

# the **CrossFit** JOURNAL ARTICLES

## The Quest to Measure Fitness

Tony Budding



The question of defining fitness has been a challenge for the fitness industry since its inception. All the established certifying agencies have so far failed to produce a clear working definition. CrossFit's 2002 manifesto "[What is Fitness?](#)" was a groundbreaking contribution in both specificity and scope. In it, founder Greg Glassman described four models of objective standards for defining fitness.



Glassman since has more concisely [defined](#) CrossFit's results, and its definition of improvements in fitness, as "increased work capacity across broad time and modal

domains." This single definition encompasses all four of his original models. The ability to accomplish all kinds (modes) of real work, at all time domains (from a few seconds of heavy or explosive effort all the way up to sustained activity for many hours) requires capacity in all ten physiological adaptations to exercise, all three metabolic pathways, and all the various hopper configurations; and will move you away from sickness on the continuum toward wellness and fitness. (See "What Is Fitness?" for further explanation of each of the four standards.)

1 of 5

Work capacity has two components: work and time. This is real work as measured by the accomplishment of some observable task. But fitness is not just about work *per se*. Increased fitness means having the ability both to get more work done in less time and to sustain high levels of work output for longer periods of time. In “Fran,” the task is to complete three rounds of thrusters and pull-ups as fast as possible. As the time it takes you to complete that work decreases, you are becoming inarguably fitter (at least for that task). In a real-world scenario such as filling and moving sandbags to a levee in a flood, you are fitter when you can move the same number of sandbags in a shorter time (or more bags in the same or shorter time), period. Fitness, in this case, is simply the capacity to move more sandbags faster.



### Average power output

It makes sense to want to quantify this. The 2005 [CFJ](#) article “[Fooling Around with Fran](#)” reported the measured work (in total ft-lbs) accomplished by an athlete doing “Fran” divided by the time it took him to finish the task. Work divided by time is average power (ft-lbs/min). Given a constant workload (95-pound thrusters and a constant height and body weight), the

faster the athlete’s time, the greater his average power output. And, as he trains, every resulting reduction in his “Fran” time is a measurable increase in his average power, and thus a quantifiable improvement in his fitness for that task.

We could theoretically perform this type of calculation with a wide variety of tasks and a wide variety of time domains and then take an average across those data points. For example, in the sub-ten-second realm, we can measure the power generated in a few max efforts (deadlift, squat, snatch, clean, for example), a few pulls on the Concept2 rower, a 40-yard dash, a hammer throw, a long jump, etc. The more, and more widely varied, the tasks, the more representational the average power calculation will be. In the three-minute realm, we use tests like “Fran,” “Diane,” max burpees, a row, a sandbag carry, a stair climb, etc. At efforts of about twenty minutes’ duration, we can measure results at “Cindy,” “Linda,” “Filthy Fifty,” shoveling a driveway full of snow, spreading gravel, etc.

This can be repeated across all time domains. The point is to use as wide a variety of functional movements (non-functional movements are simply less relevant) at each time domain. Plotting the results on a graph yields a quantitative measure of their power across broad time and modal domains. Power output will be highest in the shortest time domains, and it will decrease in descending degrees as the time increases. The integration of this curve is work, and the greater the area under the curve for an individual, the greater his or her fitness.

What would this graph mean, exactly? First of all, it is a theoretical experiment only. It is technically impossible to test an athlete’s current output across broad time and modal domains. Each measurement is a max effort, so the athlete can only perform one at a time. Some period of recovery is required between measurements and, in that recovery, there will be some adaptation. The more measurements you do, the more adaptation will occur, and the less you are measuring the same athlete.

But even if it’s only theoretical, this graph still represents a valid concept. Each athlete does in fact have a potentially measurable output level at various time and modal domains, even if you can’t accurately determine it in the real world. The purpose of CrossFit, and of all fitness programs for that matter, is to increase capacity



at each time domain. Each increase is a measurable and observable improvement in fitness. And the fitness program that most increases an athlete's performance in the least time is the most effective.

From a practical standpoint, though, the CrossFit workouts are an outstanding way to measure this. Both the main site WODs (the Workouts of the Day on CrossFit.com) and the [benchmark workouts](#) employ a wide variety of movements over a wide variety of time domains. If your times are dropping and your loads increasing, your power output is increasing and you're getting fitter. This is indisputable.

### Capacity

An important question remains however: How relevant is this fitness? In other words, how much do improvements in a broad ability to generate and sustain power across broad time and modal domains within CrossFit translate into real, improved capacity in the world outside the gym? Does the capacity developed in CrossFit improve capacity in the real world?

For longtime CrossFitters, the answer is an obvious yes. It's why we train so hard. Sure, there's often a love of fitness for its own sake, but there's little doubt the extreme growth of the program stems primarily from the real and undeniable benefit these workouts have in terms of real-world results. And one question remains to be answered: Can that improvement ever be scientifically established and measured?

Why *are* the workouts so effective? First and foremost, they use functional movements. Functional movements are natural, universal motor recruitment patterns. They are elemental, in that they are both irreducible and the building blocks of all human movement. They move the body through a full range of motion, both with and without external loads. And, they have the unique ability to move the largest loads, the longest distances, the quickest. In the thruster, for example, the knee, hip, shoulder, and elbow joints are all taken through a full range of motion (and a whole host of stabilizing musculature throughout the body is engaged), and significant amounts of average power are being generated. A 200-pound male can be capable of generating one full horsepower (33,000 ft-



lbs/min) for a minute or more with a 100-pound barbell. This kind of power is impossible with any isolation or non-functional movement.

It's common sense to conclude that someone who can generate high power output through a full range of motion in a wide variety of movements has broad physical capabilities. In fact, what more could you ask of a GPP (general physical preparedness) fitness program than the ability to generate power at a wide variety of functional movements over a wide variety of time domains?

We can term this ability "capacity." The question remains, though: How relevant is this capacity to the activities of life and to what is usually called "fitness"? All of life's activities require sets of *specific* skills. But CrossFit purposely develops *general* capacity. Does the capacity developed in high-repetition cleans, for example, lead to capacity in filling and moving sandbags to a levee in a flood? What about to bow-hunting caribou in the snow? Or to motocross racing? To combat? To firefighting? To maintaining independence into old age?

The answer is a limited yes. Limited because all these activities require specific forms of work using the same

joints through at least a portion of the same range of motion. However, each specific task requires a unique combination of coordination, accuracy, agility, and balance, at least—if not also strength, stamina, endurance, flexibility, power and speed. Fitness is thus the broad capacity that allows for the ready development of skills specific to a given task.

### The (ir)relevance of horsepower

Horsepower (ft-lbs/min) is a measure of capacity. It is the real work performed divided by the amount of time required to perform it. Improved fitness is identical to increased capacity. But how useful is the specific calculation of horsepower?

On the one hand, calculating horsepower is substantially more relevant to real life than the biological markers traditionally used in the fitness industry, such as  $VO_2$  max; resting, maximal, and recovery heart rates; and percent body fat. Generally, gains in average power by an individual correspond to increases in that person's real-world capacity, whereas changes in the biological markers often occur without corresponding changes in the capacity to get real work done (particularly across

broad time and modal domains). One widely publicized example of this is the complete absence of correlation between  $VO_2$  max levels and finishing order at the Tour de France.

But, on the other hand, high horsepower generation in itself is not the goal either. If the real world is the arena that matters to us, we have to use the ability to perform real physical work as the measurement of fitness. Let's look again at our sandbag example. Improved capacity for this task allows us to move more sandbags, but not necessarily to generate more power. The two measurements are correlated but not identical. In general, the person moving the most sandbags is the one generating the highest average power. But that's not necessarily true. You could have a smaller man moving more sandbags than a larger one and he could still have a lower total power output (ft-lbs/min), since moving bodyweight would be part of the equation.

Furthermore, because a significant amount of the work being performed in this example is moving your own body, you could gain ten pounds and move the same number of sandbags, thereby generating more calculated power but with no improvement in the task. Conversely, if you lost a bunch of weight but could still move more sandbags, you could end up reducing your mathematical average power, even though you accomplished more real-world work and made a greater contribution to stanching the floodwaters. These examples are aberrations, for sure, but they are important because they highlight a limitation of calculating output. The same principle is true in auto racing as well: the vehicle that has the highest horsepower is not necessarily the winning one. More horsepower certainly helps, but it's neither the goal nor the whole story.

The breadth of human activity is immense. CrossFit programming is the result of decades of study of the best movement and greatest capacity known combined with millions of workouts performed. Clearly, performing a wide variety of functional movements at high intensity through a full range of motion does indeed develop a broad, general, and inclusive capacity that provides a ready state for learning and improving specific skills and abilities. Precisely measuring and quantifying that fitness may or may not be possible.

### Who is fittest?

The human mind is comparative by nature. We want to rank things. CrossFit claims to be the greatest fitness program in history because it generates more capacity across broader time and modal domains than any other. This is a bold claim, but so far no one has really challenged it.

The next [CrossFit Games](#) is happening a few days after the publication of this article. The format of the Games will be different this year than last. Next year will probably be different again. The Games claims to be a legitimate test of capacity across broad time and modal domains. But can any two-day event accurately test broad capacity? I think so. Can success in any specific event predict success in any other specific event? I think so. Can the top athlete of the best fitness program in history claim to be the fittest person on the planet? I think so.

**Tony Budding** is the Media Director for CrossFit HQ.