squat exercise (Comstock et al., 2011). EMG data were recorded throughout each set. Peak EMG amplitudes were quantified from each set and used for analysis. HR was measured 30 seconds after each set. Five minutes after the last set concluded, post-exercise [BLa^-] and HR were recorded, as values usually peak for the former dependent variable at that time. Our metabolic cart provided breath-by-breath analysis of O_2 samples. Once O_2 values returned to pre-exercise levels, the mask was removed from their faces and another oral swab was placed in their mouths. Swabs were subsequently analyzed for [C] (Figure 2).

At the conclusion of the exercise bout, subjects provided a rate of perceived exertion (RPE), or their perceptual index of the workout's rigor, with the range of possible values from 1 to 10. Fourteen days after their exoskeleton workout, they performed an identical back squat protocol in our laboratory with free weights. Their data were obtained using identical procedures and methods, with measurements of the same dependent variables. The principal investigator was also present at all free weight exercise bouts to ensure there were no inter-workout differences with respect to the procedures and methods employed. In addition, subjects used the same pre-exercise preparations (meal, avoidance of stimulants, etc.) as those employed before exoskeleton workouts. Thus our study was able to assess the degree of data similarity for values from the exoskeleton squat workouts to those derived from free weight exercise bouts.

Statistical Analyses

We collected numerous dependent variables, mostly of an instantaneous nature, which depicted subjects' efforts over a brief period for each workout. The variety of dependent variables spans a large range of absolute values. Thus to assess the comparative similarity of data and the exoskeleton's validity as exercise hardware, we compared the same dependent variables from both workouts with three statistical tests. First, however, our data were analyzed with Z scores to identify outliers. They were computed as: (individual score – mean)/sd. Absolute Z score values that exceeded -1.96 or +1.96 were excluded from further analyses. Outlier data, as well as its corresponding value from the other workout, were excluded from further analyses.

Our analyses then proceeded to assess the comparative similarity of our data and thereby the validity of the exoskeleton. As used in a prior study that assessed validity, we computed Pearson correlation coefficients to compare values per dependent variable derived from the two workouts (Andresen et al., 1999). We also examined each dependent variable using paired t-tests to assess absolute (unstandardized) inter-workout differences with a Bonferroni adjustment to control for type I error. Per dependent variable, our t-test values with the Bonferroni adjustment were calculated as a ratio to the number of similar dependent variables (metabolic, HR, accelerometry, EMG, etc.) also examined in the current study. Finally, we computed Cohen's d effect size as the relative standardized difference between mean values. To exhibit a high degree of comparative inter-workout similarity and to affirm our hypothesis, current dependent variables had to yield higher (r values from 0.50-1.00) Pearson correlation coefficients, as well as lower t-test (<1) and Cohen's d effect size (<0.4) values.

RESULTS

All subjects successfully completed both workouts and none were injured through their project participation. Anthropometric dimensions (mean ± sd) were as follows: height 171 ± 5 cm, mass 70.2 ± 5.0 kg, body fat 22.8 ± 3.4%, hip width 29.3 ± 3.8 cm, hip circumference 83.6 ± 4.4 cm, torso length 40.6 ± 4.5 cm, upper leg length 41.8 ± 5.3 cm, and lower leg length 42.4 ± 2.5 cm. Z score analyses revealed approximately 10% of our total data set were outliers. Most outliers were peak power and peak EMG values produced from the exoskeleton workout; such instantaneous measurements are more prone to elicit outliers than dependent variables that exhibit more stability over time (Caruso et al., 2013; Davidson et al., 2013). EMG, in particular, is a very sensitive measure; despite the similarity of inter-workout EMG preparations to our subjects' left legs, which included replication of electrode placement over a marked skin site, there is inherently high variability when such data is obtained from dynamic exercise (Davidson et al., 2013). Table 3 and Table 4 each include a column that displays the number of subjects who provided data (with outliers and corresponding values excluded) per dependent variable. Our raw data (mean ± sd; range) from each workout appear in Table 1 and Table 2.

Our results, whereby dependent variables from the exoskeleton workout were compared to

Table 1. Metabolic, heart rate (HR), and rate of perceived exertion (RPE) data from exoskeleton and free weight workouts. Sets 1-4 entailed heavier loads (22.5, 34, 45.5, and 57 kg, respectively) for each workout.

Dependent variable (units)	Exoskeleton workout mean ± sd	Exoskeleton workout range	Free weight workout mean ± sd	Free weight workout range
pre-exercise [BLa^-] ($mmol \cdot L^{-1}$)	1.7 ± 0.8	0.8 - 3.7	1.8 ± 0.9	0.9 - 3.7
post-exercise [BLa^-] ($mmol \cdot L^{-1}$)	2.4 ± 1.4	1.0 - 6.1	2.1 ± 1.0	1.0 - 3.9
pre-exercise [C] ($\mu g \cdot dl^{-1}$)	0.21 ± 0.08	0.10 - 0.37	0.22 ± 0.08	0.10 - 0.42
post-exercise [C] ($\mu g \cdot dl^{-1}$)	0.25 ± 0.11	0.10 - 0.45	0.21 ± 0.15	0.10 - 0.76
peak O_2 ($ml \cdot min^{-1}$)	1144 ± 174	909 - 1457	1278 ± 173	947 - 1535
pre-exercise HR ($beats \cdot min^{-1}$)	65.1 ± 8.0	52 - 75	68.4 ± 9.7	46 - 80
post-set 1 HR ($beats \cdot min^{-1}$)	94.8 ± 14.0	75 - 126	104.9 ± 13.4	84 - 126
post-set 2 HR ($beats \cdot min^{-1}$)	101.1 ± 12.9	83 - 125	111.1 ± 16.3	80 - 134
post-set 3 HR ($beats \cdot min^{-1}$)	110.6 ± 18.2	83 - 151	117.2 ± 17.1	81 - 140
post-set 4 HR ($beats \cdot min^{-1}$)	107.4 ± 17.7	82 - 146	123.4 ± 20.0	89 - 153
post-exercise HR ($beats \cdot min^{-1}$)	78.6 ± 16.0	58 - 115	77.4 ± 14.2	51 - 109
RPE	6.3 ± 1.5	3 - 8	4.1 ± 1.5	2 - 7.5

Table 2. Accelerometry and EMG data from exoskeleton and free weight workouts. Sets 1-4 entailed heavier loads (22.5, 34, 45.5, and 57 kg, respectively) per workout.

Dependent variable (units)	Exoskeleton workout mean ± sd	Exoskeleton workout range	Free weight workout mean ± sd	Free weight workout range
Peak force set 1 (newtons)	283 ± 25	257 - 338	329 ± 55	213 - 440
Peak velocity set 1 ($cm \cdot sec^{-1}$)	71 ± 31	39 - 118	96 ± 13	81 - 122
Peak power set 1 (watts)	169 ± 78	85 - 317	247 ± 52	187 - 349
Peak EMG set 1 (µV)	521 ± 143	275 - 680	322 ± 101	170 - 500
Peak force set 2 (newtons)	443 ± 43	390 - 514	486 ± 122	201 - 663
Peak velocity set 2 ($cm \cdot sec^{-1}$)	69 ± 24	38 - 124	89 ± 27	53 - 140
Peak power set 2 (watts)	251 ± 99	129 - 476	322 ± 89	180 - 479
Peak EMG set 2 (µV)	413 ± 231	95 - 700	350 ± 238	110 - 810
Peak force set 3 (newtons)	618 ± 70	531 - 789	673 ± 151	358 - 879
Peak velocity set 3 ($cm \cdot sec^{-1}$)	413 ± 231	95 - 700	350 ± 238	110 - 810
Peak power set 3 (watts)	352 ± 153	135 - 576	394 ± 95	282 - 531
Peak EMG set 3 (µV)	616 ± 306	300 - 1100	511 ± 321	130 - 1100
Peak force set 4 (newtons)	761 ± 101	657 - 1060	752 ± 125	479 - 1030
Peak velocity set 4 ($cm \cdot sec^{-1}$)	77 ± 28	37 - 124	86 ± 30	30 - 124
Peak power set 4 (watts)	387 ± 111	220 - 591	526 ± 202	170 - 734
Peak EMG set 4 (µV)	645 ± 331	100 - 1005	550 ± 378	80 - 1270

Table 3. Metabolic, HR, and RPE results, whereby dependent variable values from the exoskeleton and free weight workouts were compared to note the degree of similarity. Included are the number of subjects (n) who provided paired values for analysis.

Dependent variable	r	t-tests	Cohen's d	n
pre-exercise [BLa^-]	0.01	0.11	0.04	14
post-exercise [BLa^-]	0.51	0.56	0.26	14
pre-exercise [C]	0.07	0.32	0.13	14
post-exercise [C]	0.24	0.80	0.30	14
peak O_2	0.14	1.94	0.73	14
pre-exercise HR	0.57	1.50	0.37	14
post-set 1 HR	0.61	3.10	0.70	14
post-set 2 HR	0.79	3.80*	0.65	14
post-set 3 HR	0.55	1.50	0.38	14
post-set 4 HR	0.50	3.10	0.79	14
post-exercise HR	0.78	1.70	0.37	14
RPE	0.51	5.40*	1.17	14

*: statistically ($p<0.05$) different inter-workout values

Table 4. Accelerometry and EMG results, whereby dependent variables from the exoskeleton and free weight workouts were compared to note the degree of similarity. Included are the number of subjects (n) who provided paired values for analysis.

Dependent variable	r	t-tests	Cohen's d	n
Peak force set 1	0.44	2.20	0.95	14
Peak velocity set 1	0.34	2.40	0.91	13
Peak power set 1	0.01	2.40	1.02	11
Peak EMG set 1	0.66	2.30	1.31	11
Peak force set 2	0.01	1.00	0.47	14
Peak velocity set 2	0.20	1.50	0.73	13
Peak power set 2	0.73	2.70	0.71	11
Peak EMG set 2	0.05	0.20	0.46	11
Peak force set 3	0.55	0.90	0.47	14
Peak velocity set 3	0.31	2.30	0.55	13
Peak power set 3	0.46	0.00	0.32	13
Peak EMG set 3	0.12	0.30	0.16	13
Peak force set 4	0.50	0.30	0.08	14
Peak velocity set 4	0.38	0.70	0.32	13
Peak power set 4	0.01	1.44	0.82	13
Peak EMG set 4	0.28	0.01	0.27	13

the same indices obtained from the free weight exercise bout, appear in Table 3 and Table 4. Table 3 includes metabolic, HR, and RPE results. Table 3 Pearson correlation coefficient results show post-exercise [BLa^-], all HR, as well as RPE dependent variables produced higher r values and thus the most inter-workout agreement. In contrast, pre-exercise [BLa^-] and [C], post-exercise [C], and peak O_2 display more inter-workout variability. The variability in post-exercise [C] may be due to time differences at which saliva was obtained. Post-exercise saliva collection was delayed until the mask, worn to assess O_2 uptake, was removed so that a swab could be placed in subjects' mouths. Since post-exercise collection times between workouts varied slightly, that is a potential source of variability. Table 3 t-test results exhibit non-significant inter-workout differences for [BLa^-], [C], peak O_2, and some of our HR data. However, Table 3 t-test results also include significant inter-workout differences for post-set 2 HR and RPE, with higher HR data from free weight, yet greater RPE values from exoskeleton workouts. Table 3 Cohen's d results show pre- and post-exercise [BLa^-] and [C] values produced the smallest standardized inter-workout differences. Yet most Table 3 Cohen's d results include far higher values, indicative of greater inter-workout variability and less similarity for responses obtained from both workouts.

Table 4 displays our accelerometry and EMG results. Table 4 Pearson correlation coefficient results show peak EMG set 1, peak power set 2, and peak forces for sets 3 and 4 each yielded r values of 0.50 or greater. Yet most Table 4 dependent variables exhibited far weaker inter-workout correlations. In contrast to Table 3, our Table 4 t-test results exhibit non-significant inter-workout differences for all its dependent variables. Table 4 Cohen's d results show low standardized differences for peak EMG for sets 3 and 4, as well as peak force for set 4. Yet the majority of our Table 4 Cohen's d results include far higher values, which exhibit greater inter-workout variability and less comparative similarity between data obtained from both workouts.

DISCUSSION

For our study, the similarity of results from both workouts was used to assess the validity of data provided by the exoskeleton. Validity refers to the extent data from new and established devices are similar (Keppel et al., 1992). Different forms of validity exist. Our study compared the similarity in responses derived from exoskeleton and free weight workouts, and is best described as an examination of convergent validity, which refers to the degree two data sets that in theory should be related, actually are (Keppel et al., 1992; Measurement Validity Types, 2015).

Research examined the ability of in-flight exercise hardware to mitigate muscle mass and strength losses produced by long-term stays aboard the ISS (Smith et al., 2012; Trappe et al., 2009). Despite use of in-flight aerobic hardware 5 days/week at moderate intensities, and concurrent strength training on an interim resistive exercise device (iRED) 3-6 days/week with several lower body resistive exercises, crewmembers experienced significant plantar flexor mass and strength losses after six months on the ISS (Trappe et al., 2009). In addition, there were in-flight changes to muscle fibers associated with a reduced resistance to fatigue (Trappe et al., 2009). It was concluded the ISS should be equipped with hardware that offers a greater mechanical loading stimulus to better attenuate muscle mass and strength losses (Trappe et al., 2009). That conclusion was affirmed by a bone study done on the ISS that compared the merits of different forms of in-flight hardware (Smith et al., 2012). With no crossover, crewmembers were assigned to resistive exercise done with either the iRED or ARED for their 4-6 month stays (Smith et al., 2012). Unlike the iRED, which employed elastic bands for resistance, the ARED uses pneumatic cylinders and flywheels to simulate the manner in which weights are lifted on Earth. Bone losses, measured before and after flights, were best abated in those who exercised on the ARED, which provided comparatively more resistance (Smith et al., 2012).

Ground-based studies also examined the utility of actual and potential in-flight hardware as prospective exercise countermeasures (Beck et al., 2014; Loehr et al., 2011; Rea et al., 2013). A training study compared physiological changes from strength training on the ARED to free weights (Loehr et al., 2011). Assigned to one of two groups with no crossover, subjects did identical workouts, which included the squat exercise, three days/week for 16 weeks. Results showed both groups incurred similar improvements over time. It was concluded physiological changes from ARED workouts were like those of free weights (Loehr et al., 2011). Yet it is important to interpret differences between the current and ground-based ARED results with caution. For instance, the ARED study examined chronic changes in physiology; over time, actual inter-group differences could be assessed. In contrast, the current trial quantified acute changes to dependent variables of a far more instantaneous nature, in which recorded values were attained for only a very brief time period. Thus our dependent variables inherently exhibit more data variability. In addition, the ARED study did not assess convergent validity, but rather compared changes over time in both groups (Loehr et al., 2011).

Little research exists on the convergent validity of exoskeleton data, with the goal of improved in-flight exercise hardware (Beck et al., 2014; Rea et al., 2013). Knee extension and flexion torque data from NASA's X1 exoskeleton were compared to those derived from an isokinetic dynamometer (Beck et al., 2014). Subjects performed one workout on each device. Results showed high levels of agreement for knee extension, but not knee flexion, torques obtained from each device (Beck et al., 2014). It was suggested the X1 could be used to assess lower body muscle strength (Beck et al., 2014). The validity of that data exceeds that of the current study's; yet it is important to note X1 and dynamometer torque values were derived from single joint exercises, unlike our study that entailed squats—a dynamic, multi-joint movement of far greater methodological rigor. Thus the validity of our data is expected to be less than that of the X1 paper based on the exercises examined (Beck et al., 2014).

The achievement of comparatively similar data from both of our workouts was made difficult by numerous factors inherent to the current study. They include performance of a multi-joint, multiplanar movement at relatively high velocities with a novel exoskeleton (Caruso et al., 2013; Caruso et al., 2012; Davidson et al., 2013). Yet the low degree of comparative similarity, as seen in some Table 3 and Table 4 results, is still disconcerting and the reason we did not address our study hypothesis. Due to the limited inter-workout similarity, our exoskeleton data does not exhibit an acceptable degree of convergent validity. This is most likely due to the brief amount of familiarization (a few practice repetitions) with the exoskeleton prior to actual data collection, which is a serious limitation. For our study, the exoskeleton was transported and stayed in the principal investigator's laboratory for a short (~4 days) period of time before it had to be returned to the IHMC. The brevity of the exoskeleton's stay, combined with our subjects' busy schedules, did not afford them an ideal opportunity to familiarize themselves to squats done with the exoskeleton. Clearly, limited familiarization with a complex dynamic exercise performed on a novel device appears to have impacted our results.

Anecdotal claims from 13 subjects inferred free weight squats were easier. This claim is supported by our RPE data, despite the same number of sets and repetitions, rest periods, and loads used for each workout. Due to their considerable background squatting with a barbell, subjects could perhaps put more effort into repetitions for that exercise mode, which could in part explain the higher HR values seen with free weight squats. In contrast, subjects stated the exoskeleton distributed loads in a manner they were unaccustomed to as they performed repetitions. Differences in load distributions and the resultant kinesthetic and biomechanical changes, as well as limited familiarity with the exoskeleton, likely made those squats comparatively more difficult. This is supported by our post-exercise [BLa^-] and [C] values, as well as peak EMG data, which show higher mean values from exoskeleton workouts. Higher peak EMG data from the exoskeleton exercise bout, despite generally greater performance-based values from free weight workouts, suggests subjects recruited more motor units for the exercise device they were less familiar with, which concurs with prior research (Jakobsen et al., 2013).

Familiarization requirements are based on the nature of the task and the length of inter-session time intervals (Donovan and Radosevich, 1999). The number of familiarization sessions done by human subjects prior to actual data collection certainly impacts convergent validity results. Less familiarization is required for tasks with a low methodological rigor (Sleivert and Wenger, 1994; Viitasalo et al., 1980); the opposite is true of squats, a complex dynamic motor skill that requires refined patterns of muscle activity executed in a specific sequential fashion (Donovan and Radosevich, 1999; Frost et al., 2012; Lee and Genovese, 1989). Recent work on the required number of familiarization sessions examined exercises less rigorous than the squat, and included dependent variables, such as subjects' 1RM values, which tend to exhibit less variability than our study's performance-based dependent variables, which is in part due to the speed at which repetitions are performed. Current study repetitions occurred at faster velocities than are generally seen with 1RM attempts, which lead to higher rates of movement and more data variability (Caruso et al., 2013; Caruso et al., 2012; Davidson et al., 2013). In exercise studies with low rigor (e.g., vertical jump, isometric contractions, and elastic band exercise), familiarization occurred with a single session administered prior to actual test data collection (Calder and Gabriel, 2007; Colado et al., 2014; Frost et al., 2012), or no familiarization whatsoever for vertical jumps done by physically active men (Moir et al., 2004).

Familiarization requirements were compared among young (23 ± 4 years) and old (66 ± 5 years) women who each performed multiple knee extension 1RM tests (Ploutz-Snyder and Giamis, 2001). They engaged in at least two test sessions; if their 1RM values exceeded the prior sessions' by more than 1 kg, they were required to perform an additional trial. Older women required more familiarization (8-9 sessions) than younger subjects (3-4 sessions) to achieve consistent 1RM values (Ploutz-Snyder and Giamis, 2001). While more familiarization improves validity, too many sessions may induce a training effect and thus not reflect subjects initial performance capabilities. Two such studies examined the number of familiarization sessions needed to achieve valid 1RM values for three exercises (bench press, squat, and arm curl) in women (Soares-Caldeira et al., 2009) and men (Dias et al., 2005). To derive

1RM values, subjects performed four (Dias et al., 2005) or five (Soares-Caldeira et al., 2009) sessions spaced 2-3 days apart. There was a consistent rise in 1RM values for each exercise over successive trials, which infers multiple tests induced a training stimulus and caused the required number of sessions recommended to be inflated (Dias et al., 2005; Soares-Caldeira et al., 2009). Multiple tests may evoke a 5-10% strength gain, which can be avoided by spacing sessions farther (7-10 days) apart (Schroeder et al., 2007).

Hardware development for in-flight exercise must address the adverse impacts long-term µ*g* has upon human physiology, while conforming to the novel requirements of the spacecraft environment (Carpinelli, 2014; Hargens et al., 2013). Establishment of acceptable levels of convergent validity is essential before exercise hardware can be used as an in-flight countermeasure to muscle atrophy and strength losses. The IHMC hopes to have their exoskeleton flown on future spaceflights. An acceptable level of convergent validity was not established for the IHMC exoskeleton as it was used in the current study. Based on prior outcomes, and given our subjects and their background with the squat exercise, it appears one familiarization session to accustom them to the exoskeleton before actual data collection would have certainly increased our study's convergent validity and improved the likelihood of hypothesis affirmation. However, since crewmembers are generally less fit and do not have as extensive a resistance exercise background as the current subjects, two or more familiarization sessions under ground-based and simulated µ*g* conditions may be warranted if the IHMC exoskeleton becomes in-flight hardware. Due to its novelty and its considerable promise as in-flight hardware, future research studies that involve the IHMC exoskeleton are warranted. Such research should first attempt to establish enough comparatively similarity between data obtained from exoskeleton and free weight workouts to establish acceptable levels of convergent validity.

With ~10% of our total data as outliers, which is usually higher than that seen in other trials, it appears to have foretold the rejection of our hypothesis. The lack of acceptable levels of convergent validity was made difficult by several limitations, which should be addressed in future exoskeleton studies. Some limitations are inherent to squats, which are complex motor skills with a high methodological rigor. A major limitation that prevented acceptable convergent validity, and could address the aforementioned concern with the squat exercise, is the brief amount of familiarization subjects had to the exoskeleton prior to data collection. Depending on the subjects employed and the methods by which data are obtained, future studies that involve exoskeleton squats should include one or more familiarization sessions with the device prior to actual data collection. Another current study limitation may include an order effect whereby due to its brief stay in the principal investigator's laboratory, we initially collected our exoskeleton squat data, followed by the free weight workout, 14 days later. Yet the occurrence of an order effect was likely blunted by the large disparities in the degree of familiarity subjects had with the two types of exercise devices examined. Future trials may wish to vary the sequence that subjects perform workouts in order to reduce the likelihood of an order effect. With respect to our post-exercise [C] results, the convergent validity of that data may have been higher if the time point at which those measurements were obtained was consistent. Since, for our study, we waited until O_2 uptake rates returned to pre-exercise levels in order to remove the mask and subsequently insert the oral swab for saliva collection, this is a potential source of variability. New exoskeleton trials may wish to employ different methods in order to increase the convergent validity of post-exercise [C] measurements. Finally, most of our outliers came from dependent variables that entailed instantaneous measurements, in which recorded values were attained for only a very brief time period. Future studies may wish to assess the convergent validity of the exoskeleton with dependent variables whose values show greater stability over time. Adoption of these recommendations in future exoskeleton trials should increase convergent validity and thus improve the likelihood of its use as in-flight exercise hardware.

ACKNOWLEDGEMENTS

We wish to thank our subjects for their participation. Many of the authors are participants in the TURC (Tulsa Undergraduate Research Challenge) program at the University of Tulsa.

CONFLICT OF INTEREST STATEMENT

The IHMC is a not-for-profit research center with no commercial interests in the outcomes of this study. Likewise, the universities for whom the authors are affiliated have no commercial interests in the outcomes of this study. The prototype exoskeleton examined in this study was fabricated for research and is not a commercial device, nor is it commercially available.

REFERENCES

Andresen EM, Fouts BS, Romeis JC, Brownson CA (1999) Performance of health-related quality of life instruments in a spinal cord injured population. *Archives of Physical Medicine and Rehabilitation* **80:** 877-884

Bamman MM, Clarke MSF, Feeback DL, Talmadge RJ, Stevens BR, Lieberman SA, Greenisen MC (1998) Impact of resistance exercise during bed rest on skeletal muscle sarcopenia and myosin isoform distribution. *Journal of Applied Physiology* **84:** 157-163

Beck CE, Rovekamp RN, Rea RL, English KL, Newby NJ, Hackney KJ, De Witt JK, Ploutz-Snyder LL (2014) *Comparison of Knee and Ankle Dynamometry between NASA's X1 Exoskeleton and Biodex System 4.* NASA Human Research Program Investigators' Workshop, 3058. Oral/Visual Presentation Number JSC-CN-29878

Calder KM, Gabriel DA (2007) Adaptations during familiarization to resistive exercise. *Journal of Electromyography and Kinesiology* **17:** 328-335

Carpinelli RN (2014) Exercise countermeasure to weightlessness during manned spaceflight. *Medicina Sportiva* **18:** 42-44Caruso JF, Brown LE, Tufano JJ (2012) The reproducibility of isokinetic dynamometry data. *Isokinetics and Exercise Science* **20:** 239-253

Caruso JF, Learmonth LA, Charles JL, Unruh KD, Giebel RM, Walker RH, West JO (2013) Anthropometry and high-speed elbow flexor performance. *Isokinetics and Exercise Science* **21:** 83-91

Colado JC, Garcia-Masso X, Triplett NT, Calatayud J, Flandez J, Behm D, Rogers ME (2014) Construct and concurrent validation of a new resistance intensity scale for exercise with Thera-Band elastic bands. *Journal of Sports Science and Medicine* **13:** 758-766

Comstock BA, Solomon-Hill G, Flanagan SD, Earp JE, Luk HY, Dobbins KA, Dunn-Lewis C, Fragala MS, Ho JY, Hatfield DL, Vingren JL, Denegar CR, Volek JS, Kupchak BR, Maresh CM, Kraemer WJ (2011) Validity of the Myotest in measuring force and power production in the squat and bench press. *Journal of Strength and Conditioning Research* **25**: 2293-2297

Davidson ME, Shepherd CM, Biga J, Davison SW, Giebel RM, Charles JL, Learmonth LA, Caruso JF (2013) Electromyography and high-speed elbow flexor exercise. *Isokinetics and Exercise Science* **21:** 141-150

Dias RMR, Cyrino ES, Salvador EP, Caldeira LFS, Nakamura FY, Papst RR, Bruna N, Gurjao ALD (2005) Influence of familiarization process on muscular strength assessments in 1-RM tests. *Revista Brasileira de Medicina do Esporte* **11:** 34-38

Donovan JJ, Radosevich DJ (1999) A meta-analytic review of the distribution of practice effect: now you see it, now you don't. *Journal of Applied Psychology* **84:** 795-805

Fleck SJ, Kraemer WJ (2014) *Designing Resistance Training Programs*, 4th edn, Champaign, IL: Human Kinetics

Frost LR, Gerling ME, Markic JL, Brown SHM (2012) Exploring the effect of repeated-day familiarization on the ability to generate reliable maximum voluntary muscle activation. *Journal of Electromyography and Kinesiology* **22:** 886-892

Hargens AR, Bhattacharya R, Schneider SM (2013) Space physiology VI: exercise, artificial gravity, and countermeasure development for prolonged spaceflight. *European Journal of Applied Physiology* **113:** 2183-2192

Jakobsen MD, Sundstrup E, Andersen CH, Aagaard P, Andersen LL (2013) Muscle activity during leg strengthening exercise using free weights and elastic resistance: effects of ballistic vs. controlled contractions. *Human Movement Science* **32:** 65-78

Keppel G, Saufley WH, Tokunaga H (1992) *Introduction to Design and Analysis: A Student's Handbook*, 2nd edn, New York: Worth Publishers

Lee TD, Genovese ED (1989) Distribution of practice in motor skill acquisition: different effects for discrete and continuous tasks. *Research Quarterly for Exercise and Sport* **60:** 59-65

Loehr JA, Lee SMC, English KL, Sibonga J, Smith SM, Spiering BA, Hagan RD (2011) Musculoskeletal adaptations to training with the Advanced Resistive Exercise Device. *Medicine and Science in Sports and Exercise* **43:** 146-156

Measurement Validity Types (2015) *Research Methods Knowledge Base*. Retrieved from: http://www.socialresearchmethods.net/kb/measval.php (Accessed 04/29/16)

Moir G, Button C, Glaister MZ, Stone MH (2004) Influence of familiarization on the reliability of vertical jump and acceleration sprinting performance in physically active men. *Journal of Strength and Conditioning Research* **18:** 276-280

Ploutz-Snyder LL, Giamis EL (2001) Orientation and familiarization to 1RM strength testing in old and young women. *Journal of Strength and Conditioning Research* **15:** 519-523

Rea R, Beck C, Rovekamp R, Neuhaus P, Diftler M (2013) *X1: A Robotic Exoskeleton for In-Space Countermeasures and Dynamometry*. AIAA SPACE 2013 Conference and Exposition, AIAA Technical Paper Number 2013-5510

Schroeder ET, Wang Y, Castaneda-Sceppa C, Cloutier G, Vallejo AF, Kawakubo M, Jensky NE, Coomber S, Azen SP, Sattler FR (2007) Reliability of maximal voluntary muscle strength and power testing in older men. *Journals of Gerontology* **62:** 543-549

Sleivert GG, Wenger HA (1994) Reliability of measuring isometric and isokinetic peak torque, rate of torque development, integrated electromyography, and tibial nerve conduction velocity. *Archives of Physical Medicine and Rehabilitation* **75:** 1315-1321

Smith SM, Heer MA, Shackelford LC, Sibonga JD, Ploutz-Snyder LL, Zwart SR (2012) Benefits for bone from resistance exercise and nutrition in long-duration spaceflight: evidence from biochemistry and densitometry. *Journal of Bone and Mineral Research* **27:** 1896-1906

Soares-Caldeira LF, Ritti-Dias RM, Okuno NM, Cyrino ES, Gurjão ALD, Ploutz-Snyder LL (2009) Familiarization indexes in sessions of 1-RM tests in adult women. *Journal of Strength and Conditioning Research* **23:** 2039-2045

Stauber, WT (1989) A unique problem of muscle adaptation from weightlessness: the deceleration deficiency. In *Workshop on Exercise Prescription for Long-Duration Spaceflight*, B.A. Jr. Harris, D.F. Stewart (eds), pp 57-59, Washington, DC: NASA

Trappe S, Costill D, Gallagher P, Creer A, Peters JR, Evans H, Riley DA, Fitts RH (2009) Exercise in space: human skeletal muscle after 6 months aboard the International Space Station. *Journal of Applied Physiology* **106:** 1159-1168

Viitasalo JT, Saukkonen S, Komi PV (1980) Reproducibility of measurements of selected neuromuscular performance variables in man. *Electromyography and Clinical Neurophysiology* **20:** 487-501