Striving for Perfection: Measurement of Incremental Fitness Improvement

Meg Quintero
Harvard University
514 Kirkland Mail Center, Cambridge, MA 02138
mmquint@college.harvard.edu

Vladimir Bok
Harvard University
117 Eliot Mail Center, Cambridge, MA 02138
Vladimirbok@college.harvard.edu

ABSTRACT
TBA

Author Keywords
Kinect; Metabolic Conditioning;

INTRODUCTION

Motivation
Intensive workout programs such as P90x, Insanity, and Rushfit promise users to transform their bodies into the best shape of their life by following a 60 to 90 days workout regimen. These types of routines feature methodologies such as Muscle Confusion, Max Interval Training, or High Intensity Interval Training, which all fall into the Metabolic Conditioning exercise category—exercises that increase the storage and delivery of energy for any activity through the improvement of strength and endurance. While a multitude of users report great results, many do not endure long enough to reap the fitness benefits of these programs. Prior research (see related work section) has shown that motivation improves exercise adherence. Accordingly, it is our belief that one of the main reasons people fail to persevere these exercise regimen is the lack of observable progress from one workout to the next, a crucial motivating factor. Unlike strength training exercises–such as bench press or deadlift–the incremental improvements of Metabolic Conditioning workouts are often not apparent until weeks into the program, causing many users to lose motivation and quit the program early in the 60-90 days regimen. The goal of our study is to leverage motion tracking devices such as the Microsoft Kinect to create a software tool for calculating incremental progress for Metabolic Conditioning exercises and ultimately motivate users to adhere to their Metabolic Conditioning routine of choice.

Approach
The intensity of a MC type exercise movement can be measured by the amount of resistance (either body resistance or use of external weights), the volume of repetitions performed and the power or rate at which work is performed. It is our goal to use relatively inexpensive technology, such as the Microsoft Kinect, to measure exercisers progress with a reasonable accuracy. We aim to meet with our fitness experts to create a workout in the style of P90x and Insanity that features exercises that fit the following criteria:

- Requires little to no prior skill or training
- Requires strength to accurately achieve the best stance for that particular move
- Requires stamina to complete the exercise move at a fast rate.

We will use the performance of our experts with the generated MC workout routine and use these accumulation of these and a first set of participants' performance to establish the expert fitness benchmark. We then plan to both record the footage of each participant completing the routine and use the depth sensor to measure factors such as rate of exercise move, body form (predominantly angles between limbs), speed, height (where applicable), the duration and frequency of resting time, and the volume of repetitions performed. We will then calculate a score or percentage of the expert benchmark using the collected data as well as generating more granular data specific to each workout move completed. After the participant has finished the workout, we will show them the two different reports 1.) The expert percentage score and 2.) The granular exercise move specific report and ask them to determine which report would be most helpful for them in terms of tracking their performance from workout to workout.

Contribution
If successful, our work will show that currently hard to benchmark exercises can be evaluated in terms of performance with relatively inexpensive hardware. We predict that our measurement tool could be meaningful to potential users to track minute performance improvements and, ultimately, inspire them to continue working towards their fitness goals. Much research has already been done regarding measuring exercise performance using technology that has inspired and motivated this research project. We believe our work addresses a hole in this field that has not been explored that we believe will spark future work in the field of exercise motivation and progress analysis through the use of inexpensive technology.

RELATED WORK

The notion of tracking ones progress and quantifying improvement is an old one. As early as the 18th century, Benjamin Franklin kept a daily journal to monitor his performance in thirteen personal virtues [BENJAMIN]. Modern research provides formal basis for what Benjamin Franklin seemed to know intuitively: the idea that measurement improves adherence and may ultimately promote long-term behavioral change. For example, Melody Noland showed that self-monitoring improves frequency and adherence to unsupervised exercise [NOLAN]. Despite the encouraging results, Nolands methodology may today seem outdated: the self-monitoring subjects were required to keep a written log of their workouts [NOLAN], leaving much to be desired in terms of usability and automation of what may otherwise be a very tedious process. Indeed, in a comprehensive study with over 160 participants, James J. Annesi demonstrated that automated, computer-based feedback system leads to better exercise adherence and attendance as well as lower dropout rates compared to traditional paper-and-pencil exercise-tracking system [ANNESI].

As if in response to Annesis 1998 study, the past decade has witnessed the onset of technologies that automate exercise tracking, particularly by using human motion capture. Tools were developed to measure exercise form as well as its frequency. Focusing on exercise form, Alexander et. al. developed a video-based motion tracking system to provide feedback about posture and stability of elders during exercise [ALEXANDER], and Silva et. al. used Body Sensor Network (BSN) to provide real-time corrective feedback during the performance of an exercise routine [SILVA]. Focusing on exercise frequency, Chang et. al. used accelerometer embedded in workout gloves to detect different weight training exercises and compute the rate at which they were performed.

Despite the great advances in exercise measurement, limited attention has been given to the type of feedback users would find most helpful. Indeed, the type of feedback researchers provided to their participants varied widely: from granular ones such as the one supplied by Annesi [ANNESI] or Alexander et. al. [ALEXANDER] to holistic ones such as the one Sunny Consolvo and her co-authors created when they sought to display a participants progress in the form of an animated garden [CONSOLOVO]. Similar dichotomy exists in the industry. While some productssuch as Fitbit [FITBIT] or Jawbone UP [JAWBONE]seek to provide detailed overview of the users activity, from distance travelled and stairs climbed to the number of hours slept, otherssuch as the Nike FuelBand [NIKE]seek to abstract away the wearers daily activity into a single metric, the so-called Fuel Points [NIKE]. Accordingly, the purpose of our research was to employ affordable technology to accurately track both the form and rate of a wide range of exercises, and use the output to investigate which type of feedbackholistic or granularusers would find most helpful as a means of tracking progress and improvement over time.

METHODOLOGY

TBA

Workout Routine

The participants each performed three sets of exercises moves. Each set contained three different moves and each was done for a total of 30 seconds. After a participant completed the three moves, they performed the ”resting move” for a total of 60 seconds and then repeated the set. The three sets are as follows:

1. Set 1
   (a) Move 1: Jumping Jacks
   (b) Move 2: Arm Circles
   (c) Move 3: Knee to Elbows
2. Set 2
   (a) Move 1: Squat Jumps
   (b) Move 2: Side to Side Jumps
   (c) Move 3: High Marches
3. Set 3
   (a) Move 1: Jumping Jacks
   (b) Move 2: Arm Circles
   (c) Move 3: Tuck Jumps

Micosoft Kinect Data Collection

The Microsoft Kinect utilizes an infrared emitter and a IR monochrome CMOS (complimentary metal-oxide semiconductor) sensor to measure and sense depth. [Link to microsoft webpage] The emitter emits the infrared light beams and the CMOS sensor reads these IR beams that are reflected back to the sensor. These beams are then converted into depth information measuring the distance between an object and the sensor. Due to this functionality, we were able to measure and track depth and obtain an x, y, and z point for each of the data points. During the entirety of the workout, the participant stood facing the Kinect sensor in order to accurately measure their movements. The interface we designed to track the workout movement featured the skeleton drawing of the participant as well as a ”Start” button to indicate that the participant session has begun, a button for each of the seven moves, and a ”Stop” button to stop tracking the participant while they are completing the resting move. Please see figure 2 for a screen capture of the Kinect interface we designed.

Regarding the technical specificiation of the Kinect and its capabilities, the SDK features built in functionality that provides a deep understanding of human characteristics. The SDK features skeletal tracking, facial tracking, and gesture recognition. The Kinect Fusion algorithm makes it relatively computationally inexpensive to continuously scan the scene by recovering the camera position from the previously scanned point. We took advantage of both the skeletal tracking functionality and the Kinect Fusion algorithm to store the x, y, and z points of the left ankle, right ankle, left elbow, right elbow, left foot, right foot, head, left wrist, right wrist, and spine. From this we were able to use our post processing script to determine the number of exercises moves done

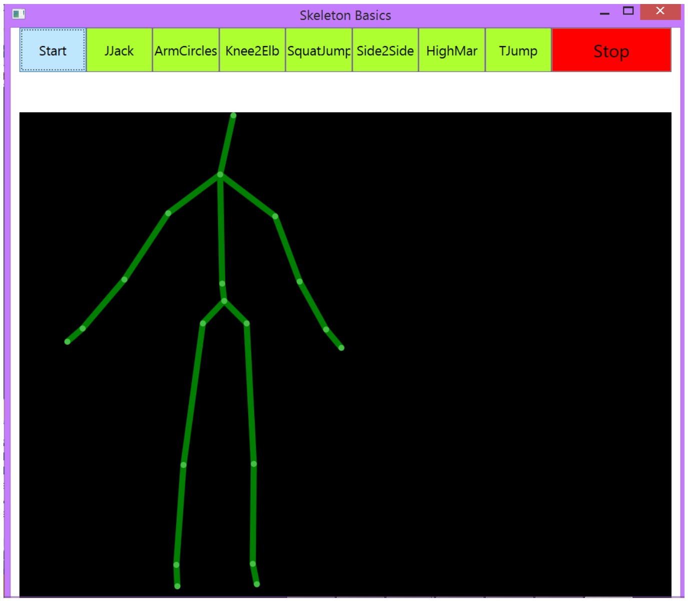


Figure 1. The Kinect Sensor Interface

within the time period and, for a sample of the exercise moves (High marches, squat jumps, and jumping jacks), the participants' from while completing that move.

Post Processing

After movement data were collected, we ran our analysis to compute volume rates for each exercise and form measurement for three selected exercises: jumping jacks, high marches, and squat jumps. Below we describe the data analysis procedures we performed. (The first step, omitted from the descriptions below, was essential data cleansing by deleting rows containing NULL values. The number of such rows did not exceed two per any of the data files, amounting to less than 0.002

Rate Calculations

Jumping Jacks

Jumping jacks consist of starting position (standing still, feet together) and jumping position (feet spread wide). To compute repetitions, we took the x coordinates of a participants left and right feet and plotted the absolute difference between the two coordinates over time, as shown in figure (). (JJ FIRGURE RAW) Each of the lows corresponds to a starting position and each of the peaks corresponds to a jumping position. Therefore, to count the number of repetitions, it suffices to count the number of peaks. Given the relative noisiness of the original data set, a naive algorithm to compute local maxima would yield an excessive number of false positives, necessitating further data processing. We experimented with two methods: smoothing and robust peak detection. In terms of smoothing, we applied a third-order spline interpolation. For robust peak detection, we classified a point as a local maximum if and only if it was preceded by a value lower than some delta, using an algorithm provided by Eli Billauer (BILLAUER). The best value of delta was determined by trial-and-error using our training data set consisting of three subjects (this data set was kept separate from our study data set). While each of theses methods led to significant improvements by itself, as determined by a dramatic reduction in the number of false positives, we found that we obtained the best results when we used both techniques: First, we smoothed-out the data, and then we applied the peak detection algorithm. Figure () shows the curve after smoothing (JJ FIGURE SMOOTH). To compute the rate, we took the number of detected peaks and divided by the duration of the exercise.

Squat Jumps

Similarly to jumping jacks, squat jumps consist of two basic positions: regular squat (starting posture) and a jump. To measure repetitions, we took the average y-coordinate values of a participant's feet and plotted the averages over time. As anticipated, the data revealed a pattern with distinct peaks, corresponding to each of the jumps. Figure () shows the curve after smoothing (SJ FIGURE SMOOTH). As in processing jumping jacks, we applied the peak detection algorithm on the smoothed-out curve, obtaining the number of repetitions, and then divided the count by exercise duration, to obtain the rate.

High Marches

High marches consist of alternatively kicking up as high as possible right and left feet, preferably while keeping knees straight. In analyzing high marches, we considered each leg separately. We took the y-coordinates of left and right ankles, smoothed-out the data and plotted it over time time. Figures () shows the curve for right ankle after smoothing (SJ FIGURE SMOOTH). As in our earlier analyses, the number of peaks gave the number of repetitions, which we divided by duration to obtain the rate.

Side-to-Side Jumps

As the name suggests, side-to-side jumps consist of jumping from left to right, while keeping feet together. To calculate the rate, we took the average x-position of a participants feet, smoothed-out the data, and plotted it over time. Then, we computed the number of peaks (corresponding to the time when the participants feet were at rightmost point) and lows (corresponding to the time when the participants feet were at the leftmost point). Figures () shows the curve after smoothing (SS FIGURE SMOOTH). We then used the sum of those numbers divided by exercise duration to obtain the rate.

Knee to Elbow

To perform knee-to-elbow, a participant twists his torso while lifting his left knee to touch his right elbow, and then repeats on the opposite side, bringing up his right knee to touch his left elbow. From the standpoint of rate calculation, this exercise was more challenging to compute than any of the prior four. We took the x and y coordinates of a participants left knee and right elbow. Then, we computed the Euclidean distance of these points over time. We repeated the same for right knee and left elbow and plotted the differences over time. Figures () shows the curve after smoothing (KE FIGURE SMOOTH). The lows correspond to the points when the participant's knee was closest to the corresponding elbowindicating a completed repetition. Accordingly, we counted the number of lows on the smoothed-out curves for both knee-elbow pairs using the same strategy we used to detect peaks. A sum of those lows was then used to compute the rate.

Tuck Jumps
Tuck jumps require the participant to jump as high as he or she can, raising his or her knees up. Here, we plotted the average y-position of the participants knees over time; accordingly, each peak corresponds to a completed jump. Figures () shows the curve after smoothing (TJ FIGURE SMOOTH). Similarly to the above-described procedures, we used the number of peaks to compute the rate.

Arm Circles
In arm-circles starting position, the participant stands upright with arms extended to his or her sides, parallel to the ground. To perform the exercise, the participant makes circles with his or her outstretched arms, keeping elbows locked. To determine the number of repetitions, we observed that participants hand follows a roughly circular trajectory in the yz-plane. During a single circular traversal, a hand reaches exactly one minimum and maximum y and z position. Therefore, we plotted the smoothed-out the z and y positions for each wrist and computed the count of peaks and lows for each, dividing by two to obtain number of repetitions per wrist. Figures () shows the curve after smoothing (AC FIGURE SMOOTH). (While it would have been computationally simpler to consider only one of the dimensionsy or zfor the purposes of counting, such estimation could be easily deceived by up-down or front-back movement instead of the proper circular one.) We then took the average of the per-wrist repetition counts to compute the overall rate for this exercise.

Form Measurement
In addition to counts, we developed ways to compute form metrics for select three exercisesjumping jacks, high marches, and squat jumps. We describe our methodology below.

Jumping Jacks
A proper form for jumping jacks demands straight knees throughout the exercise, wide leg spread in the jumping position, and complete retreat to the starting position where feet are touching. We noticed that all of these characteristics closely correlate with the magnitude of the difference between a participants right and left feet x positions. For instance, a participant cannot achieve as large a spread in the jumping position if his knees are not straight compared to when his knees are locked. In addition, the differences will be smaller if the participant fails to retreat completely to the starting position before performing another jump. Accordingly, we used the average difference of the x coordinates of a participants left and right feet, normalized by the his height, to gauge relative form between participants.

High Marches
To perform high marches properly, a participant needs to kick up as high as possible while keeping her knees straight. Unfortunately, the knee data have shown to be too noisy to make fine deductions such as the relative straightness of knees. However, we observed that bent knees necessarily lead to lower feet y-position during kick-ups compared to when the knees are straight. Therefore, we used the average foot height per each rep, normalized by the participants height, to compute the relative form metric for high marches.

Squat Jumps
Squat jumps demand that the participant jumps as high as possible while bringing his or her knees as close to his or her chest as possible. Accordingly, to compute a metric that could be used to compare squat jump form across multiple participants, we took the average in-jump knee height for each participant, normalized by height.

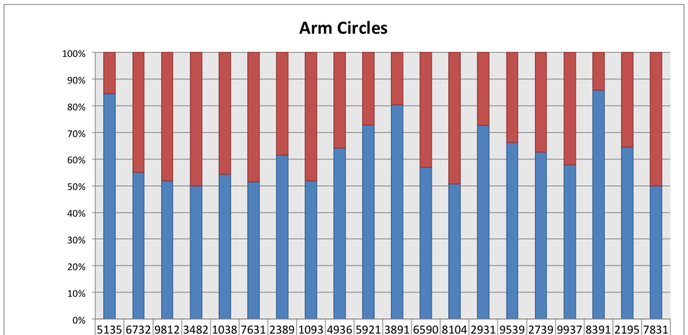
Feedback
After completion of the exercise routine, we presented each participant with two types of feedback. First, a holistic one, which consisted of two percentages summarizing how the participant performed relative to our expert benchmark: rate percentage alone and rate percentage after accounting for form. Second, a granular one comprising a detailed output of our data processing, including graphs and rates for each exercise. We then asked which type of feedback the participant would find most helpful as a means of tracking progress and improvement over time, and recorded their response. sp

Participants
We recruited 23 people total, three participants completed the first round of user testing in which the main objective was to perfect our data analysis models and 20 participants completed the final round of user testing from which we collected our main data set. Of the 20 participants who completed the study, all were undergraduate students from the age of 19 to 21. 8 of the participants are female and 12 are male.
Our participants were on varying levels of athleticism and at varying fitness levels, however, we did not ask participants to identify their perceived athletic and fitness level as we did not want to bias our subjective results nor did we trust that a self identification process would be an accurate depiction of actual ability and stamina.

RESULTS

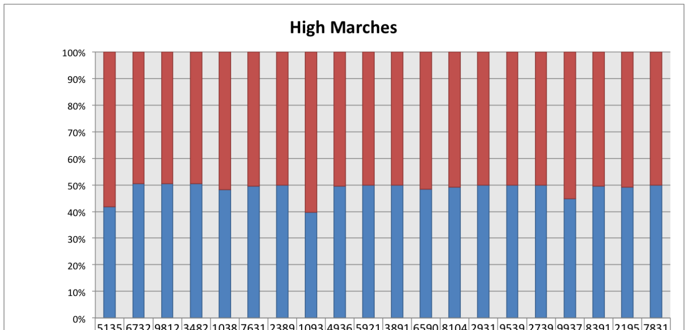
Quantitative Results
To measure quantitative accuracy of our model, we took a video recording of each of our participants. We then manually counted the number of repetitions for each of the exercises in our routine and compared them to the count estimates given by our model. The figures below show our results as 100% stacked column charts. If our model performed accurately, both the Actual and the Estimated bars would account for 50% each. If the Estimated bar comprises less than 50%, it means our model underestimated the total count; correspondingly, if the Estimated bar takes up over 50%, it means our model overestimated. For reference, we also included a table detailing the counts under each of the graphs.
Clearly, our model gave more accurate count estimates for certain exercise (such as jumping jacks) and less accurate estimates for others (such as arm circles). Another observation we can make is that the bias in our model is not random but

tends to either overestimate most of the time or underestimate most of the time. A discussion of the possible explanations for the performance disparities and inaccuracies is left for the Discussion section.



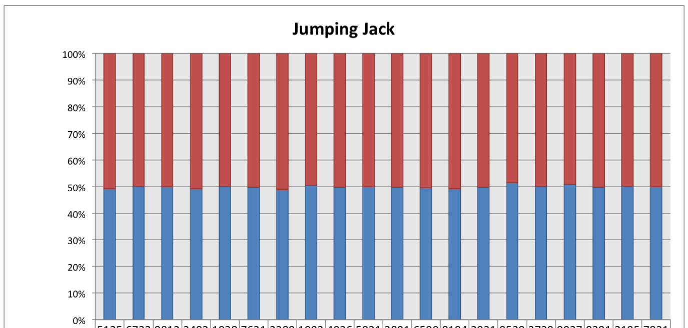
	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	43	94	119	154	112	127	96	105	84	91	48	112	105	84	84	91	95	44	90	138
Actual	233	115	128	153	133	135	153	113	150	244	196	148	108	223	164	152	130	267	163	138

Figure 2. Arm Circles Results



	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	46	58	61	43	42	51	59	47	52	76	53	62	61	57	52	25	43	55	33	53
Actual	33	59	62	44	39	50	59	31	51	76	53	58	59	57	52	25	35	54	32	53

Figure 3. High Marches Results

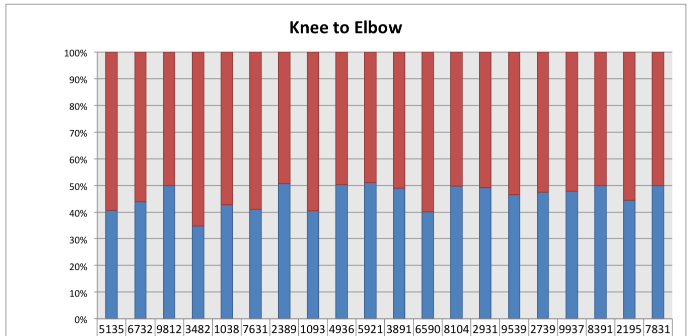


	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	132	130	126	139	128	146	149	120	138	178	101	119	103	140	148	143	119	130	129	143
Actual	128	131	126	134	129	144	142	122	137	178	100	117	100	139	157	144	123	129	130	143

Figure 4. Jumping Jacks Results

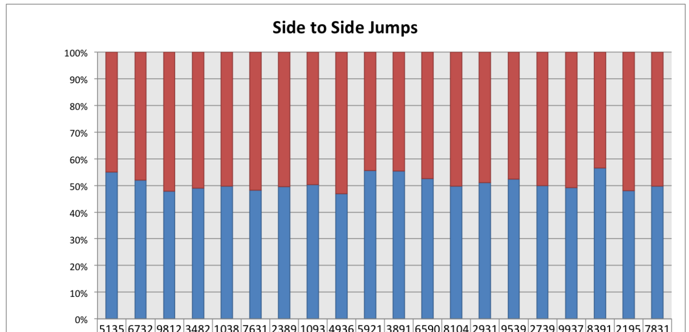
Qualitative Results

In order to measure the qualitative accuracy of our performance models, we enlisted the help of 3 experts in the field of fitness that are currently employeed by the Harvard athletic department. One of the most important goals of our project was to succeed in terms of relative performance ranking, i.e. the expert score percentage was not as important as the ability to successfully compare the form and performance of participants in relation to each other. For each of the four performances regarding form (Jumping jacks, high marches, squat jumps, and overall performance adjusted by form), the 20 participants were sorted in order of descending performance score. Groupings were then formed based on rank and the percentage difference in ranking. The three ranking pairs choosen were the following:



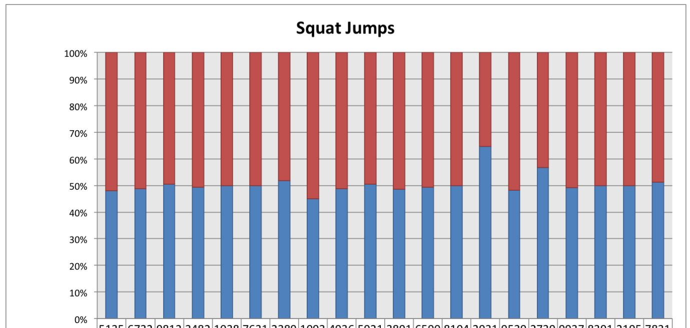
	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	117	88	111	109	91	86	79	72	63	100	77	94	80	97	93	72	83	113	96	70
Actual	80	69	111	58	68	60	81	49	64	104	74	63	79	94	81	65	76	113	77	70

Figure 5. Knee to Elbow Results



	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	102	110	138	134	120	144	136	116	142	112	96	124	112	152	116	128	128	94	132	142
Actual	125	119	127	129	119	134	134	117	125	140	119	137	111	158	128	128	124	122	122	141

Figure 6. Side to Side Results



	5135	6732	9812	3482	1038	7631	2389	1093	4936	5921	3891	6590	8104	2931	9539	2739	9937	8391	2195	7831
Estimated	39	41	47	39	48	39	38	38	41	57	35	45	37	29	44	29	32	47	26	41
Actual	36	39	48	38	48	39	41	31	39	58	33	44	37	53	41	38	31	47	26	43

Figure 7. Squat Jumps Results

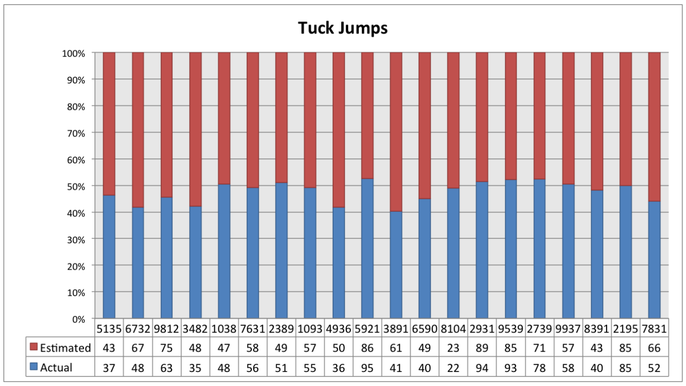


Figure 8. Tuck Jumps Results

- 1st and 20th, 95% difference in ranking
- 5th and 16th, 55% difference in ranking
- 9th and 12th, 15% difference in ranking

For the three specific exercise moves, the experts were shown 30 seconds of each of the two participants performing the move in the ranking pair and asked to indicate which participant had the better form for that specific move. In order to judge the best performance adjusted for form, the experts were shown 30 seconds of each of the participants performing the moves not specified (Arm circles, Knee to Elbows, Side to Sides, and Tuck Jumps) and then asked to indicate which participant had the better form overall based on what they were shown. The following four tables demonstrate the results from the expert session in order of Jumping Jacks, High Marches, Squat Jumps, and Overall Form respectively.

Table 1. Jumping Jacks Results

Ranking Pair	Expert #1	Expert #2	Expert #3
1 and 20	20	1	20
5 and 16	16	5	5
9 and 12	9	9	9
# Correct (out of 3)	0	3	2

Table 2. High Marches Results

Ranking Pair	Expert #1	Expert #2	Expert #3
1 and 20	1	1	1
5 and 16	5	5	5
9 and 12	12	12	12
# Correct (out of 3)	2	2	2

Feedback Results

When presented with the two different forms of feedback (design A being the percentage expert score feedback design and design B being the granular workout move specific design), 15 participants (75%) indicated that they preferred design B to track their performance from workout to workout and 5 participants (25%) indicated that they preferred design A to track their performance. Some of the reasons given to prefer design B over design A include the approval of the inclusion of the specific workout details. Participants found it useful to see the areas in which they needed the most improvement as they were unsure of the specifics of their form. However, one participant indicated that she preferred design A because she was already cognizant of her form and would prefer a tangible number to serve as an incentive for improvement. Three participants indicated that they would have preferred a synthesis of the two reports and found much of the information portrayed in design B to excessive, however they still preferred design B as design A did not include enough information for them to accurately reflect on their performance.

Table 3. Squat Jumps Results

Ranking Pair	Expert #1	Expert #2	Expert #3
1 and 20	1	1	20
5 and 16	5	5	5
9 and 12	9	12	9
# Correct (out of 3)	3	2	2

Table 4. Overall Form Results

Ranking Pair	Expert #1	Expert #2	Expert #3
1 and 20	1	1	1
5 and 16	5	5	5
9 and 12	9	9	9
# Correct (out of 3)	3	3	3

DISCUSSION

TBA

CONCLUSION

TBA

ACKNOWLEDGMENTS

We would like to extend our gratitude to the following people serving as our fitness experts: Luke Walczewski (Assistant Director of Athletics at Harvard University), Beth Zeitlin (Strength and Conditioning Coach at Harvard University), and Phil Thomson (General Manager at the Malkin Athletic Center at Harvard University). We would also like to thank our participants for donating their time to participate in the study and providing us with our rich data set. Lastly, we would like to thank Professor Krzysztof Gajos for acting as our mentor during the entirety of the project.

REFERENCES FORMAT

References must be the same font size as other body text.

REFERENCES

1. Consolvo, S., Klasnja, P., McDonald, D. W., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., and Landay, J. A. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing*, UbiComp '08, ACM (New York, NY, USA, 2008), 54–63.

2. Consolvo, S., McDonald, D. W., and Landay, J. A. Theory-driven design strategies for technologies that support behavior change in everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, ACM (New York, NY, USA, 2009), 405–414.
3. Noland, M. P. The effects of self-monitoring and reinforcement on exercise adherence. *Research Quarterly for Exercise and Sport 60*, 3 (1989), 216–224.