Karen Coyne

INTRODUCTION

The need to protect workers from the inhalation of airborne contaminants has been recognized for many centuries. In 77 AD, Pliny the Elder wrote about red lead refiners wearing animal bladders to avoid breathing the lead dust (Roach, 1992). People such as Leonardo da Vinci (1452 - 1519) and Bernardino Ramazzini (1633 - 1714) recognized also the need for respiratory protection (Rajhans and Blackwell, 1985). However, it wasn't until the 1800s and the industrial revolution that significant advances were made. In 1814, the "precursor to the modern day air-purifying respirator was developed" and in 1825, John Roberts developed a smoke filter for firefighters (Rajhans and Blackwell, 1985). In 1910, the Mine Enforcement Safety Administration (MESA), the predecessor of the Mine Safety and Health Administration (MSHA), began specifying regulations about the design and certification of different respiratory protection sold in the United States (Teresinski and Cheremisinoff, 1983). Design progressed rapidly during WWI when toxic gases were first used as a military weapon (Rajhans and Blackwell, 1985).

In 1970, the concern for worker health came to the forefront. Former Labor Secretary Schultz testified before Congress that 14,500 Americans died and 2.2 million workers were disabled due to industrial accidents each year and, the U. S. Public Health Service stated that there were approximately 390,000 new cases of occupational diseases each year (Wang, 1993). The total monetary cost to the

American public was estimated at $8 billion annually. Due to the large numbers of workers killed or injured in industrial accidents every year, the Williams and Steiger Occupational Safety and Health Act (OSHA) was enacted in 1970. OSHA made employers responsible for the safety and health of their workers in the workplace. Engineering controls, such as increased ventilation, should be used first in protecting against the health risks from hazardous substances in the workplace. When these controls fail or are not technically feasible, personal protective equipment becomes necessary.

Respirator masks are an essential component of the personal protective equipment and are used to protect workers against the inhalation of various contaminants - dust, mist, vapor, gas, and fume – that are found in the manufacture of chemicals, automobiles, steel, batteries, furniture, adhesives, and many other products. Additionally, there are individuals in small factories, offices and laboratories who are exposed to hazardous substances. Painters, soldiers, firefighters, miners, wood workers, construction workers, asbestos removal personnel and others must wear these masks. These workers perform activities of a physical nature at varying intensities while wearing respirators.

Respirator design currently involves making a prototype and then testing it on humans. Adjustments to the respirator are made based on those tests and then a new prototype is made and is tested. This process continues until an adequate respirator is developed. A model that predicts the effects of a respirator on a person would allow respirator design to proceed more rapidly. Such a model would be an important design tool that would provide valuable information on the potential physiological and psychological compatibility of a respirator with the wearer. The model would not eliminate the need for human testing, but would decrease the number of prototypes and testing required. Much time and money could be saved.

There are thermal, metabolic, cardiovascular, respiratory, and psychological effects of respirator wear that need to be considered. Information on these effects is found in many different sources. A model would bring this information together and quantify these effects. The development of the model should also indicate areas where more information is necessary.

A successful model would be very complex because of the many factors to consider. And, because of the variability of human response to exercise, work, and respirator wear, the initial development of the model will include many assumptions and this may limit the expected accuracy of the predictions. As more research is done that quantifies the effects of respirators on humans, this information should be included in the model.

The purpose of this research was to develop a model that examined the effects of a respiratory protective mask on the pulmonary system during constant-rate exercise. This model could form the foundation for the larger model. If the intensity is not severe, constant-rate exercise will eventually result in a physiological steady-state (Wasserman et al. 1967; Poole and Richardson, 1997). Although a steady-state may not be possible physiologically, parameter values may still be determined because these steady-state values will determine the rate of rise of the parameter and will be important when transient effects are included (Givoni and Goldman, 1972).

REVIEW OF LITERATURE

When investigating the behavior of large-scale biological systems, it is often difficult to determine the effect changes in the system parameters have on the overall system. This difficulty may be due to the scale of the system or to problems collecting data. To overcome these problems, a mathematical model of the system may be developed. “[Mathematical models] provide a concise description of complex dynamic processes, indicate ways in which improved experimental design could be achieved and enable hypotheses to be tested (Finkelstein and Carson, 1985).” This approach has become more common in recent times due to the increase in the computational power of computers and the use of the systems approach to problem solving (Murthy et al., 1990).

A model is a representation of a system in the real world. This system is analyzed to determine the important components and interactions between these components. These observations are then translated into a set of mathematical equations that describe the relationships between a system’s behavior and its properties (Finkelstein and Carson, 1985). The resultant model is only an approximation of the whole system. The degree to which the model corresponds to the real-world system will depend on the purpose for which the model is designed. If a great degree of accuracy is required, the model necessarily becomes more complex and subsequently more difficult to evaluate. A less complex model would be simpler to evaluate, but would contain less information.

Developing mathematical models is not just a science, it is an art as well (Finkelstein and Carson, 1985; Murthy et al., 1990). Science is evident in the principles and equations used to formulate the model. However, artistic aspects such as creativity, ingenuity, intuition, and foresight are needed to make the model more than just a group of related equations. Because of the degree of personal choice in specifying a model, no two models will be the same.

Model Development

Model development depends in part on the type of model being used. Mathematical models may be classified as either empirical or theoretical, although there may be an overlap between the two (Murthy et al., 1990; Shirmohammadi et al., 2001). A theoretical model results when well-established theories are used in determining the equations for a model. These models are called also physical or mechanistic models because they are based on the physical system. When the modeler fits equations to a set of data without considering the theory behind the relationship, an empirical model results. However, even when an empirical model is developed, it is important that the model not contradict established theory. So, an empirical model does have some theoretical basis. Theoretical or physical models have a broader application than empirical models because the theoretical models are not based on any one data set (Shirmohammadi et al, 2001).

In developing a model, it is important that a systematic approach be used. While various authors (Finkelstein and Carson, 1985; Hunt, 1999; McCuen, 1993; Murthy et al., 1990) use different nomenclature to describe the modeling process, the approach should involve the following steps: problem formulation, factor specification, data collection, assumption making, system characterization and mathematical description, model formulation, model calibration, and model validation. Because each of the stages is interrelated, the overall process is inherently iterative (Finkelstein and Carson, 1985).

The techniques of aggregation, abstraction, and idealization must be employed during each stage of model development (Finkelstein and Carson, 1985). Aggregation involves grouping many common objects into one composite object. For example, the resistances of the arteries in the leg may be considered as a circuit of many single resistances or as one lumped equivalent resistance. The choice would depend on the intended use of the model. Abstraction concerns the "degree to which only certain aspects of a system are included in a model (Finkelstein and Carson, 1985)." For instance, a model of stream health may include industrial pollution but not surface runoff. Approximation of system characteristics, or idealization, is also performed. An example of idealization would be assuming that all gases in a system are mixed instantaneously, even though this takes some finite time to occur.

Problem Formulation

The problem formulation stage involves determining the objectives or purpose and scope of the model. It is important that the purpose be stated explicitly with as much detail as possible because the form of the model will depend on the purpose. "Thus the form of a model which is simply being used to describe some experimental test data is unlikely to be the same as one used for examining alternative hypotheses regarding the precise quantitative nature of the chemical and neural control of breathing or as that used for predicting the growth of a dysmature infant in response to a particular regime of feeding (Finkelstein and Carson, 1985)."

Models may be developed to be descriptive, predictive, or explanatory (Finkelstein and Carson, 1985). Descriptive models attempt to find relationships between data. An example would be determining the equation relating the change in heart rate at increasing levels of exercise to the work rate. Predictive models are used to determine how a system will respond to a stimulus or change in the system, for example to predict the response of a person to a new drug. Finally, explanatory models provide insight into "the ways in which different features of system behavior and structure depend upon each other (Finkelstein and Carson, 1985)." Many models are a combination of the three.

Factor Specification

At this stage it is important to list all the important factors in the model. Simplification and elimination of some factors will occur later. Factors can be classified into three categories (Edwards and Hamson, 1990): constants, parameters, and variables. Constants are factors that have fixed values (speed of light) and factors that are essentially the same in all cases of interest (acceleration due to gravity). Parameters have constant values for a particular problem but can change from problem to problem (Edwards and Hamson, 1990). In a fluid pumping model, the fluid density, the pipe diameters, and the pipe lengths would all be parameters. While these factors may vary from system to system, they are constant for the particular system being investigated. Variables will have values that change throughout the model. For the fluid pumping system, the velocity of the flow in the pipes would be a variable because its value will change depending on factors such as the pipe diameter.

After listing all the possible factors, it is useful to group related factors together (Edwards and Hamson, 1990). This will help later when relationships between factors are formed. Each of the factors needs to be identified as a constant, parameter, or variable. Variables should then be divided into inputs and outputs. It is often easiest to first identify the constants and parameters and then the variables can be separated. To distinguish between input and output variables, it is helpful to look at the possible relationships between factors in each group. If a variable is a direct consequence of other variables, then that variable is an output (Edwards and Hamson, 1990). If a variable's value is independent of all other variables, then that variable is an input to the model. To complete the list, each factor should be assigned a variable name and units.

Data Collection

Data concerning and knowledge about the various factors involved in the model must be obtained. This information helps to define the scope of the model and may also cause the objectives to be altered if, for instance, there is not enough information available. The required data may be available from various reference sources or new experiments may need to be conducted to obtain the data.

Data are necessary for many stages of the modeling process. Plotted data can give insight into the mathematical form of a model or part of the model. Data are used in the calibration stage to approximate model parameters. They are used also in the validation stage to determine whether the model results agree adequately with real situations.

System Characterization and Mathematical Description

Because the model is only an approximation of the actual system, the modeler must decide which “features or characteristics of the system are relevant and significant for the goal in mind (Murthy et al, 1990).” The systems approach requires first a functional and then a mathematical description of the biological processes and systems involved. The degree of detail included in a model is a compromise and is a part of the art aspect of modeling (Murthy et al., 1990). Including too much detail results in a cumbersome model, while having too little detail gives an incomplete model.

The relationships between the factors must next be specified. This involves deriving equations based on the gathered data. In many cases, such equations already exist. The modeler then must choose which equations fit the particular problem. In later stages, it may be necessary to return to this point to either include more information or eliminate some factors.

The end result of this stage is a collection of equations describing the procedures and processes that characterize the system. This collection is still far from being a model. It is during the next stage that these equations are combined and formed into a model.

Model Formulation

An inductive, deductive, or pragmatic approach is used in formulating a mathematical model. The inductive approach involves observing system behavior and trying to model its characteristics. With this method, it is unlikely that the model parameters will have any physical significance. The deductive approach breaks a large system down into its component parts. Equations are developed for each of the parts and for the interaction between the parts. A model is then formed from this system of equations (Barreto and Lefevre, 1984). The engineering approach is frequently the pragmatic one. That is, the model is determined with a definite purpose in mind (Barreto and Lefevre, 1984). Physiological models typically use the deductive method because of the need to understand each of the parts and its relationship to the whole system.

These approaches lead to empirical and theoretical models, or to combinations of the two. An empirical model results when an inductive approach is used. These models are typically viewed as “black boxes” because the resultant model is based only on the data, not on any theory or knowledge about the system. Empirical models are generally used only for descriptive purposes.

The deductive approach leads to a theoretical model. This type of model is based on *a priori* knowledge about the system’s structure and function. These models can be used for descriptive or predictive purposes.

If the deductive approach is used, it is necessary to couple together the individual equations determined in the system characterization stage. This process is not as simple as connecting the equations together. Care must be taken that the resultant model is not redundant and does not contain any incompatibilities such as two voltage sources connected in parallel (Barreto and Lefevre, 1984). Once the model has been formed by relating the equations, the model must then be evaluated.

Calibration

The next step in the model development is to calibrate the model. This involves fitting the model to the data by adjusting the coefficients of the predictor, or independent, variables, so that accurate model output is obtained. The values of the coefficients that give the best agreement between the model output and collected data are considered the optimal values (McCuen, 1993). In addition to the model itself, an objective function and a set of measured data are needed to calibrate the model. The objective function is an explicit mathematical function that specifies the optimal solution. Often, the least squares fit of a model is used as the objective function.

Not all of the data should be used for calibrating the model. Some of the data should be saved for the next stage, model validation.

Validation

Validation consists of assessing whether the model is accurate and achieves the purpose for which it was designed. It is not possible to verify a model. “[Models] are essentially hypotheses, which are tested by subjecting them to crucial experiments designed to falsify them and they are accepted to the extent that they are not falsified. (Finkelstein and Carson, 1985).” Validity concerns not just the final output, but the purpose, current theories, experimental test data, and other relevant knowledge (Finkelstein and Carson, 1985). When new theories are accepted and more experimental data are obtained, the model must be validated again.

Validation of the model should take place throughout the development of the model, not just at the end. If any validation assessment indicates errors or inaccuracies in the model, it is necessary to return to the system characterization and formulation stages to make changes. It may even be necessary to modify the initial conceptual model (Finkelstein and Carson, 1985).

If it is not possible to validate the completed model, then model reduction must be used. This process begins by reviewing the initial conceptual model. Systematic model reduction is then accomplished by making assumptions based on physiological and mathematical principles (Finkelstein and Carson, 1985). Although it may seem better to start with this simplified model, “there is the danger, particularly if the model is formulated simply on the basis of conforming to test response data, that it will lack physiological realism (Finkelstein and Carson, 1985).”

Determining the level of acceptance of the model and the degree to which the model replicates experimental data is subjective and often determined by the model purpose. Specifying the validation criteria explicitly will reduce this subjectivity (Cobelli et al., 1984).

The validity of a model is assessed using both internal and external criteria (Finkelstein and Carson, 1985). Internal criteria include consistency and algorithmic validity. The model is considered to be consistent if it does not have any mathematical, logical, or conceptual contradictions (Finkelstein and Carson, 1985; Cobelli et al., 1984). Algorithmic validity requires that the algorithm be appropriate for the model and that it lead to accurate and logical solutions (Finkelstein and Carson, 1984; Cobelli et al., 1984).

External criteria include empirical, theoretical, pragmatic, and heuristic validity. Empirical and theoretical validity concern current knowledge. The model is empirically correct if it agrees with experimental data and is theoretically correct if it follows currently accepted theories. Pragmatic validity assesses whether or not the objectives of the model have been met. Heuristic validity concerns determining the "potential of the model for scientific explanation, discovery, and hypothesis testing (Finkelstein and Carson, 1985)."

Methods of Validation

No model should be used before it has been validated thoroughly. Validation consists of assessing whether the model is accurate and achieves the purpose for which it was designed. The model is subjected to input data over the range expected in the physical system to ensure that rational output is obtained. However, it is not possible to verify a model. The model is accepted to the extent that it cannot be proven incorrect.

Both qualitative and quantitative methods are used to assess empirical and theoretical validity. Care must be taken when using any validation method. No single method should be used to determine validity. A combination of qualitative and quantitative methods should be performed with the results being used in conjunction with knowledge, experience, and common sense to determine the validity of the model.

***Qualitative Analysis*.** Qualitative assessment consists primarily of observing the output response and comparing it to the expected response. Such parameters as magnitude and sign of the output should be checked to determine if they are physiologically reasonable. Trends in the data, such as expected increases and decreases in the output should also be checked.

***Quantitative Analysis*.** Quantitative evaluation generally involves goodness-of-fit tests to determine how closely the model output agrees with experimental data. The correlation coefficient, modified correlation coefficient, and standard error of estimate can all be used. In many cases, including time-dependent models, these goodness-of-fit criteria should be considered goodness-of-fit indices and not statistical measures because the underlying statistical assumptions, such as independent observations of the data, do not hold. The indices are still measures of variance, but "they should not be used with standard tests of significance (McCuen, 1993)."

However, Finkelstein and Carson (1985) argue that "due to the considerable physiological variation within the human population and the errors involved in measurements on the cardiovascular system, it is not appropriate to use integral of error squared or other similar performance criteria in the comparison of this model with the real cardiovascular system." They recommend feature matching of the principle responses as the primary validation procedure. Murthy et al. (1990) state that goodness of fit tests can be used if they are adapted to the particular evaluation and have suggested specifying individual indices for each part of the model to be validated.

The approach of McCuen (1993) is more practical. Seven criteria are described that should be considered when assessing a model's reliability; not all seven should be used with all models. These criteria are coefficient rationality, meeting the assumptions of the model, standard error of the estimate, correlation coefficient, model and relative bias, accuracy of fitted coefficients, and the analysis of variance (McCuen, 1993).

Model rationality concerns both whether the output is reasonable and whether the coefficients provide an accurate relationship between the predictor and criterion variables. All coefficients should be rational in sign and magnitude. The intercept coefficient has the same units as the dependent variable so its rationality can be assessed directly. However, slope coefficients have units that are a function of both the independent variable and the dependent variable. Slope coefficients may be converted to dimensionless standardized partial regression coefficients:

$$t_i = \frac{b_i S_i}{S_y} \tag{1}$$

where: b_i is the slope coefficient

S_i is the standard deviation of predictor variable i

S_y is the standard deviation of the criterion variable.

A standardized partial regression coefficient has an absolute value between one and zero, with one indicating an important predictor variable. If the absolute value exceeds one, then intercorrelations are significant and the coefficient is irrational. McCuen (1993) stated that an irrational model should be used with caution and should not be used beyond the range over which it was developed.

The model bias is found by summing the differences between the model and experimental values. A positive bias means that the model consistently overestimates, while a negative bias indicates the opposite. Small biases are tolerable if other criteria are met. The t-test can be used to determine if model bias is significantly different from zero.

The standard deviation is a measure of the spread of the data and the accuracy of the mean. To reduce the error variance, the criterion variable is related to the predictor variables. The goal is to provide an unbiased relationship that has a minimum sum square of errors. The error variance is the sum square of errors divided by the degrees of freedom. The standard error of estimate is the square root of the error variance. If the S_e is less than the S_y of the population, then the model provides a better estimate of the criterion variable than the mean. The ratio, S_e/S_y, is used to determine if any improvement has occurred. If the ratio is near zero, a significant improvement has occurred. Conversely, if the ratio nears one, no improvement has occurred.

The correlation coefficient is a measure of the degree of the relationship between a criterion and predictor variable; it does not specify the relationship. The square of the correlation coefficient is a measure of the amount of variance of the criterion variable explained by the predictor variable. McCuen (1993) states that the standard error of estimate is a better measure of goodness of fit than correlation coefficient because the standard error of estimate has the following advantages: it has the same units as the criterion variable, the degrees of freedom are accounted for properly, and it is valid for nonlinear and linear models.

Model coefficient accuracy can be assessed by examining the standard error of the regression coefficient. McCuen (1993) has found from experience that the coefficient is of questionable accuracy if the ratio $S_e(b_i)/b_i$ exceeds 0.3 to 0.4.

The sum of the residuals is examined to determine if there is a bias in the model. If the sum differs from zero, a bias exists. While R^2 is the amount of variation in the criterion variable explained by the predictor variable, the residuals are the variation not explained by the predictor variables.

The principle of least squares assumes a constant error variance. A plot of the residuals versus the independent variable should be obtained to determine if there is any pattern to the residuals. If a pattern exists then the residuals do not have a constant variance.

Respiratory System Background

The main function of the respiratory system is to provide oxygen to the tissues and remove carbon dioxide. This is accomplished through external and internal respiration. External respiration occurs in the lungs whereas internal respiration takes place at the tissue level. External respiration begins as the diaphragm and external intercostal muscles contract, expanding the chest cavity and creating a resultant pressure that is lower than atmospheric (Jensen and Schultz,1970). Due to the lower pressure inside the chest cavity, air rushes into the lungs to equalize pressure. Air is returned to the atmosphere with the subsequent relaxation of the diaphragm and intercostal muscles that increases the pressure within the chest cavity and forces the air out of the body. Thus, at rest, inhalation is considered active whereas exhalation is passive. During exercise, exhalation also becomes active requiring the internal intercostal and abdominal muscles to contract and further reduce the size of the thorax.

The air that is forced into the lungs first enters either through the nose or the mouth and then passes to the pharynx. From the pharynx, the air passes the larynx and enters into the trachea, the start of the tracheobronchial tree. From this point on, the air flow will divide among a set of dichotomously branching tubes in both the left and right lobes of the lung. At each branching, the diameter of the tubes becomes smaller, although the total cross-sectional area increases. From the original branchings off the trachea, the main stem bronchi, through the bronchioles, and into the terminal bronchioles, the air will eventually reach the alveoli. The alveoli are tiny, thin-walled sacs that lie among a bed of capillaries, small diameter blood conduits. It is in the alveoli that gas exchange with the blood occurs. The inspired air carries oxygen to the alveoli and the blood while the expired air carries carbon dioxide from the blood and delivers it to the atmosphere.

The respiratory muscles are controlled by respiratory centers located in the medulla, a part of the autonomic nervous system. As such, breathing is involuntary. An individual may hold his or her breath for a while, but eventually, the person will be forced to take a breath. Factors influencing the control of respiration include: muscular activity, emotions, carbon dioxide concentration, oxygen deficiency, and heart rate (Jensen and Schultz, 1970).

The amount of air that is inhaled or exhaled during each breath is termed the tidal volume. In an average, healthy, resting human, this value is approximately 500 mL (Johnson, 1991). The typical respiration rate of the same typical human is approximately 17 breaths per minute (Johnson, 1991). The minute volume, the amount of air inspired or expired in one minute, is the product of the tidal volume and the respiration rate.

Respiration and Physical Activity

Physical activity begins at some external work rate. This work rate requires a certain amount of internal or physiological work. The increased amount of oxygen required by the body is dependent on the physiological work rate. In response to the increased oxygen consumption, minute volume rises immediately. It then rises at a slower rate to a steady-state value (Johnson, 1991). The increase is exponential with a time constant of 65-75 seconds (Whipp, 1981). More capillaries open in the lung increasing the area for gas diffusion and thus the diffusing capacity of carbon dioxide and oxygen (Berne and Levy, 1988). At a constant moderate rate of exercise below the anaerobic threshold, the minute volume will level off at a steady-state value (Johnson, 1991). Above the anaerobic threshold, a steady state may not be achieved. Tidal volume and respiratory rate also increase. The inhalation and exhalation times shorten.

Wearing a respirator has been shown to affect the pulmonary response to exercise (Johnson et al., 1999). Hypoventilation can occur with a decreased oxygen consumption. The effects of the respirator need to be considered.

External Work

External work is the amount of mechanical work being accomplished. It is equal to the product of force and distance. Work rate, or power, is the work divided by the time to accomplish that work. Work is expressed in units of N·m while work rate is expressed in N·m/s, or Watts (W). So, the external work accomplished by a person with a mass of 70 kg who climbs a set of stairs (total distance: 3 m) is:

$$\mathrm{W}_{\mathrm{ext}} = (70\mathrm{kg})(9.8\frac{\mathrm{m}}{\mathrm{s}^2})(3\mathrm{m}) = 2058\ \mathrm{N} \cdot \mathrm{m} \tag{2}$$

The work rate would depend on how fast the person climbed the stairs. If the person took 3 seconds to ascend the stairs then the external work rate would be 686 W. Taking ten minutes to climb the stairs would result in an external work rate of 3.43 W. So, the time to accomplish the task is an important factor in how hard the person is working. Therefore, it is common to use external work rate instead of external work to make comparisons between activities.

Physiological studies often use activities where it is easy to determine the external work rate of a subject. These activities include walking or running on a treadmill, cranking an arm ergometer, pedaling a cycle ergometer, or stepping up and down a block. The work rate when using a bicycle ergometer is (Robergs and Roberts, 1997):

$$\mathrm{WR}_{\mathrm{ext}} = \frac{\mathrm{cadence} \cdot \mathrm{load} \cdot \frac{\mathrm{distance}}{\mathrm{revolution}} \cdot g}{60} \tag{3}$$

where: WR_{ext}, external work rate, W

cadence, rev/min

load, kg

distance/revolution, m

g is the acceleration due to gravity, m/s^2

60 is a conversion from min to sec

For a Body Guard or Monark ergometer, the distance/revolution is 6 m, while for a Tunturi it is 3 m (Robergs and Roberts, 1997).

The work rate of stepping (W) is:

$$WR_{ext} = h_{step} \cdot mass \cdot n_{step} \cdot g \qquad (4)$$

where: h_{step}, height of the step, m

mass, the mass of the person, kg

n_{step}, number of steps, dimensionless

g, acceleration due to gravity, m/s^2

The work rate of walking is more difficult to assess. In fact, Wasserman et al. (1999) stated that “probably the greatest disadvantage of the treadmill is the difficulty in quantifying the work rate.” The external work rate of walking or running on level ground is usually taken to be zero. It’s not that work is not being done. Work is done as the body is raised and lowered, but the positive and negative work are usually assumed to offset one another.

Webb et al. (1988) performed a study to determine if the work rate during walking was actually zero. Five male and five female subjects wore a suit calorimeter in a respiration chamber while walking on a level treadmill for 70 to 90 min at speeds of 0.69, 1.28, and 1.86 m/s. The suit calorimeter consisted of a mesh of water-filled tubes that covered the body. The amount of heat transferred to the water in the suit was determined. Subjects also pedaled a bike ergometer for 70 to 90 minutes against loads of 53 and 92 W. For cycling, the energy expenditure calculated from respiratory gas exchange equaled the heat produced plus the external work rate on the bike. However, the heat balance for walking showed that the energy expenditure did not equal the heat produced. This indicated that external work was done in walking. The amount of work done during walking increased with walking speed and was found to be an average of 12% of the transformed energy. The authors concluded that work was done bending the sole of the shoe and in other interactions between the foot and the treadmill surface.

The work of Webb et al. (1988) was continued by Nagle et al. (1990). These investigators had ten male subjects walk on a treadmill while wearing a suit calorimeter. Subjects walked at 1.5 m/s at grades of 10, 5, 0, -5, and 10%. Similar to their previous work (Webb et al., 1988), a non-thermal energy term was found at all grades. So, there is physical work done in grade walking as well as level walking that cannot be accounted for by external work or heat produced. This non-thermal term was significant at grades of 0, 5, and 10% but not at the negative grades. On average, this non-thermal energy term accounted for 6% of the transformed energy, which is half of that reported previously (Webb et al., 1988). While Webb et al. (1988)

proposed that the energy was expended in the compression of the heal of the shoe and in bending the sole, the current investigators offered a different explanation. They theorized that a portion of the energy externalized during the positive phase of walking is only partially recovered as heat energy during the negative phase (Nagle et al., 1990).

The external work done in level walking was investigated also by Snellen (1960). Three subjects walked on a level treadmill in a climatic chamber for one hour. The air and wall temperatures were kept close to skin surface temperature so that heat loss through radiation and convection was kept to a minimum. Heat lost through evaporation was calculated. The final heat balance showed that heat gained equaled heat lost. The investigator determined that level walking did not involve external work. It was noted in the article that there were errors in the measurements. Air and wall temperatures did not exactly match weighted skin temperature. Evaporation was determined through weight loss of the subject. Some of the water evaporated comes from the respiratory tract, but the heat of vaporization was determined at average skin temperature.

The different results obtained by Snellen (1960) and Webb et al. (1988) and Nagle et al. (1990) may be due to technique. Webb et al. (1988) and Nagle et al. (1990) used a suit calorimeter to measure the heat loss by the subject. As the external work rate represented 6% of the energy, it is possible that this non-thermal term was not seen in the study done by Snellen (1960) because of the errors involved in the calculations of heat loss and heat production. In fact, a study conducted by Johnson et al. (2001a) that investigated the heat production in level and grade walking found that there was a difference between the metabolic rate and heat production. While a calorimeter was not used, subjects were thermally insulated from the environment by clothing that consisted of light underwear, a neoprene wet suit, military fatigues, sneakers, sock, two pairs of gloves, a full-facepiece respirator mask, and a neoprene hood. This was done to decrease the heat loss by conduction and evaporation. So, when heat loss and heat gain are monitored carefully, it appears that there is indeed work done in level walking.

A number of approaches have been used to deal with the problem of determining external work during walking. Lakomy (1984) and Cheetham et al. (1986) have used an ergometer system that allows power to be determined during running. Givoni and Goldman (1972), Pandolf et al. (1977), and Aoyagi et al. (1995) provided equations for calculating external work rate. Other authors (Groot et al., 1994) have filmed various activities and determined the work performed.

A treadmill ergometer system was developed by Lakomy (1984). The subject ran on a non-motorized treadmill to which a small generator was attached. The generator gave a voltage proportional to the belt speed. A transducer was mounted at the back of the treadmill. The subject wore a harness around the waist that attached to the transducer. The harness held the subject in place and ensured that the force measured by the transducer was the same as the force applied horizontally on the belt. Instantaneous power was found from the treadmill speed and the force applied to the transducer. Similar types of systems have been used for rowing (Hagerman and Lee, 1971) and swimming (Toussaint et al., 1990).

Equations for external work were presented by Givoni and Goldman (1972) and Aoyagi et al. (1995). The equation provided by Givoni and Goldman (1972) was:

$$\mathrm{WR}_{\mathrm{ext}} = 0.098 \cdot \mathrm{m}_{\mathrm{t}} \cdot \mathrm{v} \cdot \mathrm{G} \tag{5}$$

where: $\mathrm{WR}_{\mathrm{ext}}$, external work rate, W

m_{t}, total mass, kg

v, velocity, m/s

G, grade, percent

The term 0.098 is the acceleration due to gravity divided by 100. So, equation (5) may be written as:

$$\mathrm{WR}_{\mathrm{ext}} = \mathrm{m}_{\mathrm{t}} \cdot g \cdot \mathrm{v} \cdot \frac{\mathrm{G}}{100} \tag{6}$$

The equation provided by Aoyagi et al. (1995) was:

$$\mathrm{WR}_{\mathrm{ext}} = \frac{3.6 \cdot \mathrm{m}_{\mathrm{t}} \cdot \mathrm{g} \cdot \mathrm{v} \sin\theta}{A_D} \tag{7}$$

where: $\mathrm{WR}_{\mathrm{ext}}$, external work rate per area, kJ/(m^2 h)

θ, angle of inclination with respect to the vertical,

(= arctan (G/100)), degrees

If equation (7) is expressed in Watts, it becomes:

$$WR_{ext} = m_t \cdot g \cdot v \sin\theta \tag{8}$$

where: WR_{ext} is the external work rate, W

The difference between equations 6 and 8 is the G/100 and sin θ terms. These terms are equivalent for grades up to 25%.

Other researchers (Groot et al., 1994) filmed subjects during exercise and then determined the individual joint moments and angular velocities. The joint power was found as the product of the joint moments and velocities. The sum of these joint powers reflected the external work rate for the task.

Muscular Efficiency

The amount of power input to a machine is greater than the power output. This is because machines are not 100% efficient. Mechanical efficiency is the power output divided by the power input. Humans are also not 100% efficient. In physiology, mechanical efficiency is referred to as overall or gross efficiency. It is found by dividing the external work rate by the physiological work rate. The physiological work rate, sometimes called the metabolic cost of exercise, is the internal energy required to produce external work.

There are other definitions of efficiency encountered in the literature. There are net efficiency, work or apparent efficiency, delta efficiency, and activity specific efficiencies such as propelling efficiency for swimming. Net efficiency is external work rate divided by the difference of physiological work rate and resting metabolic work rate (Fukunaga et al., 1986). The resting metabolic work rate, or basal metabolic rate, is the amount of energy required by the body for the chemical and metabolic processes required to sustain life. The work or apparent efficiency is found by dividing external work by the difference of physiological work rate and the energy expenditure during non-working conditions. The term work efficiency is used typically for bicycle exercise while apparent efficiency is used for treadmill walking or running (Stainbsy et al., 1980). Delta efficiency is the increment in work rate performed above the previous work rate divided by the increment in physiological work rate above the previous work rate (Fukunaga et al., 1986).

Stainbsy et al. (1980) discussed the validity of base-line subtractions for determining efficiency. The authors indicated that there were differences between exercise efficiency and muscle efficiency. Muscle efficiency should be determined from the processes that provide and convert energy to work (Stainbsy et al., 1980). Exercise efficiency was the external work divided by the energy required to perform that work.

It was suggested (Stainbsy et al., 1980) that muscle efficiency be determined as the product of phosphorylative coupling efficiency and contraction coupling efficiency. The energy for muscular contraction comes from the oxidation of nutrients. Part of this energy is saved in the ATP molecule. This process was termed phosphorylative coupling. The phosphorylative coupling efficiency was found by dividing the free energy conserved as ATP by the free energy of oxidized foodstuff (Stainbsy et al., 1980). Some of the energy from the ATP was used to perform work and was termed contraction coupling. Contraction coupling efficiency was calculated by dividing the external work accomplished by the free energy of ATP hydrolysis (Stainbsy et al., 1980).

The authors argued against using base-line subtractions in determining efficiency because the base line values have been found to change as exercise intensity increases and are thus invalid. They stated that gastrointestinal processes decreased, splanchnic metabolism increased, and energy required by the lungs increased with increasing work rate. Additionally, body temperature increases which then increases the metabolic rate. These factors all caused changes in the base-line values and, according to Stainbsy et al. (1980), precluded the use of efficiencies using base-line subtractions. The authors further stated that while none of the widely used and widely accepted efficiencies (gross, net, apparent, work, and delta) really represent muscle efficiency, there were no errors in using gross efficiency as long as it was referred to as exercise efficiency.

Gross efficiency depends on the work rate, type of work, and which muscles are used. There is a lot of error in the efficiency calculation due to human variability and the fact that external work rate alone does not determine efficiency. Muscular efficiency is influenced by the subject’s coordination and familiarity with the activity being performed (Robergs and Roberts, 1997). Wasserman et al. (1999) found that experience in treadmill walking may lead to an increase in efficiency. Activities involving fine movements generally have low efficiencies while activities such as running that involve gross movements and large muscle mass have higher efficiencies (Johnson, 1991). As the resting metabolic demands become a smaller proportion of overall energy requirements, gross efficiency approaches a maximum value of 20% (Johnson, 1991).

Efficiency Studies

Many studies have been performed that investigated the efficiency of various activities. Fukunaga et al. (1986) found that for college oarsmen, the gross efficiency of rowing in the external work rate range of 124 – 182 W was 17.5%. The efficiency of swimming in competitive male and female swimmers ranged from 5 to 9.5% (Toussaint et al., 1990). The authors found that as power output increased, gross efficiency increased also.

Webb et al. (1988) investigated the work done by five males and five females during bicycle ergometer work at 53 and 92 W. Heat production was measured using a suit calorimeter. Gross efficiency for the 53 W and 92 W workloads were 13% and 17%, respectively.

The effects of speed and work rate on muscular efficiency during steady-rate exercise on a bicycle ergometer were investigated (Gaesser and Brooks, 1975). Gross muscular efficiencies were reported at work rates of 33, 65, 98, 131 W at pedaling rates of 40, 60, 80, and 100 rpm. Their results are shown in Table 1. They found that as pedaling frequency increased, efficiency decreased.

Table 1. Gross efficiencies at four work rates and four pedaling rates. Efficiencies are reported as mean ± standard deviation. Data are from Gaesser and Brooks (1975).

	33W	65W	98	131
40 rpm	12.0 ± 0.3%	17.0 ± 0.3%	19.3 ± 0.2%	20.2 ± 0.4%
60rpm	12.1 ± 0.3%	16.6 ± 0.3%	19.2 ± 0.4%	20.4 ± 0.4%
80 rpm	10.2 ± 0.2%	14.8 ± 0.2%	17.6 ± 0.2%	18.8 ± 0.3%
100 rpm	7.6 ± 0.3%	12.1 ± 0.3%	15.1 ± 0.2%	16.6 ± 0.3%

The physiological responses of nineteen subjects to arm, leg, and combined arm and leg ergometry at work rates of 49, 73.5, and 98 W was investigated by Eston and Brodie (1986). Physiological work rate was calculated. The average gross efficiencies for the work rates of 49, 73.5, and 98 W for arm ergometry were 11.8%±0.6%, 12.5%±1.20%, and 12.5%±1.20%. These efficiencies were significantly different from the leg and combined arm and leg efficiencies. For leg ergometry the efficiencies were 13.5%±0.80%, 15.60%±1.40%, 17.11%±1.20% while for the combined arm and leg ergometry the efficiencies were 12.90%±1.30%, 15.20%±1.10%, and 16.8%±1.60%. All efficiencies are reported in order of increasing work rate. There were no statistically significant differences between the leg and combined arm and leg ergometry.

Luhtanen et al. (1987) reported gross efficiencies for subjects on a bicycle ergometer. For work rates of 146±15, 190±4, 225±12, 254±11, and 283±17 W, the gross efficiencies were 19.7%±3.7%, 19.7%±2.8%, 18.9%±2.8%, 18.2%±2.8%, and

17.4±1.0%, respectively. On average, the efficiency of the subjects decreased as external work rate increased.

Nagle et al. (1990) investigated the work done in grade walking on a treadmill. If the non-thermal energy term is ignored (an average of 6% of transformed energy), the efficiencies for walking at a speed of 1.5 m/s at grades of 5, 10, -5, and –10% were 10.6%, 15.8%, -20%, and –48.8% respectively.

Haembraeus et al. (1994) adapted the suit calorimeter used by Webb et al. (1988) and Nagle et al. (1990) so that the suit could be used for exercise intensities of 250W or higher. Unfortunately, external and internal work rates were reported only for two male subjects, one twenty-nine year old and one fifty-five year old. The fifty-five year old subject completed one trial on a bicycle at 100W. The efficiency of the activity was 22%. The twenty-nine year old completed two bike sessions at 200 W and one at 100W. The efficiencies for these activities were 19.4%, 18.9%, and 15.2%.

The muscular efficiency of uphill and downhill walking at a constant speed of 1.1 m/s was investigated by Johnson et al. (2001a). The authors found that the efficiency of downhill walking was negative two times the efficiency of uphill walking. These results were supported by the work of Orsini and Passmore (1951), Pivarnik and Sherman (1990), and Nagle et al. (1990).

Hesser (1965) examined the efficiency of 10 male and 10 females climbing up and down stairs at speeds of 88 steps/min and 160 steps/min. The author found that for the lower speed, the ratio of oxygen cost of positive work to negative work was 8:1. When the speed was increased, the ratio decreased to 5:1. These results contrasted with those of Abbott et al. (1952) and Asmussen (1953) who found that for bicycle ergometer work the ratio increased with speed. The fact that negative work in running or walking is more efficient than positive work is supported by Pimental et al. (1982) and Davies et al. (1974).

Equations Relating Efficiency to External Work Rate

Johnson (1992) developed a series of equations relating gross muscular efficiency to external work rate. Maximum efficiency was assumed to be 20%. The equations were:

$$\eta = \frac{WR_{ext}}{200} \qquad 0 \le WR_{ext} \le 10 \qquad (9)$$

$$\eta = 0.05 + 0.001(WR_{ext} - 10) \qquad 10 \le WR_{ext} \le 140 \qquad (10)$$

$$\eta = 0.18 + 0.0002(WR_{ext} - 140) \qquad 140 \le WR_{ext} \le 240 \qquad (11)$$

$$\eta = 0.2 \qquad 240 \le WR_{ext} \qquad (12)$$

where η, muscular efficiency, dimensionless

WR_{ext}, external work rate, W

Physiological Work Rate

Physiological work rate is the internal energy required to produce external work. Physiological work rate is equal to the external work rate divided by the muscular efficiency. For steady-state exercise when there is no change in the body temperature and thus no change in the rate of heat stored, physiological work rate is the sum of the heat produced during exercise and the external work produced.

If the external work is zero, then determining physiological work rate in the above way would give a physiological work rate of zero. If the person is resting, the physiological work rate would equal the basal metabolic rate. However, if the person is running on level ground, the person has a physiological work rate much higher than basal metabolic rate. An alternative to calculating external work rate would be to use a look-up table that provides physiological work rates for walking, running, and other tasks. When a look-up table is used, it is important to consider the conditions under which the values were obtained. Factors such as age, body mass, gender, and fitness level would be important. Tables of physiological work rates for leisure, work, and military tasks can be found in many sources including Johnson (1992), Johnson (1991), and McArdle et al. (1996).

The physiological work rate can also be calculated in another manner. Givoni and Goldman (1971) developed an empirical equation for predicting the metabolic energy cost of level and grade walking, with and without loads. The equation was found to apply to walking speeds of 0.7 m/s to 2.5 m/s at grades up to 25% and for running speeds from 2.22 m/s to 4.72 m/s at grades up to 10% with loads up to 70 kg. The authors suggested empirical coefficients to modify the equation for different terrains, for load placement, and for very heavy work levels. The results showed a correlation of 0.95 between predicted and measured values.

Pandolf et al. (1977) continued the work of Givoni and Goldman (1971) by adjusting the equation to make predictions for subjects who were standing or walking very slowly (less than 0.7 m/s). The authors validated the equation with two studies. The first involved six males walking at speeds of 1.0, 0.8, 0.6, 0.4, and 0.2 m/s while carrying loads of 32, 40, and 50 kg. In the second experiment, ten males stood while wearing backpacks that had masses of 0, 10, 30, and 50 kg. Good agreement was found between the empirical model and the experimental results.

Myles and Saunders (1979) had nine male subjects walk on a treadmill with loads equal to 10% and 40% of body weight. They used the equation developed by Pandolf et al. (1977). Good agreement was found between the predicted and measured values.

Physiological work rate, or metabolic energy cost, can also be determined using either direct or indirect calorimetry. Direct calorimetry measures the amount of heat produced by the body. Indirect calorimetry relates the total metabolic heat production of the body to oxygen consumed and carbon dioxide produced.

With direct calorimetry, the subject is placed typically in a thermally isolated chamber (Ferrannini, 1988) for periods of 24 hours or more. The heat lost through evaporation, radiation, conduction, and convection is measured. These chambers are expensive and are not common. A suit calorimeter was developed that enabled a subject to perform activities outside of a chamber (Webb et al., 1988; Nagle et al., 1990; Hambraeus et al., 1994).

The use of indirect calorimetry began over two hundred years ago when Adair Crawford in England and Antoine Lavoisier in France proved that respiratory gas exchange represented combustion similar to that of a burning candle (Webb, 1991). The energy production results from converting nutrients (carbohydrate, fat and protein) into the chemical energy of ATP minus the energy used in the oxidation process (Ferrannini, 1988). Indirect calorimetry assumes that all of the oxygen consumed is used to oxidize fuel and that all the evolved carbon dioxide is recovered (Ferrannini, 1988). So, measuring oxygen consumed and carbon dioxide produced gives an estimate of the energy production of the body. Swyer (1991) has found that estimations of metabolic energy production made with indirect calorimetry agreed with direct calorimetry values for steady-state conditions if proper procedures were followed.

Many equations have been developed that relate the metabolism of nutrients to the oxygen equivalent of the metabolism. The equations are based on the same theory but differ in their assumptions and intended applications. These equations have been used by many investigators to estimate metabolic energy production.

The theoretical Weir (1949) equation was based on the caloric equivalent of oxygen and carbon dioxide. The equation used total respiratory quotient which includes metabolism of carbohydrate, fat, and protein. Weir assumed that the total percentage of protein calories was between 10 and 14%. If this assumption held, the error in the equation was less than 0.2%. The equation was:

$$\frac{\text{kcal liberated}}{\text{L O}_2\text{ consumed}} = 3.9 + 1.1\text{RQ} \tag{13}$$

where: RQ, total respiratory quotient, dimensionless

Garby and Astrup (1987) developed a theoretical equation based on the metabolism of carbohydrate and fat. Protein metabolism was assumed to be zero. Thus, the respiratory quotient used is termed the non-protein respiratory quotient. The equation was:

$$\mathrm{O_2 - eq.} = \mathrm{A} \cdot \mathrm{NPRQ} + \mathrm{B} \tag{14}$$

where: O_2 – eq., energy equivalent of oxygen, J/L

A, B, coefficients that depend on the amounts of carbohydrate and fat consumed, J/L

NPRQ, non-protein respiratory quotient, dimensionless

The most commonly used values for the coefficients A and B in equation 14 were 4,940 J/L and 16,040 J/L, respectively (Garby and Astrup, 1987).

A third theoretical equation was presented by Lusk (1928). The equation was:

$$\frac{cal}{L\,O_2\ consumed} = 4.686 + \frac{0.361 \cdot (RER - 0.707)}{0.293} \tag{15}$$

where: RER, respiratory exchange ratio, dimensionless

The Weir (1949), Garby and Astrup (1987), and Lusk (1928) equations all determined the energy equivalent of oxygen. Physiological work rate was determined from these equations by multiplying the energy equivalent of oxygen by the oxygen consumption and converting units. The equations for predicting physiological work rate from the Weir, Garby and Astrup, and Lusk equations, respectively were:

$$\mathrm{WR}_{\mathrm{phys}} = \frac{(4606\mathrm{RQ} + 16329)V_{\mathrm{O2}}}{60} \tag{16}$$

$$\mathrm{WR}_{\mathrm{phys}} = \frac{(4940\mathrm{NPRQ} + 16040)V_{\mathrm{O2}}}{60} \tag{17}$$

$$\mathrm{WR}_{\mathrm{phys}} = \frac{(5155\mathrm{RER} + 15962)V_{\mathrm{O2}}}{60} \tag{18}$$

where: $\mathrm{WR}_{\mathrm{phys}}$, physiological work rate, W

V_{O2}, oxygen consumption, L/min

Gagge and Nishi (1983) presented an equation for predicting metabolic energy from oxygen consumption and carbon dioxide production:

$$WR_{phys} = (0.23RER + 0.77)(5.873)V_{\mathrm{O2}}(60) \tag{19}$$

where: $\mathrm{WR}_{\mathrm{phys}}$, physiological work rate, W

RER, respiratory exchange ratio, dimensionless

V_{O2}, oxygen consumption, L/min

5.873, energy equivalent of oxygen, W·hr/L

The authors recommended that the equation not be used for transient conditions. Putting their equation in the same format as above yielded:

$$\mathrm{WR}_{\mathrm{phys}} = \frac{(4863\mathrm{RER} + 16280)V_{\mathrm{O2}}}{60} \qquad (20)$$

where: $\mathrm{WR}_{\mathrm{phys}}$, physiological work rate, W

RER, respiratory exchange ratio, dimensionless

V_{O2}, oxygen consumption, L/min

The four equations (16 – 18, 20) have similar coefficients. Some of the equations used respiratory quotient while others used respiratory exchange ratio.

The respiratory quotient is defined as the ratio of the rate of carbon dioxide produced to the rate of oxygen consumed. The respiratory exchange ratio is defined as the ratio of the rate of carbon dioxide exhaled to the rate of oxygen consumed. So, the difference is in the carbon dioxide term. RQ deals with cellular respiration and is used to calculate the caloric value of oxygen consumption. RER is related to external respiration and is an indication of the work intensity. RQ and RER can be considered to be equal except under the following conditions: metabolic acidosis, non-steady state conditions, hyperventilation, excess post-exercise oxygen consumption, and extremely heavy exercise (Robergs and Roberts, 1997; Johnson, 1991).

The RQ cannot exceed 1.0 because the carbon dioxide produced by cells cannot exceed the oxygen consumed. However, when excess acid is produced (metabolic acidosis) such as during heavy exercise, the body produces increased levels of carbon dioxide separate from oxygen consumption due to buffering of the carbon dioxide. Because of the excess carbon dioxide produced, RER can exceed 1.0 under conditions of metabolic acidosis.

Under non-steady state conditions, oxygen consumption has not had a chance to increase to levels that account for ATP produced during metabolism. Instead, the ATP comes from creatine phosphate hydrolysis and glycolysis. So, a lower metabolic intensity would be indicated during the transition than if the person had already achieved a steady state (Robergs and Roberts, 1997).

During hyperventilation, the volume of carbon dioxide exhaled from the lung increases. This can occur without increases in oxygen consumption, so the RER may be increased. The RQ would remain the same because the carbon dioxide produced by the cells had not increased.

Finally, after exercise, the amount of carbon dioxide exhaled decreases rapidly while the oxygen consumption remains elevated above resting levels. Thus, the RER may decrease below resting values (Robergs and Roberts, 1997).

The actual physiological work rate can be less than the predicted when there are connections between a subject and the test apparatus other than, for example, the connection between the shoes and the treadmill belt. Wasserman et al. (1999) stated that railings, armboards, mouthpieces, blood pressure measuring devices, and steadying hands could all reduce the patient’s metabolic requirement. The mass of shoes and stiffness of their soles may affect the physiological work rate (McArdle et al., 1996). Loads carried on the foot increase the physiological work rate more than loads carried on the torso. So, heavy shoes would cause a greater increase in physiological work rate than lightweight shoes. Softer-soled shoes reduce the physiological work rate compared to stiffer soled shoes.

Oxygen Consumption

As long as the work rate is not too high during constant-rate exercise, oxygen consumption will reach a steady-state. A secondary rise in oxygen consumption, or oxygen drift, may occur for extended periods of exercise (Poole and Richardson, 1997; Kearon et al., 1991). There appears to be both a fast and slow component to oxygen consumption during work. The fast component is responsible for the initial steady-state reached. The oxygen drift is thought to be related to the slow component. The slow component may also cause a greater than linear increase in the oxygen consumption with work rate above the anaerobic threshold. The presence of oxygen drift would be important to consider for a model of steady-state exercise.

There are exercise levels for which oxygen consumption will continue to rise until the maximum oxygen consumption is reached, fatigue occurs, and exercises stops. While it may not be possible physiologically for a subject to attain the steady-state, the theoretical steady-state value is still necessary to determine the response (Givoni and Goldman, 1972).

Slow Component of Oxygen Consumption

Poole and Richardson (1997) stated that the four most important determinants of oxygen consumption response during exercise were external work rate, work efficiency, whether the work was incremental or constant load, and the intensity level of the work (above or below the anaerobic threshold). The heavy exercise domain starts at the anaerobic threshold. The highest exercise level in this domain is the highest work rate at which blood lactate production can be stabilized, albeit at an elevated level (Poole and Richardson, 1997). The slow component of oxygen consumption is evident in this domain 80-100 seconds after the start of exercise (Poole and Richardson, 1997). Work efficiency is reduced in this domain.

The severe exercise intensity domain begins around 50% of the difference between the anaerobic threshold and V_{O2max} (Poole and Richardson, 1997). In this domain, blood lactate levels continue to increase and the slow component pushes the oxygen consumption towards V_{O2max}.

Gaesser and Poole (1996) suggested that the increase in oxygen consumption for exercise above the anaerobic threshold (slow component) not be confused with oxygen drift. The authors suggested that oxygen drift occurs during prolonged moderate intensity exercise and is a small increase (200mL) in the oxygen consumption. The slow component of the oxygen consumption response on the other hand is only seen for exercise above the anaerobic threshold and is of much greater magnitude. It is the increase in oxygen consumption beyond the third minute of exercise (Gaesser and Poole, 1996).

Whipp and Wasserman (1972) investigated the oxygen uptake kinetics for various intensities of constant-load work. They found that for low work rates, the oxygen consumption reached a steady-state within three minutes. At higher work rates, the steady-state was progressively delayed. A difference was found between the oxygen consumption measured at three minutes and that measured at six minutes.

The authors found that this difference was a useful indicator of the slow component of oxygen consumption.

Kearon et al. (1991) investigated oxygen consumption, minute ventilation, tidal volume, and respiratory rate during prolonged exercise at work rates of 34%, 43%, 63%, and 84% of maximal capacity in six healthy subjects. Subjects exercised for 60 minutes or until they could not continue. The average (± standard error of the mean) performance times at the four work rates were 60 ± 0 min, 56 ± 4.0 min, 37 ± 6.6 min, and 12 ± 3.7 min, respectively. A regression line was fit to the average oxygen consumption data versus time for each of the work conditions. Data collected in the first four to six minutes was ignored as this was considered to be the time it took for the subjects to reach a steady state. A statistically significant increase in the oxygen consumption was declared if the slope of the regression line was significantly different from zero.

At the lowest work rate, there was a small but statistically significant increase in the oxygen consumption from 1.47 to 1.52 L/min. Oxygen consumption increased during the 43% work rate from 1.76 to 1.93 L/min. the differences at these two work rates were in the 200 mL range that Gaesser and Poole (1996) suggested indicates oxygen drift rather than the slow component. For the third work rate, oxygen consumption increased from 2.35 L/min to 2.84 L/min. Finally, oxygen consumption values for the highest work rate increased from 3.13 to 3.59 L/min. All of the increases were statistically significant.

Barstow and Mole (1991) investigated oxygen uptake kinetics during heavy exercise. Four trained cyclists completed four replications of cycle exercise at four work rates, two of which were below the anaerobic threshold. The four work rates were 35, 55, 85, and 100% of maximal oxygen consumption. Each test consisted of four minutes of pedaling at 33W followed by eight minutes at the selected work rate and then ten minutes of recovery at 33W. Two exponential models were fit to the data:

$$\Delta \dot{V}_{O_2}(t) = A_1\left[1-e^{-(t-TD)/\tau_1}\right] + A_2\left[1-e^{-(t-TD)/\tau_2}\right] \tag{21}$$

$$\Delta \dot{V}_{O_2}(t) = A_1\left[1-e^{-(t-TD_1)/\tau_1}\right] + A_2\left[1-e^{-(t-TD_2)/\tau_2}\right] \tag{22}$$

where: $\Delta V_{O2}(t)$, oxygen consumption response above baseline, L/min
t, time starting from the onset of exercise, sec
A_1, first steady-state oxygen consumption, L/min
A_2, second steady-state oxygen consumption, L/min
TD_1, TD_2, time delays for phase two and three, respectively, sec
τ_1, τ_2, time constants for phase two and three, respectively, sec

The difference between the two equations was that the second equation was a more general form that allowed a second independent time delay.

A single-exponential function of the form:

$$\Delta \dot{V}_{O_2}(t) = A_3\left[1-e^{-(t-TD)/\tau_3}\right] \tag{23}$$

where: A_3, the sum of A_1 and A_2 from the first equation
τ_3 equals τ_1 and τ_2

fit the data for all eight exercise cases below the anaerobic threshold (two work rates for each of four subjects). So, for the oxygen consumption response below the anaerobic threshold, there is only one steady-state value, A_3.

For seven of the eight responses above the anaerobic threshold, a two-exponential function (equation 22) was found to fit the data. For the eighth case, the single exponential function (equation 23) was the best fit. The better fit of the two-exponential model indicated that for exercise above the anaerobic threshold there was a second component to the oxygen consumption that did not begin at the same time as the first exponential, but began later into the exercise. The authors concluded that this was evidence of a slow component of the oxygen consumption response. Equation 22 was modified in a later study (Mole and Hoffmann, 1999) to include baseline oxygen consumption in the response:

$$\Delta \dot{V}_{O_2}(t) = \alpha_R + \alpha_F\left[1 - e^{-(t-TD)/\tau_F}\right] + \alpha_S\left[1 - e^{-(t-TD)/\tau_S}\right] \qquad (24)$$

where: α_R, initial resting oxygen consumption, L/min

α_F, steady-state V_{O2} due to the fast component, L/min

α_S, steady-state V_{O2} due to the slow component, L/min

τ_F, time constant for the fast component, sec

τ_S, time constant for the slow component, sec

Similar results were found by Paterson and Whipp (1991). Six healthy subjects performed two to four repetitions of cycle exercise from a baseline of unloaded pedaling to one of two selected work rates, one at 90% of the anaerobic threshold and the other at the halfway point between the anaerobic threshold and V_{O2max}. A single-exponential function was the best fit equation for the oxygen consumption response for the exercise below the anaerobic threshold. For exercise above the anaerobic threshold, the authors found that a two-exponential model, with separate time constants and time delays was the most accurate model. It was concluded that the slow component of the oxygen consumption response was a delayed-onset process. The two-exponential model has been shown to be accurate for predicting the steady-state oxygen consumption for exercise intensities above the anaerobic threshold (Bernard, et al., 1998). The two exponential response of oxygen consumption with time has been shown also in untrained subjects (Camus, et al., 1988).

The physiological reason or reasons for the slow component of oxygen consumption are still under debate. Possible reasons include lactate, epinephrine, cardiac and ventilatory work, temperature, potassium, and recruitment of lower-efficiency fast-twitch muscle fibers (Gaesser and Poole, 1996). Poole et al. (1992) showed that most (86%) of the increase in oxygen consumption beyond the third minute was due to a increase in leg oxygen consumption. So, Gaesser and Poole (1996) suggested that factors that do not involve working muscles probably make only small contributions to the slow component. They suggested that muscle temperature and more importantly, the recruitment of lower efficiency fast-twitch muscle fibers were the major factors contributing to the slow component.

Steady-State Oxygen Consumption

Equations that related physiological work rate to RQ, NPRQ, or RER and oxygen consumption have been discussed previously. If the physiological work rate were calculated using a separate method, the above equations could be solved for oxygen consumption in terms of physiological work rate and RQ, NPRQ, or RER.

Johnson (1992) fit equations to experimental data in Hurley et al. (1984) that related respiratory exchange ratio to percent of maximum oxygen consumption for trained and untrained subjects:

$$RER = 0.842 \qquad 0 \le \frac{V_{O2}}{V_{O2\max}} \le 0.1 \qquad \text{untrained} \qquad (25)$$

$$RER = 0.778 \qquad 0 \le \frac{V_{O2}}{V_{O2\max}} \le 0.1 \qquad \text{trained} \qquad (26)$$

$$RER = 0.826 + 0.160(\frac{V_{O2}}{V_{O2\max}}) \qquad 0.1 \le \frac{V_{O2}}{V_{O2\max}} \le 0.8 \qquad \text{untrained} \qquad (27)$$

$$RER = 0.756 + 0.220(\frac{V_{O2}}{V_{O2\max}}) \qquad 0.1 \le \frac{V_{O2}}{V_{O2\max}} \le 0.9 \qquad \text{trained} \qquad (28)$$

$$RER = -0.230 + 1.480(\frac{V_{O2}}{V_{O2\max}}) \qquad 0.8 < \frac{V_{O2}}{V_{O2\max}} \qquad \text{untrained} \qquad (29)$$

$$RER = -0.810 + 1.960\left(\frac{V_{O2}}{V_{O2\max}}\right) \qquad 0.9 \le \frac{V_{O2}}{V_{O2\max}} \qquad \text{trained} \qquad (30)$$

where: RER, respiratory exchange ratio, dimensionless

V_{O2}, oxygen consumption, L/min

V_{O2max}, maximum oxygen consumption, L/min

These equations assumed RER=1.25 for untrained and RER=1.15 for trained individuals.

Johnson's (1992) RER equations could be substituted for NPRQ in the physiological work rate equation that could then be solved for oxygen consumption. For very heavy exercise, errors in calculating the oxygen consumption and subsequent parameters would result when substituting RQ for RER. These errors should be evaluated.

Other methods of determining oxygen consumption from work have been developed. Astrand and Rodahl (1970) showed in their Figure 13-2 that oxygen consumption was related linearly to physiological work rate. ACSM (2000) provided equations for estimating oxygen consumption for treadmill walking or running, ergometry, and stepping. Van der Walt and Wyndham (1973) developed equations to predict oxygen consumption for level treadmill walking and running. Their equations were of the form:

$$\dot{V}_{O_2} = A_1 + A_2 m + A_3 m v^2 \qquad (31)$$

where: A_1, A_2, and A_3, empirically derived regression coefficients m, mass, kg

v, velocity, m/s

The authors did not investigate the effects of loads carried, grade, or of ambulating on surfaces other than a treadmill. Equations such as those developed by ACSM (2000) and Van der Walt and Wyndham (1973) are useful for predicting oxygen consumption of specific activities, but have no use in predicting the oxygen consumption of other activities such as painting or wood working. The Astrand and Rodahl (1970) plot may show an idealized relationship, but is worth considering.

Astrand and Rodahl (1970) showed that the absolute oxygen consumption required by the body depended on the physiological work rate (their Figure 13-2). Logically, the higher the work rate, the greater the amount of oxygen consumed. Because their graph showed a completely straight line with no regression equation, the graph may show an idealized relationship.

Effects of Age and Training

The following factors may cause the actual oxygen consumption to differ from the predicted: faulty ergometer calculation, obesity, cardiovascular disease, pulmonary disease, fitness, exercise protocol, handrail holding, stride length, training specificity, habituation, and coordination (Robergs and Roberts, 1997; Wasserman et al., 1999). For trained individuals, steady-state oxygen consumption is lower at a given work rate than for untrained individuals due to an attenuation of the slow component (Gaesser and Poole, 1996). The reason for the decrease in the V_{O2} slow component may be due to the increase in mitochondria in all fibers that occurs with endurance training. Training also can speed up the transient response while detraining and cardiopulmonary disease can decrease the response (Poole and Richardson, 1997). Children have a greater gain for the fast component than adults and exhibit little or no slow component (Barstow, 1994).

Anaerobic Threshold

The point at which the lactate levels in the blood begin to rise during incremental exercise has been termed the anaerobic threshold (AT) (Wasserman, 1973). When the oxygen required by the muscles can be supplied by ventilation alone, metabolism occurs aerobically. If the oxygen demand of the exercising muscles cannot be supplied by ventilation alone, then ATP production does not occur at the mitochondrial level (Claiborne, 1984) but is instead produced anaerobically (Sady, et al., 1980). Thus, around the anaerobic threshold, non-oxidative metabolism plays more of a role in energy production (Sady et al., 1980). Lactic acid production increases and is buffered by the bicarbonate system (Weltman and Katch, 1979), resulting in an increase in the production of non-metabolic carbon dioxide. The increase in CO_2 production acts as an strong ventilatory stimulus (Sady et al., 1980), causing the minute ventilation-oxygen consumption relationship to increase beyond linear (Wasserman, 1973).

There are invasive and non-invasive techniques for determining the anaerobic threshold. Wasserman et al. (1973) stated that the AT was the point of: "1) nonlinear increase in minute ventilation, 2) nonlinear increase in carbon dioxide production, 3) an increase in end-tidal oxygen without a corresponding decrease in end-tidal carbon dioxide, and 4) and increase in the respiratory exchange ratio, as work rate was increased during exercise." The term "lactate threshold" is sometimes used to describe the point at which lactic acid begins to accumulate in the blood (Johnson, 1991; Johnson et al., 1995). The point at which minute ventilation increases beyond linear is sometimes called the "ventilation threshold" (Johnson, et al., 1995; Mahon and Vaccaro, 1989).

Other researchers disagreed with the description of anaerobic threshold provided in Wasserman et al. (1973). Skinner and McLellan (1980) labeled the set of responses observed by Wasserman et al. (1973) as the "aerobic threshold". They contended that there were really three phases to exercise, not two. The second breakaway point was described as the point at which lactic acid increased from 4 mmol/L, FE_{CO2} decreased, and hyperventilation increased. This point occurred between 65-90% of V_{O2max} and was termed the "anaerobic threshold" (Skinner and McLellan, 1980).

There has been some disagreement about whether the AT as determined by blood analysis is the same as that determined from respiratory gas exchange. Powers et al (1984) compared the onset of AT measured by blood lactate and estimated by the point where ventilation increased non-linearly. They found that the two points did not always occur simultaneously and suggested that there may be limitations to estimating the AT using respiratory gas exchange. However, Ivy et al. (1980) found that there were no significant differences between the two methods of estimating the

AT. Davis et al. (1976) found a correlation coefficient of 0.95 between the two methods.

One of the major problems with determining AT using respiratory gas exchange is the subjectivity involved (Davis et al., 1976). Computer programs that use objective methods of determining AT from respiratory gas exchange have been developed (Herbert et al., 1982; Orr, et al., 1980). This eliminates the problem resulting from researcher subjectivity in detecting the point at which the curve departs from linearity.

A new method of detecting the AT from gas exchange variables was presented by Caprarola and Dotson (1985). They plotted FE_{CO2} versus percent of maximal oxygen consumption and fit a quadratic equation to the data. The point at which the curve was a maximum was the anaerobic threshold. The authors found good agreement between this method and standard techniques.

Johnson et al. (1995) investigated the effects of full-facepiece masks and half-masks on ventilation threshold and lactate threshold on fourteen subjects undergoing incremental bicycle exercise. These researchers found that mask condition did not affect either the lactate or ventilation thresholds.

Many studies investigating the relationship between anaerobic threshold and oxygen consumption have been performed. Subjects of these studies have been male and female, trained and untrained.

Weltman et al. (1978) reported that for thirty-three female college students, the AT occurred at an average of 50% of V_{O2max}. These researchers paired 22 subjects according to their V_{O2max} values. Paired members had similar V_{O2max} values but different AT values. The average V_{O2max} for the two groups (36.66± 7.62 and 38.36±6.28 for the low and high AT groups respectively) were not significantly different statistically. The AT values were significantly different. The AT values for the low and high AT groups were 16.23 ±4.57 L/min and 21.35±4.14 L/min, respectively. This corresponded to an AT% of 44% and 56% for the low and high AT groups respectively. So, even though the two groups had similar subjects, the AT (ml/kg/min) was quite different.

Dwyer and Bybee (1983) reported that the AT occurred at an average of 70±7% of V_{O2max} for twenty female recreational runners and cyclists. Average V_{O2max} was 38.4±4.7 ml/kg/min. They found a high correlation (r = 0.87) between AT (L/min) and V_{O2max} (L/min).

Fifteen trained female cross-country skiers aged fifteen to twenty with an average V_{O2max} of 47.3±3.6 ml/kg/min were studied (Rusko et al., 1980). The AT (40.9±3.3 ml/kg/min) occurred at 85.7±6.6% of V_{O2max}. A correlation (r=0.6) was found between V_{O2max} (ml/kg/min) and AT (ml/kg/min). An insignificant correlation was found between AT expressed as a percent of V_{O2max} (AT%) and V_{O2max} (ml/kg/min).

Eighteen overweight females were studied by Sady et al. (1980). Subjects were split into three groups for different exercise treatments. Pre-training V_{O2max} and AT values are reported for each of these three groups separately. The V_{O2max} values for the three groups (n=7, n=7, and n=4) were 2.23±0.07, 2.09±0.18, and 2.31±0.12 L/min while the AT values were 1.02±0.06, 0.97±0.04, and 1.28±0.08 L/min, respectively. These AT values corresponded to 46, 46, and 55% of VO2max, respectively.

Thorland et al. (1980) studied ten trained female collegiate cross-country runners. The AT occurred at average of 80% of V_{O2max}. The anaerobic threshold expressed in ml/kg/min was highly correlated with maximal oxygen consumption (r = 0.81).

Weltman and Katch (1979) found that thirty-one male subjects with an average V_{O2max} of 51.36±6.36 ml/kg/min had an AT of 59.5±7.70% of V_{O2max}. They found a high correlation (r=0.81) between V_{O2max} (L/min) and AT (L/min).

Thirteen trained men were studied by Powers et al. (1984). The AT of these subjects occurred at an average of 56% of V_{O2max}. Balsom (1988) studied fourteen male college soccer players with an average V_{O2max} of 57.4±6.18 ml/kg/min. The average AT occurred at 70.5±5.99% of V_{O2max}. Robbins et al. (1982) found that for healthy adult males with a mean V_{O2max} of 59.2 ml/kg/min, the AT occurred at an average of 65.3% of VO2max. For male college students performing arm-cranking, leg cycling, and treadmill walk-running, the AT occurred at average values of 46.5, 63.8, and 58.6% of VO2max, respectively (Davis et al., 1976). Jones (1984) found that the AT occurred at 50% ± 4.8% of the V_{O2max} for inactive, young, adult male smokers (average V_{O2max}: 34 ml/kg/min). For males aged 24-35, Bradley (1982) found that AT occurred at 58.6%±10.7% of V_{O2max} (average V_{O2max} was 45.7 ± 7.9 ml/kg/min).

These studies show that the occurrence of the AT is highly variable even among subjects of the same gender and similar ages and training statuses. Skinner and

McLellan (1980) reported that the anaerobic threshold occurs between 65 and 90% of V_{O2max}. The studies discussed here have shown values outside this range. These studies have reported that the anaerobic threshold can occur between 29 and 95% of maximal oxygen consumption. The anaerobic threshold for trained athletes occurs at a higher percentage of V_{O2max} than for untrained subjects. While the AT can be elevated after training even if there is not an increase in V_{O2max} (Claiborne, 1984), generally the higher the V_{O2max}, the higher the AT.

A significant relationship between the AT and V_{O2max} was reported by Dwyer and Bybee (1983), Rusko et al. (1980), Thorland et al. (1980), and Weltman and Katch (1979). The other researchers did not report on this relationship. Only one paper reported regression of AT% on V_{O2max} (Rusko et al., 1980); no correlation was found.

The anaerobic threshold is important because relationships below the anaerobic threshold differ from the relationships above the anaerobic threshold (Johnson, 1991). Martin and Weil (1979) found that for incremental exercise below the anaerobic threshold, the minute volume increased linearly while above the threshold it increased at a greater rate. And, the time to reach a steady-state in oxygen consumption is longer above the anaerobic threshold (Wasserman et al., 1973).

Wasserman et al. (1973) suggested that for patients with severe respiratory impairment, an AT may not be present because these subjects might not be able to exercise at a high enough rate to elicit lactic acidosis. For subjects with cardiovascular impairment, the anaerobic threshold will occur at lower values than for healthy subjects (Wasserman, et al., 1973).

Minute Ventilation

Minute ventilation is the amount of air exhaled in one minute. It is found as the product of respiration rate and tidal volume. At rest, the minute ventilation is around 5-6 L/min. During mild exercise, this can increase to 75 L/min while during maximal exercise values up to 160 L/min occur. For endurance athletes, the minute ventilation may increase to as much as 27 times the resting value (Robergs and Roberts, 1997). As long as exercise intensity is not too high, the minute ventilation will reach a steady state. Because minute ventilation is related to oxygen consumption, an increase in the oxygen consumption due to oxygen drift or the slow component would cause a concomitant increase in the minute ventilation. During exercise with a progressive work rate, below the anaerobic threshold, minute volume increases linearly with oxygen consumption. Above the anaerobic threshold, minute volume increases exponentially (Martin and Weil, 1979).

For constant rate work below the anaerobic threshold, minute ventilation reaches a steady-state (Wasserman et al., 1980). For exercise above the anaerobic threshold, the time to reach steady state is prolonged. For very heavy exercise, a steady state may not be reached before the subject has to cease exercise.

When constant rate work below the anaerobic work begins from rest, there is an initial abrupt rise in minute ventilation (Whipp et al., 1982; Johnson, 1991). The abrupt rise is thought to be neurogenic in nature (Johnson, 1991; McArdle et al., 1996). There may be a short duration plateau (20 seconds) immediately after the abrupt rise. Minute ventilation then increases exponentially to a steady state if the exercise is not too intense (McArdle et al., 1996). The steady state value attained depends on the intensity of exercise. If the work rate is very high, a steady state will not be achieved and the minute ventilation will increase progressively until the person ceases exercise (Wasserman et al., 1980).

There is a large variability in the response of minute ventilation, and other respiratory parameters, to exercise. In fact, Johnson (1991) states that "respiratory responses are difficult to reproduce" and recommends that applications to individuals be made with caution. The variability of the minute ventilation response is less when related to carbon dioxide production instead of oxygen consumption (Wasserman et al., 1980). This indicated the importance of carbon dioxide in the control of respiration (Johnson, 1991).

At low levels of exercise, increases in minute ventilation are brought about mainly by an increase in tidal volume, while at higher intensity levels, minute ventilation increases as a result of increased respiration rate (Johnson, 1991; McArdle et al., 1996).

Factors That Affect Minute Ventilation

Age, training, and gender affect minute ventilation. Maximal minute ventilation decreases with age. Additionally, for a given submaximal oxygen consumption (e.g., 2 Lpm), older subjects will have a higher minute ventilation than younger subjects (Robergs and Roberts, 1997). Training results in a higher maximal minute ventilation during maximal exercise. During submaximal exercise, there is a reduction in the minute ventilation at a particular oxygen uptake after training. This indicates a lower oxygen cost of exercise for breathing (McArdle et al., 1996). Because minute ventilation is related to body mass, male subjects generally have higher minute ventilations than female subjects (Johnson, 1991).

Tidal Volume

Tidal volume is the amount of air exhaled with each breath. While some authors (McArdle et al., 1996; Robergs and Roberts, 1997) define tidal volume as the volume of air either inhaled or exhaled, these two volumes are not the same. The difference results mainly from the different temperatures of the inhaled and exhaled air. The different water vapor addition and different gas composition are smaller factors (Johnson, 1991).

Tidal volume varies with age, gender, and size (McArdle et al., 1996). Males generally have larger tidal volumes than females. An average resting tidal volume for men is 600 mL while that for a woman is 500 mL. During exercise, tidal volume can reach values of 2 – 3 L. Tidal volume can be quite variable even when the subject is at steady state (Johnson, 1991). The interbreath variation is caused predominantly through changes in the inspiratory time.

During exercise, tidal volume is increased by using parts of the inspiratory and expiratory reserve volumes. These volumes are the amount of air present in the lungs after a normal inhalation or exhalation. At low intensity exercise, the tidal volume increases causing an increase in the minute ventilation. Once the tidal volume reaches 50 – 60% of the vital capacity, the minute ventilation is further increased through an increase in the respiratory rate (Wasserman et al., 1999). Vital capacity is the sum of the inspiratory reserve volume, expiratory reserve volume, and tidal volume. Maximum tidal volume has been reported to range from 45 – 58% of vital capacity (Wasserman et al., 1999).

For constant rate exercise below the anaerobic threshold, the tidal volume is relatively constant with time. For exercise above the anaerobic threshold, the tidal volume may decrease slightly with time (Wasserman et al., 1980).

Exhalation and Inhalation Times

The prediction of inhalation and exhalation times can be accomplished using different approaches. Caretti et al. (1992) investigated the effects of exercise modality on breathing patterns. Subjects exercised on a bicycle ergometer and a treadmill. Other investigators have average consecutive breaths with different breathing frequencies and then evaluated the inhalation and exhalation times. Caretti et al. (1992) examined individual breathing frequencies and inhalation and exhalation times. Their rationale was that when consecutive breaths were averaged, the variability in breathing patterns and timing differences related to breathing frequency was masked. Individual breathing frequencies were grouped together into bins to aid in the analysis.

The authors plotted inhalation and exhalation time versus breathing frequency. A regression curve was not fitted to the data, but a the relationship was observed to be similar to a power-law relationship. This relationship was qualitatively similar for both treadmill and bike exercise except for respiration rates below 12 breaths/min. Below 12 breaths/min, the investigators found that exhalation time was significantly longer for treadmill exercise compared to bike exercise. A large variability in inhalation and exhalation times was observed for breathing frequencies below 18 breaths/min. Above 18 breaths/min, the variability decreased.

So, Caretti and Whitley (1998) showed that inhalation and exhalation times could be predicted from respiratory rate. Johnson and Masaitis (1976) took a different approach.

By minimizing total respiratory work during a complete respiratory cycle, Johnson and Masaitis (1976) derived an equation to predict the ratio of inhalation time to exhalation time:

$$\tau^3 - \left(\frac{\lambda}{1+\mu}\right)\tau - \left(\frac{\mu\eta}{1+\mu}\right) = 0 \tag{32}$$

where:

a) $$\tau = \left(\frac{t_i}{t_e}\right) \tag{33}$$

where τ = inhalation time/exhalation time ratio, dimensionless

t_i = inhalation time, seconds

t_e = exhalation time, seconds

b) $$\lambda = \left(\frac{K_{1i}}{K_{1e}}\right) \tag{34}$$

where λ = ratio of first inhalation and exhalation Rohrer coefficients, dimensionless

K_{1i} = first Rohrer coefficient for inhalation, (cm H_2O·sec)/L

K_{1e} = first Rohrer coefficient for exhalation,

(cm H_2O·sec)/L

c) $$\eta = \left(\frac{K_{2i}}{K_{2e}} \right) \quad (35)$$

where η = ratio of first inhalation and exhalation Rohrer coefficients, dimensionless

K_{2i} = second Rohrer coefficient for inhalation,

(cm H_2O·sec)/L

K_{2e} = second Rohrer coefficient for exhalation,

(cm H_2O·sec)/L

d) $$\mu = \left(\frac{2K_{2e}V_T}{K_{1e}t_e} \right) \quad (36)$$

where μ = dimensionless ratio

V_T = tidal volume, L

The Johnson and Masaitis (1976) model assumes: "1) inhalation/exhalation times are determined by respiratory work during one cycle; 2) expiratory work is important in determining inhalation/exhalation times; 3) energy stored during inhalation due to respiratory system compliance or inertance is fully recovered during exhalation." Equation 32 is a cubic equation and can be solved using a method such as Cardan's solution (Korn and Korn, 1961). Inhalation and exhalation times were determined using an iterative process. The authors showed that the model had good qualitative and quantitative agreement between calculated and experimental results.

Effect of Inspiratory and Expiratory Loading

Resistance loading of the respiratory system causes changes in the respiration rate and in the duration of inhalation and exhalation. Inspiratory loading leads to an increased inhalation time and a decreased respiration rate. The subsequent exhalation is affected as well, with an increased exhalation time following an increased inhalation time (Cherniack and Altose, 1981). Expiratory loading leads to increased exhalation times and decreased respiration rates. These effects were shown by Caretti and Whitley (1998), Johnson et al. (1999) and Caretti et al. (2001).

The effect of inspiratory resistance breathing on respiratory rate was investigated by Caretti and Whitley (1998). Subjects exercised on a treadmill at 80-85% of V_{O2max} while wearing a half-respirator with one of four inspiratory resistances ranging from 0.2 kPA to 0.49 kPa, measured at a steady airflow rate of 1.42 m/s. Treadmill speed and grade were adjusted for each resistance condition so that the subject was at 80-85% of V_{O2max}. Respiratory rate decreased from the control condition 4.6%, 10.7%, 16%, and 32% for the four resistance conditions. The respiratory rate for R4 (highest resistance) was significantly different from the control, R1, and R2 conditions.

Johnson et al. (1999) investigated the effects of inspiratory resistance on work performance. Subjects exercised on a treadmill at constant speeds and grades that were chosen to elicit respiratory stress. A full-facepiece respirator was worn for each of six tests with different levels of inspiratory resistance. The inhalation resistances ranged from 0.78 to 7.64 cm $H_2O \cdot sec/L$. The exhalation resistance for all tests was

1.3 cm $H_2O \cdot sec/L$. It was found that minute volume decreased as inhalation resistance increased:

$$\dot{V}_{min} = -.0687R + 1.325 \tag{37}$$

where: $\dot{V}_{min}$, minute volume, L/sec

R, resistance, cm $H_2O \cdot sec/L$

Caretti et al. (2001) conducted a similar study investigating the effects of exhalation resistance on work performance. Exhalation resistances ranged from 0.27 to 27.35 cm $H_2O \cdot sec/L$. Average minute volumes decreased as expiratory resistance increased:

$$\dot{V}_{min} = -1.76R + 73.16 \tag{38}$$

So, in both cases (Johnson et al., 1999; Caretti et al., 2001), as resistance increased, minute volume decreased. Caretti and Whitley (1998) found that tidal volume did not change with resistance for exercise at 80-85% of V_{O2max}. So, assuming a constant tidal volume during steady state work, a decrease in the minute volume would lead to a decreased respiratory rate.

Oxygen Deficit

Because the oxygen consumption does not rise immediately to the steady state value, there is a difference between the oxygen required by the body (the steady state value) and the actual oxygen consumption (that during the exponential increase). This difference is termed the oxygen deficit. The oxygen deficit is found as the product of steady-state oxygen consumption and the time constant of the exponential rise (Whipp et al., 1982). During the deficit, mitochondrial respiration is supplemented through energy generated by creatine phosphate and glycolysis (Robergs and Roberts, 1997). The increase in oxygen consumption due to the slow component means that the oxygen deficit as a percentage of the total oxygen required increases as the work load increases above the anaerobic threshold (Whipp and Wassermnan, 1972).

Whenever there is a difference between the actual and required oxygen consumption, there is a deficit. As a trained person will reach steady state faster than an untrained individual, the trained person incurs less of an oxygen deficit. Performing a warm-up can also decrease the oxygen deficit (Robergs and Roberts, 1997).

When a steady-state can be reached, the oxygen deficit is the difference between steady state and non-steady state oxygen consumption. However, a deficit may occur also when a respirator is worn. Respirators have been shown to cause hypoventilation (Johnson et al., 1999; Caretti et al., 2001), so a respirator wearer has a lower minute ventilation and thus a lower oxygen consumption than that required by the body. The greater the deficit, the shorter the performance time.

Respiratory Work Rate

"It has long been assumed that respiration is physiologically adjusted to yield optimum respiration ratio, ratio of inhalation time to exhalation time, expiratory reserve volume, dead volume, airways resistance, and airflow waveshape (Johnson, 1993)." These adjustments are especially important during exercise when there is a competition among the skeletal, cardiac, and respiratory muscles for the limited oxygen available. Because of the limited oxygen supply and the fact that respiratory work does not contribute to the activity being performed, it is logical that respiratory work should be minimized during exercise. Data taken during exercise support this contention (Johnson, 1993).

At rest, respiratory work accounts for 1-2% of the total oxygen consumption (Johnson, 1991). This increases up to 10% during exercise. Changes in airflow waveshape could have a significant effect in a model of respiratory work. Indeed, Yamashiro and Grodins (1971) found a 23% lower work rate for a rectangular waveshape compared with a sinusoidal waveshape. They used a simple model that had only had one resistance and one constant compliance.

Respiration occurs with different flow patterns that depend on exercise intensity. At rest, inhalation has a sinusoidal waveshape while exhalation occurs with an exponential waveshape. Both inhalation and exhalation waveforms are trapezoidal with rounded corners during moderate exercise. Inhalation waveforms remain trapezoidal during heavy exercise, but exhalation waveforms return to an exponential shape.

At rest, both inhalation and exhalation waveshapes appear to be unrelated to work rate. Yamashiro and Grodins (1971) found that the sinusoid resulted from a mean squared acceleration criterion. They reasoned that the sinusoidal waveshape resulted in improved gas transport efficiency and a uniform ventilation of the lungs.

The resting exponential exhalation waveshape is due to passive exhalation. There is little muscle activity required during exhalation at rest. The energy comes instead from elastic energy stored in the chest wall, which is expanded during inhalation. Additional energy comes from air that is compressed in the lungs during inhalation.

During moderate exercise, both inhalation and exhalation are active. The trapezoidal waveshapes appear to be related to respiratory work rate, although they differ from the rectangular waveshape that minimizes respiratory work (Yamashiro and Grodins, 1971; Johnson and Masaitis, 1976). Both Hamalainen and Sipila (1984) and Ruttiman and Yamamoto (1972) gave possible reasons for the trapezoidal shape. Hamalainen and Sipila (1984) got a trapezoidal waveform when they included an additional term in their optimization criteria that is equal to the square of muscular pressure times the volumetric flow rate. This term accounts for the decreased muscular efficiency seen at higher loads. Ruttimann and Yamamoto (1972) also obtained a trapezoidal waveform, although the slope was in the opposite direction. Their waveform resulted when they minimized respiratory work while using an airways resistance that increased as volume decreased. Johnson (1986) found that a part of lower airways resistance has this inverse effect. The reason for the rounded corners may be that rapid accelerations are penalized to avoid damage or loss of control (Johnson, 1991; Johnson, 1993). Or, the rounded corners may indicate that the strength of the respiratory muscles is limited (Johnson, 1993).

The same inhalation optimization criteria during moderate exercise is in effect during heavy exercise (Johnson, 1991). Thus, the waveshape remains trapezoidal. The exhalation waveshape returns to exponential although the reason for the exponential waveform differs from that at rest. During heavy exercise, exhalation flow rate is limited. Johnson and Milano (1987) plotted transpulmonary pressure against expiratory flow rate along lines of equal lung volume. They found that a point was reached beyond which the flow could not be increased. The limiting flow rate was inversely related to the lung volume. The very abrupt transition to the exponential waveform only occurs during a maximal effort when the respiratory system is extremely taxed (Johnson and Milano, 1987; Johnson, 1993). Because flow rates and respiratory muscle pressure were so high, much more energy was required by the exponential waveform (Johnson, 1993).

There is one other characteristic of the moderate and heavy exercise waveforms that needs to be discussed. There are dimples that often appear in the waveforms. The reason for these dimple is not clear (Johnson, 1991). However, when minimizing the Hamalainen and Viljanen (1978) inhalation optimization criteria, the dimples appear in the waveform under certain conditions (Johnson, 1991).

Respiratory Work Rate Model

The work rate of breathing with different waveshapes was investigated by Johnson (1993). The model of the respiratory airways that was used contained a small number of elements with nonlinearities resulting from the airways and mask (Johnson, 1992). The model used the modified Rohrer equation:

$$p = K_1 \dot{V} + K_2 \dot{V}^2 + \frac{K_3 \dot{V}}{V} + \frac{V - V_r}{C} + I \ddot{V} \tag{39}$$

where p = respiratory muscle pressure, N/m^2

$\dot{V}$ = respiratory flow rate, m^2/sec

V = lung volume, m^3

$\ddot{V}$ = volume acceleration, m^3/sec^2

V_r = resting volume of the lung, m^3

K_1 = first Rohrer coefficient for the respiratory system, N·sec/m^5

K_2 = second Rohrer coefficient, N·sec^2/m^8

K_3 = “third” Rohrer coefficient, N·sec/m^2

C = respiratory compliance, m^5/N

I = respiratory inertance, N·sec^2/m^5

This model was sufficient for both inhalation and exhalation if different values were used for the parameters K_1 to K_3 , C, and I (Johnson, 1992). An extra term must be added when flow rate nears maximum exhalation flow rate (Johnson, 1993):

$$p_e = K_1 \dot{V} + K_2 \dot{V}^2 + \frac{K_3 \dot{V}}{V} + \frac{V - V_r}{C} + I \ddot{V} + \frac{K_4}{\left(1 - \frac{\dot{V}}{\dot{V}_L}\right)} \qquad (40)$$

where p_e = respiratory muscle pressure at airflow limitation, N/m^2

K_4 = additional coefficient, N/m^2

$\dot{V}_L$ = limiting flow rate, m^3/sec

Waveshapes. Johnson (1993) developed the equations for the respiratory work rates when breathing with a sinusoidal, rectangular, truncated exponential, hybrid exponential, and trapezoidal breathing pattern. Linear, quadratic, volume dependent, compliant, and inertial pressure terms were included.

Variable Lung Volume. The expiratory reserve volume changes during exercise thus changing the initial lung volume. The correct lung volume needs to be included in the volume dependent and compliant work rate terms. For the volume dependent term, the correct volume is simply inserted into the formula. For the compliant term, it was not necessary to change the equation as long as exhalation was active and the whole breathing cycle was considered. This was because the added term would be the same magnitude but opposite sign for inhalation and exhalation, thus canceling its effect.

Maximum Expiratory Flow. Expiratory flow rate can become limited during maximal exertion. This limitation can cause respiratory distress and early termination of exercise for people wearing respirators (Johnson and Berlin, 1974).

A term for maximum respiratory rate of work must be added to the limited flow hybrid exponential work rate equations (Johnson, 1993):

$$\dot{W}_R(6) = \frac{1}{T}\int_0^T p_{\max} \dot{V}_L dt \tag{41}$$

where $\dot{W}_R(6)$, average respiratory work rate during flow limitation,

N·m/sec

T, duration of waveform, sec

t, time, sec

Effect of Waveshape on Respiratory Work Rate. The work rate while breathing with each of the five waveshapes was investigated during rest and light, moderate, heavy, and very heavy exercise. The lowest work rates occurred with the rectangular waveform. Comparisons were made to the rectangular waveform. The increased cost of the sinusoid for inspiration ranged from 9% at light exercise to 16% at very heavy exercise. The inspiratory trapezoid had an increased cost of 3% at light exercise and 7% during heavy exercise. The truncated exponential costs 30% more at light exercise and 9% during heavy exercise for inspiration. Finally, the hybrid exponential for inspiration was 29% higher for light exercise and 12% higher during very heavy exercise. For exhalation, the work rates were lower than for inhalation because of longer inhalation times.

Waveform Transition

Little work has been done on the transition between waveshapes during exercise. This is important, because as shown in Johnson (1993), the work rate is dependent on the breathing waveform.

Hamalainen and Viljanen (1978) developed a model of the control of the breathing pattern during respiration based on optimization criteria. The performance criteria were chosen to minimize the oxygen cost of breathing. Both criteria have an average square of volume acceleration term. The inspiratory criterion is the weighted sum of that term and the mechanical work performed by the inspiratory muscles. The expiratory criterion includes an integral square driving pressure in place of the mechanical work term.

For inhalation, the authors found that when the ratio of pressure times flow to the square of volume acceleration became large, a transition occurred from a sinusoidal to a trapezoidal waveform. Similarly for exhalation, when the ratio of pressure squared to volume acceleration squared became large, the waveshape changed from exponential to trapezoidal.

Their method is not practical for this model because the weighting functions, α_1 and α_2, are specific to the individual being tested and have no known physiological basis. The authors noted that different alpha parameters made sense because “the airflow patterns of any given individual look as unique as fingerprints (Hamalainen and Viljanen, 1978).” But, this means that each person must be tested and the actual breathing waveforms compared to the predicted waveforms. The weighting functions are adjusted until the differences between the two sets of waveforms are minimal. A better means is necessary to determine when transitions in the respiratory waveforms occur.

Respiratory Protective Masks

Respiratory protective devices have a profound impact on the wearer. Vision, communications, and personal support (wiping of nose, drinking) are all hindered. Problems occur due to sweat accumulation inside the mask and reduced heat loss through the mask. Sore neck muscles and skin irritation become a concern with extended wear. The physical characteristics of the respirator, the inspiratory and expiratory resistance, the dead volume, and the weight, affect the physiological response and impede performance. The influence of each of these factors depends in part on the work intensity and the type of task. Other important factors to consider are variability, anxiety, and hypoventilation.

Physical Characteristics

Resistance. A person wearing a mask must overcome the resistance to breathing caused by the filter and the inspiratory and expiratory valves in the mask. A number of studies have investigated the effects of external resistance on pulmonary function.

Flook and Kelman (1973) investigated the effects of increased inhalation resistance on seven subjects exercising on a bicycle ergometer for ten minutes at 35, 50, and 70% of V_{O2max}. The inhalation resistances were 8.9, 16.5, and 53.1 $cmH_2O/L/s$ measured at a steady flow of 1 L/s. These resistances were chosen to represent resistances seen in patients with pulmonary disease. Regression equations fit to their data showed that minute ventilation decreased with increased resistance.

The slope coefficients for these equations for work done at 35, 50, and 70% V_{O2max} were –0.0023, -0.005, and –0.0214, respectively. Regression equations fit to the tidal volume data indicated that at 35%V_{O2max}, the tidal volume increased with increased resistance while tidal volume was virtually unaffected by resistance at the other two work rates (r = 0.05 and r = 0.005). The slopes of these equations in order of increasing work rates were 0.0078, 0.0011, and 0.0009, respectively.

The effects of three inhalation resistances on subjects performing steady-state bicycle exercise was investigated by Demedts and Anthonisen (1973). Exercise periods lasted five minutes if possible or three minutes when the work load could not be tolerated for the full five minutes. The work loads were 82, 131, 196, 245, and 270 W. The resistances read off a pressure-flow graph at approximately 1.4 L/s were 1.6, 3.1, and 12.4 cmH_2O/L/s. The dead space for all conditions was 350 mL. The authors found that minute ventilation was not decreased by the lowest resistance. A statistically significant 12% decrease occurred for the middle resistance at the highest work load while the highest resistance caused a 50% decrease at the higher work rates.

Silverman et al. (1951) investigated the effects of two combinations of inhalation and exhalation resistance on 18 healthy males during bicycle exercise at constant rates of 0, 34, 68, 102, 136, 181, 226, and 271 W. Not all subjects completed all conditions. Data were recorded at six, eight, and ten minutes into the exercise. The inhalation and exhalation resistances were 0.4 and 0.2 cmH_2O/L/s for the low condition, and 4.5 and 2.9 cmH_2O/L/s for the high condition. A third condition was tested at the 68 W work rate only. The inhalation and exhalation resistances for this condition were 4.5 and 1.9 cmH_2O/L/s. The authors found that the minute ventilation was reduced almost 20% at the highest two work rates. The authors stated that the resistance used did not affect tidal volume at work rates below 181 W. The percent change in the minute ventilation, respiratory rate, and tidal volume from the low to high resistance conditions was determined at each work rate. Tidal volume was determined by dividing the mean minute ventilation by the mean respiratory rate. The results are shown in Table 2.

Table 2. Percent changes in minute ventilation, respiratory rate, and tidal volume from the low to high resistance conditions. Data are from Silverman et al. (1951).

	V_E (L/s)	RR (b/s)	V_T (L)
	% change	% change	% change
Rest	-13.2	1.4	-14.7
0	-7.6	-12.0	4.0
34	-5.1	-13.4	7.3
68	-10.7	-9.7	-1.0
102	-3.0	-2.2	-0.8
136	-11.9	-10.9	-0.8
181	-16.9	-7.1	-9.2
226	-27.9	-19.0	-7.5
271	-26.0	-13.3	-11.2

It can be seen from the table that tidal volume was affected at low work rates. In fact, the tidal volume increased at the two lowest work rates. At work rates of 68, 102, and 136 W, the tidal volume does not appear to be affected by the resistances used. At work rates above 181 W, the tidal volume decreased with added resistance.

Cerretelli et al. (1969) assessed the effects of two resistances on two subjects during treadmill exercise at work rates ranging from about 70 to about 210 cal/kg min. Subjects inhaled and exhaled against the same two resistances of 8.5 and 16.9

$cmH_2O/L/s$. The minute ventilation for the two subjects decreased at all work rates as the resistance increased.

Hermansen et al. (1972) investigated the effects of a respirator mask and breathing valve on minute ventilation and tidal volume on ten healthy subjects performing on a bicycle ergometer at work rates of 49, 98, 147, and 196 W. The inhalation and exhalation resistances of the mask were 9 and 2.6 $cmH_2O/L/s$, respectively while those of the valve were 1.7 and 1.7 $cmH_2O/L/s$.

Minute ventilation was always lower with the mask than with the valve. At the highest work load, the decrease in minute ventilation was 43%. Tidal volume was greater with the mask up to a minute ventilation of approximately 70 L/min. After that, tidal volume decreased with added resistance.

The effect of inspiratory resistance on breathing parameters was investigated by Caretti and Whitley (1998). Subjects exercised on a treadmill at 80-85% of V_{O2max} while wearing a half-respirator with one of four inspiratory resistances ranging from 0.2 kPA to 0.49 kPa, measured at a steady airflow rate of 1.42 m/s. Treadmill speed and grade were adjusted for each resistance condition so that the subject was at 80-85% of V_{O2max}.

Tidal volume was shown to be relatively constant across the respirator conditions. No significant differences among the conditions were found. The differences from the control condition were +1%, 0%, +1.1%, and –2.7% for the R1, R2, R3, and R4 conditions respectively. However, minute ventilation decreased as resistance increased. The differences were significant between the control and R4 conditions. The decreases in minute ventilation from the control condition were 2.4%, 9.8%, 14.9%, and 35.4% for the R1, R2, R3, and R4 conditions respectively.

Johnson et al. (1999) quantified the effect of increased inhalation resistance on minute ventilation. Twelve subjects exercised at 80-85% V_{O2max} until their volitional end-point while wearing a U.S. Army M-17 respirator with one of six different inhalation resistances. Plugs with different size holes bored through the center were placed in the inhalation ports to modify the resistance. The inhalation resistances were 0.78, 1.64, 2.73, 3.32, 6.47, and 7.64 cm H_2O/L/s at a flow of 1.42 L/s (85 L/min). The exhalation resistance for all tests was 1.3 cm H_2O/L/s. The relationship between minute volume and inhalation resistance was found to be:

$$V_E = -0.0687 \cdot R_{inh} + 1.325 \tag{42}$$

where: V_E, minute volume, L/s

R_{inh}, inhalation resistance, cmH_2O/L/s

A similar study was conducted to examine the effect of increased exhalation resistance on work performance and ventilation (Caretti, et al., 2001). Subjects wore a U.S. Army M40 respirator with one of five exhalation resistances while exercising on a treadmill at 80-85% V_{O2max}. The exhalation resistances were 0.47, 1.81, 4.43, 12.27, and 27.35 cm H_2O/L/s. The inhalation resistance for all conditions was 3.17 cm H_2O/L/s. Lower minute volumes were found for increasing exhalation resistance:

$$V_E = -0.0299 \cdot R_{exh} + 1.2365 \tag{43}$$

where: V_E, minute volume, L/s

R_{exh}, exhalation resistance, $cmH_2O/L/s$

So, in both cases (Johnson et al., 1999; Caretti et al., 2001), as resistance increased, minute volume decreased. The effects of the inhalation resistance were three times that of the exhalation resistance (Caretti et al., 2001).

The above studies indicated that at all work rates, inhalation and exhalation resistance caused a decrease in minute ventilation. Only one study contradicted this. Demedts and Anthonison (1973) found that minute ventilation was not decreased at their lowest resistance.

Flook and Kelman (1973), Hermansen et al. (1972), and data from Silverman et al. (1951) indicated that at low work rates tidal volume was increased by resistance. Resistance at higher work rates has been reported to not have an effect on tidal volume (Flook and Kelman, 1973; Caretti and Whitley, 1998) or to decrease tidal volume (Silverman et al., 1951; Hermansen et al., 1972).

In addition to increasing the inhalation and exhalation resistance, the valves also require an additional amount of pressure to open the valves. Cummings (1968) investigated the pressures required to open the valves in an M17 mask. The inspiratory pressure was found to be:

$$p_i = 3.227x10^5 \dot{V} + 5.609x10^7 \dot{V}^2 \qquad (44)$$

where p_I, inspiratory pressure inside the mask, N/m^2

$\dot{V}$ = flow rate, m^3/sec

The expiratory pressure for the same mask was (Cummings, 1968):

$$p_e = 59.93 + 6.629x10^4 \dot{V} + 1.376x10^7 \quad (45)$$

where p_e, expiratory pressure inside the mask, N/m^2

$\dot{V}$ = flow rate, m^3/sec

The constant term in the p_e equation is the pressure needed to open the valve. This results in an addition to the respiratory work rate (Johnson, 1992):

$$\dot{W}_R(7) = 0.05 p_o \dot{V}_{\max} T(1 + e^{-0.8T/\tau}) \quad (46)$$

where $\dot{W}_R(7)$, respiratory work rate due to constant pressure term, W

$p_{o,}$ constant term, N/m^2

$\dot{V}_{\max}$, maximum flow rate during breathing waveform, m^3/sec

T, waveform duration, sec

τ, respiratory time constant, sec

Dead Volume. Dead volume, or dead space, is the amount of air present that does not take place in respiration, including air in the nasal passages and throat. This volume is increased when an object, such as a snorkel, mask, or breathing tube, is placed over the mouth and/or nose. Carbon dioxide accumulates in the dead volume, causing it to act as a respiratory stimulant.

As airflow increases, so does dead volume. This occurs because when the flow rate increases, the airflow becomes more turbulent, causing a greater mixing of gases. Thus, air that was trapped at corners and around objects becomes mixed with the airflow, increasing the dead volume. The volumetric space inside a respirator is termed the nominal dead volume while dead space as a function of tidal volume is termed effective dead volume.

Breathing through an external dead volume causes a performance decrement. Johnson, et al. (2000) investigated this effect by having subjects walk on a treadmill at 80-85% VO2max with respirator configurations giving a range of dead volumes. While performance time was affected, no effect of dead volume on minute ventilation, tidal volume, or oxygen consumption at termination was found.

Stannard and Russ (1948) studied the effects of increasing dead volume on minute ventilation and tidal volume for seven subjects at rest and during light exercise. The light exercise was chosen as the work rate at which the resting oxygen consumption doubled. No indication of V_{O2max} was given. Nominal dead spaces of 250, 350, 420, 450, and 540 mL were used.

At rest, the tidal volume increased as dead volume increased. During light exercise, tidal volume increased with added dead volume, but the changes were smaller. For the lowest dead volume, the change in tidal volume was not significant.

The minute ventilation increased with added dead volume for resting and lightly exercising subjects. The authors noted that the regression lines fit to the data had similar slopes. The near constant difference between the two lines was reported to be approximately 2 L/min.

In 1980, Ward and Whipp studied the effects of dead volume on minute ventilation of three subjects. The authors concluded that minute ventilation increased during rest and moderate exercise as a result of added dead space.

The three studies noted above only looked at rest, light exercise, and heavy exercise. Harber and colleagues have completed a number of studies in which they investigated the effects of inhalation resistance and dead volume on breathing parameters at rest and during moderate exercise. Unfortunately, most of their information can not be used in a model. In one study (Shimozaki et al., 1988) only subjective responses were reported. In another study (Harber et al., 1982) subjects were allowed to pick their own work rate so that it was consistent with long-term work. Finally, three studies (Harber et al., 1984; Harber et al., 1988; Harber et al., 1990) were conducted in which one load, a combination of inhalation resistance and dead volume, was applied. The effects of the resistance and dead volume on the breathing parameters could not be separated.

Mass and Load Placement. The mass of the mask will increase the external work rate. The equation developed by Pandolf et al. (1977) and the external work rate equation presented by Aoyagi et al. (1995) included total mass (body mass plus load mass) in the calculations. If the external work is specified and not calculated, the external work rate without the mask will be increased. The increase will equal the percentage increase in mass represented by the mask. Thus, a typical mask has a mass equal to 1.4% of the normal body mass of a man. The work rate for that mask would be increased by 1.4% to account for the added mass of the mask.

The respirator mass is not distributed evenly over the head. An eccentricity factor takes into account this fact.

Other Factors

Variability. The variability in response to respirators wear across the population underscores the necessity of using large sample sizes in conducting studies and in calibrating and validating models. The study by Johnson et al. (1999) showed that three of the twelve subjects were not sensitive to inspiratory resistance and indeed showed little performance decrement. A study examining the effects of exhalation resistance (Johnson, et al., 1997) found that three of ten subjects could perform no treadmill work when the resistance was very high, but that the other seven were able to perform for two to ten minutes. Finally, the performance of subjects who scored an anxious rating on the Spielberger State-Trait Anxiety Test was dependent on the numerical score, while those classified as non-anxious had performances unrelated to their score.

Anxiety. Psychological factors can play a large role in whether or not a person can tolerate respirator wear. To determine the amount of influence such factors have, Johnson, et al. (1995) conducted a exercise study in which subjects took the Spielberger State-Trait Anxiety Inventory (STAI) to assess their anxiety level.

Twenty subjects exercised at 80-85% of their age-predicted maximum heart rate until their volitional end-point. The performance times of subjects classified as non-anxious (STAI scores less than 34) were unrelated to the STAI score. However, for anxious subjects, the performance rating was related to the anxiety score. Someone with a STAI score of 40 would suffer a 25% decrement in performance. A highly anxious person (STAI score of 70) would have a 79% decrement and would therefore only achieve a 21% performance rating.

Hypoventilation. Hypoventilation is a condition in which the subject is breathing at a lower minute volume than normal. This may be due to either more shallow breaths or less frequent breaths, or both. The hypoventilating person must extract more oxygen from each breath as the oxygen requirements of the body are unaffected by the decreased minute volume. As less air is exhaled during hypoventilation, the carbon dioxide concentration in the exhaled air must increase. Thus, high concentrations of carbon dioxide and low concentrations of oxygen in the exhaled air indicate that a person is hypoventilating. Hypoventilation has been evident in two respirator studies. The first study (Johnson, et al., 1995) involved incremental bicycle exercise while wearing an M17 respirator. Hypoventilation was indicated by high F_{ECO2} and low F_{EO2} values during respirator wear.

Subjects participating in a study on the effects of inspiratory resistance on performance time also evidenced hypoventilation (Johnson et al., 1999). Subjects had decreasing minute volumes and oxygen consumption as resistance increased.