
On the reliable assessment of cardiovascular recovery: An application of curve-fitting techniques

NICHOLAS CHRISTENFELD,^a LAURA M. GLYNN,^b AND WILLIAM GERIN^{c,d}

^aDepartment of Psychology, University of California, San Diego, USA

^bDepartment of Psychology, University of California, Irvine, USA

^cDepartment of Psychology, Cornell University Medical Center, New York, New York, USA

^dNational Institutes of Health, Office of Behavioral and Social Sciences Research, Bethesda, Maryland, USA

Abstract

The underuse of cardiovascular recovery as an adjunct to reactivity may stem from a lack of research on how to assess the process reliably. We explore the test–retest reliability of three simple, intuitive approaches to measuring recovery, and of a more sophisticated curve-fitting technique. Eighteen young normotensive subjects experienced three stressors twice each, with 10-min baseline, 3-min task, and 20-min recovery periods and continuous monitoring of heart rate and blood pressure. Reactivity showed moderate reliability, but the three simple approaches to measuring recovery revealed essentially none. However, the curve-fitting approach, using a three-parameter (amount, speed, and level of recovery) logistic function was reliable. This approach, capturing the inherently dynamic process of cardiovascular recovery, may allow researchers to usefully add the assessment of recovery to paradigms exploring reactivity as a risk factor for cardiovascular disease.

Descriptors: Cardiovascular recovery, Reactivity, Logistic equation, Curve fitting, Cardiovascular disease

There has been recently a renewed interest in examining not just cardiovascular reactivity, or how aroused people become by particular stressors, but also recovery, or how much time it takes people after the stressor to return to resting levels (Hocking Schuler & O'Brien, 1997; Linden, Earle, Gerin, & Christenfeld, 1997). Several studies suggest that cardiovascular recovery may be of some clinical utility. For example, delayed recovery has been implicated as a risk factor for the development of cardiovascular disease (Haynes, Gannon, Orimoto, O'Brien, & Brandt, 1991) and hypertension (Borghi, Costa, Boschi, Mussi, & Ambrosini, 1986; Franz, 1986; Gerin, & Pickering, 1995). In addition, it has been shown that people who are physically fit show the same reactivity to psychological stressors as those who are not fit, but recover from the stressors more rapidly (Jamieson & Lavoie, 1987; McCubbin, Cheung, Montgomery, Bulbulian, & Wilson, 1992). Delayed recovery from laboratory stressors has also been linked to anger, both as an individual difference (Faber & Burns, 1996; Lai & Linden, 1992), and as a situational characteristic (Earle, Linden, & Weinberg, 1999; Lai & Linden, 1992).

Despite positive findings using measures of recovery, many studies of cardiovascular reactivity do not include recovery periods. Of those that do include such periods, only a fraction mention having done any analyses on these data, and fewer still report having found anything (Linden et al., 1997). One likely explana-

tion for the discrepancy between the promise of recovery and the practice of laboratory reactivity studies is that there is no guiding empirical work on which measures of recovery are reliable. The purpose of this study was to explore the reliability of a variety of measures of cardiovascular recovery, and to propose ways of assessing recovery that are likely to reveal relationships with other variables.

Demonstrating that recovery can be assessed reliably does not mean that the recovery measures capture critical aspects of the functioning of the cardiovascular system. Answering that question requires studies linking the particular measures of recovery to long-term health outcomes, or to markers for heart disease, such as hypertension or left-ventricular hypertrophy (Wilson, 1997). However, the development of reliable methods of assessing recovery should encourage researchers who examine the role of reactivity in predicting, or causing, later disease to include measures of recovery.

The reactivity hypothesis suggests that people who show exaggerated cardiovascular responses to stress are at risk of later hypertension and cardiovascular disease. The weak version of this hypothesis suggests that reactivity is a marker for disease vulnerability, but the strong, or causal version of this hypothesis suggests that frequent large blood pressure excursions actually damage the circulatory system and can lead to the development of hypertension and coronary artery disease. One implication of this latter view is that it may not be just the magnitude of the cardiovascular response to stress that is important, but also its duration. That is, people who remain elevated for longer periods of time following stress may be at greater risk than those who show identical reactivity but recover promptly. Reliably identifying such people is

We thank John Wixted for curve-fitting advice and inspiration.

Address reprint requests to: Nicholas Christenfeld, Department of Psychology, University of California–San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, USA. E-mail: nicko@ucsd.edu.

thus of considerable potential value. Clearly, measures of recovery that are not stable cannot usefully predict any later disease outcome.

Reliability depends not just on characteristics of the measure, but also on characteristics of the sample tested. Generally, the standard for psychological measures is that they be sensitive to the naturally occurring differences between people. This is the theoretical foundation of the test–retest measure of reliability, which assesses whether relative differences between people are preserved from one testing session to another, or nonparametrically, whether they come out in roughly the same order on the scale. Of course, the range of people in the sample is still a determining factor in the assessed reliability of the measure. If the people are all similar on the critical dimension, the reliability is likely to be small, whereas the same measure will be more reliable when a broad range of people is included. For example, a relatively crude measure of resting blood pressure may show reasonably high test–retest reliability if the sample includes both normotensive and hypertensive individuals. A more sensitive measure would be needed to distinguish between normotensive subjects. However, a sensitive measure is what is required if the measure is to be useful for predicting future disease outcomes. Because of the emphasis in cardiovascular reactivity work on individual difference markers of future pathology, a good measure will be one that is sensitive to differences between currently healthy young people.

In the present study, we measure heart rate and blood pressure recovery following a series of stressful tasks, and then evaluate the test–retest reliability of a variety of ways of quantifying this recovery. Although there are generally accepted ways of assessing resting levels and reactivity, the same is not true for recovery. Because resting and stress levels are each fairly constant states, they can be characterized reasonably by an average reading over appropriate periods. Reactivity is generally assessed by taking the difference between the mean over the stress period and subtracting from that the mean over the prestress resting period. However, recovery is an inherently dynamic process, with no clear end point, and so there is no one obvious way to capture the phenomenon numerically.

In addition to examining the reliability of three standard methods of measuring recovery, we also explored more novel curve fitting approaches. The three standard methods are (1) to measure how much time elapses between the end of the stressor and the parameter's return to prestress levels; (2) to measure the difference between the prestress baseline and the measure at a fixed time following the end of the stressor; and (3) to measure the difference between the average over the entire recovery period and the prestress level. The final measure that we examine is one derived from more sophisticated curve-fitting procedures, in which the parameters of the best-fitting equation are the individual difference measures. We used a logistic curve, which describes the size and rate of transition between two relatively stable levels. Such curves are used frequently to characterize dose–response relationships (Suhnel, 1998; Arcangeli et al., 1998). They have also been used to describe other sorts of transitions between two stable states, such as children's acquisition of grammar (Kemper, Rice, & Chen, 1995), people's awareness of alcohol warning labels (Hankin et al., 1996), and the relationship between latency and magnitude of eye-movements in schizophrenic populations (McDowell, Clementz, & Wixted, 1996).

There are several potential benefits to assessing cardiovascular recovery using a curve-fitting procedure. One is that the curve can capture the dynamic nature of the process. Other approaches, such as measuring how much recovery has occurred at a fixed time

point following the stressor, characterize the process with measurements of a level. Another advantage is that the curve is based on the entire range of data, rather than on a subset of the scores, for example those that are taken 3 min after the end of the stressor. Additionally, a curve-fitting approach can provide multiple parameters that allow the simultaneous assessment of several different aspects of the process, such as the amount and the speed of recovery. For all of these reasons, the curve-fitting-derived parameters have some potential to display reliability greater than that obtained from more conventional means of assessing recovery.

Methods

Participants

Participants were 18 college students (12 women and 6 men), all between the ages of 19 and 24 years. They were all normotensive (resting blood pressure less than 140/90 mmHg), with no history of heart disease. The participants were all research assistants, involved in a variety of other projects. Most were familiar with the blood pressure apparatus and the basic procedures of reactivity experiments.

Recording of Cardiovascular Activity

Blood pressure and heart rate were collected using an Ohmeda Finapres 2300 blood pressure monitor, which takes beat-to-beat pressures in a noninvasive manner, using the Peñáz method (Wesseling et al., 1985). Systolic and diastolic pressures, as well as heart rate, were recorded. The Finapres uses a finger cuff, worn on the third finger of the nondominant hand. This apparatus has been demonstrated as a useful alternative to intraarterial blood pressure measurement in laboratory testing (Imholz et al., 1988; Imholz, Settels, van der Meiracker, Wesseling, & Wieling, 1990) and in clinical practice (Gorback, Quill, & Lavine, 1991; Imholz, Wieling, Langewouters, & van Montfrans, 1991; Wieling, ten Harkel, & van Lieshout, 1991). In addition, the Finapres has been shown to track intraarterial readings extremely well, even during sudden changes of blood pressure (Parati, Casadei, Groppelli, Di Rienzo, & Mancia, 1989), making it a good candidate for use during reactivity and recovery testing. For the entire collection period, the participant was seated with the monitored hand resting at heart level on a table to the participant's side.

Stressor Tasks

Participants experienced each of three different tasks twice. The three tasks each lasted 3 min and were a physical task, a mental arithmetic task, and a speech task. In the physical task participants walked in place. They did this while seated to eliminate postural changes between this and the baseline and recovery periods. They were instructed to lift each foot 8 inches with each step, and this level was marked on the desk next to their seat. They were further instructed to keep their steps synchronized to the beats of a metronome, set at the brisk rate of 150 beats per minute (bpm).

For the mental task, participants did a serial subtraction task in which they subtracted 13s from a large starting number out loud. They were told to go as quickly as possible, but were given no feedback during the task unless they made a mistake, in which case they were informed of the correct answer, and allowed to continue.

For the speech task, participants gave a brief talk on drug legalization. To make the event more stressful, a faculty member served as the audience, and the speech was also recorded with a clearly visible tape recorder. Participants were told they would give a speech, and would have 3 min to prepare their talk. They

were told they would be given the position they would have to support, as well as several arguments that they could use in the talk. For one speech they were asked to argue in favor of drug legalization, and in the other they were asked to argue against this position. In both cases, they did not know the topic of the speech until the start of the preparation period, and the order of the two speeches was counterbalanced across participants. The person serving as the audience remained attentive but neutral during the speech. The same faculty member was present for both speeches.

Procedure

Participants came to the laboratory on six separate occasions, and performed each of the three tasks twice. The two sessions of each task occurred 1 week apart at the same time of day. This schedule was planned to minimize the effects of outside influences, such as school and work schedules. The order in which participants participated in the three different pairs of tasks was determined randomly. Participants were asked not to discuss anything about the tasks with anyone until the experiment was complete.

When the participants arrived, they were seated in a chair adjacent to a small table, and instrumented with the blood pressure cuff. They were told they should sit quietly, with little movement, for the 10-min baseline period. The experimenter left the room during the baseline period, and returned at the end of 10 min. The experimenter then explained the particular task the participant should do, and started the participant on the 3-min task. At the end of this period, the experimenter told the participant that he or she should sit quietly for another 20 min, again without moving. The experimenter left the room for the recovery period. At the end of this period, the blood pressure cuff was removed, and the session was over.

For the speech tasks, the procedure was slightly different. Instead of the regular experimenter returning after the baseline period, the faculty member who was to serve as the audience came into the room. He explained the speech task briefly, gave the participant the sheet with the topic and the possible arguments, and allowed the participant 3 min to prepare. At the end of this period, the experimenter told the participant to begin the speech.

Analyses

Baseline Levels, Task Levels, and Reactivity

Baseline blood pressure and heart rate levels were estimated for each session by averaging over the final 5 min of the 10-min pretask period. Discarding the first 5 min of the baseline period allows for adaptation to occur and thus provides a more stable estimate of the resting level (Manuck, Kasprowicz, Monroe, Larking, & Kaplan, 1989). The task level was calculated as the average over the 3-min stressor period, and the measure of reactivity was the difference between stress and baseline levels. (Averages were computed using the pulse-based technique; Glynn, Christenfeld, & Gerin, 1997.) The test-retest reliability of baseline measures was assessed by computing correlations between scores for the pairs of tasks for each physiological parameter. Thus, for example, we correlated baseline levels during the first session of the physical task with baseline levels during the second session of that task for systolic blood pressure, diastolic blood pressure, and heart rate. Similar analyses were done for the other tasks, and also for reactivity scores.

Recovery Measures

We examined the reliability of four possible measures of recovery. The first three are measures that have been used in pub-

lished studies, and are intuitively appealing. These are time to recovery, recovery at a fixed point, and total carryover. The final measure examined was one derived from more sophisticated curve-fitting procedures, in which the parameters of the best-fitting equation were the individual difference measures.

Time to recovery. How quickly people recover from a stressor can be assessed by the number of seconds that elapse between the time the stressor ends and the time the measure returns to its resting level. Here, we assume that the prestress level is the resting level to which the physiological parameter (systolic blood pressure, diastolic blood pressure, or heart rate) will return. Because the Finapres produces a reading for each heartbeat, we could measure the time that elapses before a single reading is back to baseline. However, single readings tend to be unreliable. Instead, we computed 30-s running averages, and measured the time until the first of these periods was back to baseline. The score is then the number of seconds that elapses from the end of the stressor to the first time that the average over the next 30 s is at, or below, the prestress baseline.

Recovery at a fixed point. A second intuitive way of assessing recovery is to measure the amount of residual arousal at a specific time after the end of the stressor. We used the average for each measure taken over a 1-min period starting 13 min after the stressor ends. The baseline level is subtracted from this value to calculate the residual arousal during this period.

Total carryover. The final intuitive method of calculating recovery that was evaluated involves calculating the total carryover during the entire 20-min recovery period. The carryover is assessed by averaging over the recovery period for each measure, and subtracting from that score the baseline level. This calculation is mathematically equivalent to computing the area between the recovery curve and the resting level, starting at the end of the stressor and ending at 20 min.

Curve fitting. The final method of assessing recovery modeled the process using mathematical equations, adjusting the parameters to optimize the fit for each participant on each physiological measure, for each of the six sessions. The set of parameters that produce the least squared error from the fitted equation represents the measures of recovery. To calculate these parameters, we used a Quasi-Newton minimization algorithm to minimize squared deviations between the model and the observed data (c.f. Fletcher, 1972).

The process to be modeled, recovery from a stressor response, is the transition between two fairly stable levels: the stressor level and the posttask recovery level. An equation that captures such transitions is the logistic function. Here we use a version of this function with three parameters:

$$\text{measure}(\text{time}) = \frac{a}{1 + 0.026e^{(7.33 \cdot \text{time}/b)}} + c$$

The three parameters capture psychophysiological meaningful aspects of the recovery process. The first one (parameter "a") represents the amount that the measure drops between the stressor level and the posttask recovery level. The second parameter ("b") represents the time it takes the measure to drop from the stressor level to the recovery level. The final parameter ("c") represents the

posttask recovery level to which the measure drops. The meaning of each parameter is illustrated in Figure 1, which shows logistic curves generated by this equation with the parameters varied systematically.

Although the logistic equation describes the transition between two levels (here, the stressor level and the posttask recovery level), it actually never reaches either one, but only approaches them asymptotically. Because of this asymptotic nature of the equation, there is technically no point when the drop from the stress level starts, and no point when the measure actually reaches the recovery level. Accordingly, the time of recovery parameter (“*b*”) measures the time it takes the measure to go 95% of the way between the two levels. The moment when the curve has fallen 2.5% from the stress-level asymptote is set at the end of the stressor, and recovery is considered complete when the measure is within 2.5% of the posttask level. With this particular equation, the start of recovery is when “time” = 0, and time before this is negative. The constants in the equation serve to convert the transition-time parameter (“*b*”) to seconds and fix the start of recovery at the moment that the stressor ends.

For the purposes of analysis, the data were averaged into 2-s blocks, and the curve was fit to a 12.5-min period, starting 30 s after the onset of the stressor, and ending after the first 10 min of

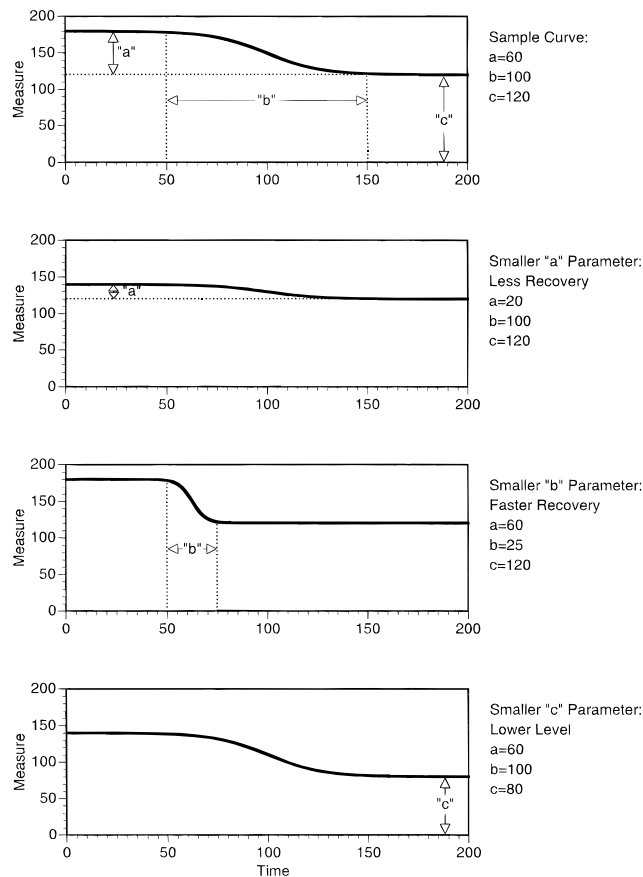


Figure 1. Effects of individually varying each of the three parameters of the logistic function. The top panel shows a typical curve. The second shows the effect of reducing only the amount of recovery parameter “*a*.” The third panel differs from the first only in that it has a faster rate of recovery parameter “*b*.” The bottom panel differs from the first only in that it has a lower level of recovery parameter “*c*.”

Table 1. Test–Retest Reliability of Resting Cardiovascular Levels

	Task		
	Physical	Mental	Speech
Systolic BP	.26	.61	.63
Diastolic BP	.33	.76	.38
Heart rate	.81	.79	.71

Note: BP = blood pressure.

recovery. Two-second blocks were used rather than 1-s blocks because, with a heart rate below 60 bpm, 1-s blocks would leave seconds for which there would be no readings. Using 2-s blocks, then, eliminates these missing values, and ensures that all curves will be fit with the same number of data points. The first 30 s of the stressor are not included, because during this time the measures are not stable, but are rising from resting to stress levels.

The parameters are estimated independently for each of the three physiological measures for each of the six sessions for each of the 18 participants. The parameter of greatest importance is the rate of recovery measure—parameter “*b*”. The amount of recovery is likely to be largely a function of the reactivity; how far participants come down will depend heavily on how far they went up. The recovery level parameter is likely to be similar to the prior measure of the participants’ baseline; where they return to be close to where they started. The recovery time parameter, however, is not dependent on the amount of recovery or the level to which the physiological measure returns.

Results

Test–Retest Reliability

Reliability of Baseline

The reliability of the resting baseline levels for the three tasks for systolic and diastolic blood pressure and heart rate are shown in Table 1. Over the nine measures, the average reliability coefficient was $r = .62$.¹

For heart rate, the reliability was high for all three tasks. The blood pressure resting levels were more mixed. This result is due at least in part to the nature of the equipment used. The Finapres has been shown to be accurate at tracking changes that occur within a single session, during which the cuff is never removed (Gerin, Rosofsky, Pieper, & Pickering, 1993), but changes in cuff placement may affect levels, producing an unusual situation in

¹This and all subsequent computations of mean correlations are based on the Fischer r to z transformation (Hays, 1973). There are several ways to think about significance levels for these analyses. All correlations have 16 degrees of freedom, and so individual reliability coefficients greater than $r = .40$ are significant $p < .05$, one-tailed. The significance of the distribution of nine test–retest coefficients can also be tested. With a .5 probability of any one coefficient being positive, finding eight or nine positive scores is significant, $p < .05$ using the binomial test. Such a test probably falsely assumes independence of the nine coefficients, but still provides a rough guide to significance. However, in general, for the assessment of reliability, statistical significance is not the critical issue. The effect size, which provides information about how much stability there is from session to session, is far more useful. Accordingly, we will report the size of the reliability correlations. Readers may compute significance levels, using one of the above techniques, if they wish.

Table 2. Reactivity Scores for the Three Tasks for Each Physiological Measure

	Task					
	Physical		Mental		Speech	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Systolic BP	39.3 (11.6)	37.1 (14.6)	19.6 (8.0)	14.5 (9.3)	34.7 (17.8)	25.8 (10.6)
Diastolic BP	19.9 (8.6)	18.6 (12.9)	10.9 (4.7)	10.2 (5.3)	19.3 (5.9)	17.3 (7.2)
Heart rate	39.7 (13.4)	35.0 (13.0)	7.6 (7.9)	4.3 (4.8)	13.8 (13.8)	8.7 (7.4)

Note: Standard deviations are in parentheses. BP = blood pressure.

which test–retest correlations for changes may be almost as high as those for levels.

Reliability of Reactivity

The reactivity means for the three cardiovascular measures for both sessions of all three tasks are shown in Table 2.

For each of the three tasks, we computed the test–retest reliability of the change from baseline to stress levels. The reliability of the reactivity is shown in Table 3. The reliability of the reactivity change scores falls within the typical range for these measurements (Gerin et al., 1993; Manuck et al., 1989) although the physical task appears slightly less reliable than the others. The lower reliability scores for the walking-in-place task may be due in part to movement artifacts resulting from the greater physical activity for this task. However, the fact that the reliability of the baseline levels before the physical task (where there is no activity) was also slightly lower suggests that task differences in the reliability of reactivity may be just random variability. The mean reliability coefficient for reactivity was $r = .54$.

Reliability of Recovery

Time to recovery. We assessed how many seconds elapsed between the time the stressor ended and the time a 30-s running average of the measure was at or below the baseline level. The reliability of this measure is shown in Table 4A.

The time-to-recovery measure was far from reliable. The average coefficient was $r = .11$. Only one of the nine correlations was large, and this was probably a fluke, given the small size of the other eight correlations. On the basis of these results, time to recovery does not appear to have sufficient reliability to be a useful

Table 3. Test–Retest Reliability of Cardiovascular Reactivity

	Task		
	Physical	Mental	Speech
Systolic BP	.28	.62	.60
Diastolic BP	.22	.68	.62
Heart rate	.64	.48	.60

Note: BP = blood pressure.

Table 4. Test–Retest Reliability of Three Measures of Recovery

Recovery measure	Task		
	Physical	Mental	Speech
A. Time to recovery			
Systolic BP	.26	.73	-.19
Diastolic BP	-.20	.01	.14
Heart rate	.31	-.16	-.08
B. Recovery at a fixed point			
Systolic BP	.15	.51	.17
Diastolic BP	.24	.31	.53
Heart rate	.26	.17	.17
C. Total carryover			
Systolic BP	.10	.11	.09
Diastolic BP	.42	.20	.39
Heart rate	.25	.30	.29

Note: BP = blood pressure.

measure of individual differences.² On average, it took 322 s for systolic blood pressure to return to baseline, 293 s for diastolic blood pressure, and 155 s for heart rate.

Recovery at a fixed point. We examined residual arousal for each measure by computing the difference between baseline and an average taken over a 1-min period starting 3 min after the stressor ends.³ Averaging across the three tasks, the residual elevation above baseline at the third minute was 6.5 mmHg for systolic blood pressure, 2.6 mm Hg for diastolic blood pressure, and 1.9 bpm for heart rate. The reliability for this measure is shown in Table 4B. Again, this measure does not produce impressive reliability, with an average of $r = .21$.

Total carryover. We measured the difference between baseline and the average over the recovery period for each measure.⁴ The reliability of this measure is shown in Table 4C. The reliability for this measure, as for the other two, was not high, with an average of $r = .24$. The mean carryover was 100 mmHg-minutes for systolic blood pressure, 52 mmHg-minutes for diastolic blood pressure, and 14 pulses for heart rate.

We conclude, on the basis of these results, that the simple, intuitively appealing methods of calculating recovery scores do not provide measures with satisfactory levels of reliability.

²There are two parameters that can be varied with this measure. One can, instead of waiting for complete recovery, measure elapsed time until the participant is some specified fraction of the way back to baseline. For example, rather than the number of seconds to return to baseline, one can measure the number of seconds until the parameter is 50% of the way back from the stress level to the resting level. The other parameter is the number of seconds in the running average. Instead of waiting until the average over a 30-s period is back to a predetermined level, one could wait until the average over a 10-s period is back to that level. Varying these parameters, however, does not produce reliability any higher than that shown in the table for 100% recovery and a 30-s running average.

³For this method of assessing recovery, there are two basic parameters that can be varied: how long after the stressor offset the measurement is taken, and how long a period goes into that measure. Altering these parameters did not appreciably improve the reliability of this measure.

⁴For this recovery measure, it is possible to calculate the total carryover for periods other than the entire 20-min posttask recording. However, reducing the carryover period does not improve the reliability of the measure.

Curve fitting. Fitting the recovery data with a logistic curve produces three parameters that represent: (a) the amount that the measures drop from stress to recovery levels; (b) the time that transition takes; and (c) the level to which the parameter drops. The equation was fit to the recovery data for all 18 participants, for each of the two sessions of the three tasks for all three cardiovascular measures, resulting in a total of 324 fits. On average, the equation accounted for 58% of the variance in readings. Figure 2 shows a sample curve fit to the data from one participant's heart-rate recovery following the physical task.

The test-retest correlations for the three curve-fitting parameters are shown in Table 5. All three parameters showed reasonable reliability. The amount of recovery—parameter “a”—showed a reliability, averaged over the three tasks and three measures, of $r = .61$. The parameter that assesses the speed of recovery—“b”—showed an average reliability of $r = .56$. The final parameter—“c”—which assesses the level to which people asymptotically return, showed an average reliability of $r = .65$. Although these reliability coefficients are not huge, they are at roughly the same level as the reliability of reactivity for these three tasks ($r = .54$). It seems clear that this curve-fitting approach to assessing recovery is capable of producing individual difference measures that are stable over time.

The three curve-fitting-derived parameters are conceptually independent—capturing the amount, speed, and level of recovery. The three parameters are also empirically largely independent. The average correlation between the “a” and “b” parameters, across the three tasks, two trials per task, and three measures, was $r = .23$ (range: $-.19$ to $.56$). The average correlation between the “a” and “c” parameters was $r = -.06$ (range: $-.53$ to $.21$). Finally, the average correlation between “b” and “c” was $r = .17$ (range: $-.21$ to $.65$).

The Relationship of Recovery and Reactivity Measures

Aside from the basic reliability of the curve-fitting approach to the measurement of recovery, these data can also address the relationship between the measures of recovery and measures of reactivity. The parameter that describes the size of the drop from stress levels to recovery levels—“a”—ought to be closely related to measures of reactivity. How far the levels go up during stress should be

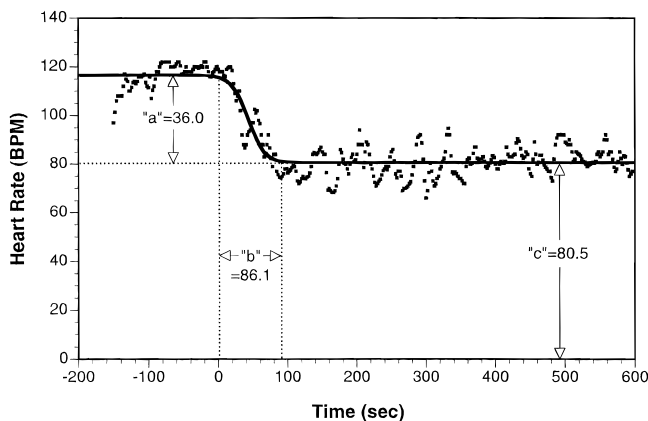


Figure 2. The logistic curve fit to one participant's heart rate data during the last 2.5 min of the physical task, and the first 10 min of the recovery period. The recovery period starts at time = 0. The three parameters indicate the amount, speed, and level of recovery.

Table 5. Test-Retest Reliability of Parameters Derived From Curve-Fitting Procedures

Recovery measure	Task		
	Physical	Mental	Speech
A. Parameter “a”—amount of recovery			
Systolic BP	.60	.63	.53
Diastolic BP	.48	.61	.62
Heart rate	.58	.62	.77
B. Parameter “b”—rate of recovery			
Systolic BP	.67	.69	.59
Diastolic BP	.38	.27	.52
Heart rate	.69	.77	.22
C. Parameter “c”—recovered level			
Systolic BP	.38	.73	.56
Diastolic BP	.15	.80	.47
Heart rate	.84	.85	.69

Note: BP = blood pressure.

similar to how far they come down afterwards. This prediction was the case. Across the three tasks, two trials per task, and three measures, the average correlation between the “a” parameter and the reactivity score was $r = .84$ (range: $.62$ to $.94$). The other two parameters were much less closely related to reactivity. The speed-of-recovery measure—parameter “b”—showed an average correlation of $r = .11$ (range: $-.13$ to $.42$). The measure that captures the level to which people return after the stressor—“c”—correlated only $r = .15$ with reactivity (range: $-.33$ to $.58$). The level-of-recovery measure was closely related to the original resting baseline, average $r = .92$ (range: $.55$ to $.98$). Given that how much the measures went up was closely related to how much they came down, it is clear that the level to which they come down should be similar to the level from which they started.⁵ Thus, whereas the amount of recovery is closely related to reactivity, and the level of recovery is similar to the resting level, the speed of recovery measure does seem to be providing information that is independent of that provided from the assessment of only reactivity.

The three intuitively appealing, though largely unreliable, measures of recovery were not independent of reactivity. The time-to-recovery measure on average correlated with reactivity $r = .48$ (range: $-.14$ to $.76$). The recovery-at-a-fixed-point measure correlated with reactivity $r = .39$ (range: $-.20$ to $.65$). The final measure, total carryover, revealed an average correlation with reactivity of $r = .32$ (range: $-.21$ to $.62$). Partialling these measures for their dependence on reactivity did not appreciably change their test-retest reliability. After removing the influence of reactivity, the test-retest correlations for time-to-recovery, recovery-at-a-fixed-point, and total carryover were $r = .14$ (range: $-.20$ to $.69$), $r = .25$ (range: $-.03$ to $.69$), and $r = .09$ (range: $-.18$ to $.41$), respectively.

⁵Although the correlations between asymptotic recovery levels and pretask levels were high, there was a slight difference at the level of means. Averaging across the six sessions, participants' systolic blood pressure returned to a level that was 4.3 mmHg above the baseline level. Diastolic blood pressure also was slightly elevated, 2.0 mmHg above baseline. Heart rate, on the other hand, showed complete recovery, returning, on average, to within 0.06 bpm of pretask levels. Consistent with the unreliability of the recovery-at-a-fixed-point measures, there was no appreciable consistency in the difference between the baseline and asymptotic recovery level (with an average test-retest correlation of $r = .13$; range: $-.43$ to $.47$). Whatever the small amount of residual arousal that is visible in blood pressure represents, it was not a stable individual difference.

Discussion

The data presented suggest that it is possible to derive reliable measures of cardiovascular recovery from stress using curve-fitting techniques. The more traditional approaches did not appear to tap anything stable about the individual. It is important to note that the test of reliability is conservative. The range of individual differences in cardiovascular functioning in the participant sample was presumably small because all were young, healthy, fit, normotensive individuals, and such similarity makes it more difficult to find stable differences between participants. The measures that appeared worthless in these data may be used profitably when the range of participants is much greater, or when the effects of significant manipulations are investigated. However, the conservatism of the present test makes it even more remarkable that the curve-fitting-derived parameters showed test-retest stability across tasks and across physiological measures. In situations in which simpler methods of assessing recovery have produced positive results, curve-fitting approaches should be even more fruitful.

It is possible that not only is speed of recovery important in understanding cardiovascular functioning, but that speed of reactivity is also useful in this regard. It may be the case that how quickly people respond to stress could be assessed using curve-fitting procedures. As we have shown with cardiovascular recovery following stress, static measures may not be useful for capturing inherently dynamic processes. Mathematical models describing the

transition between a stress and resting levels appear to capture this process reliably, and similar models of the transition from resting to stress levels may also be fruitful.

Although the measure of recovery explored here has the potential to aid in the understanding of cardiovascular responses to stress, there are some limitations to this approach. First, assessing recovery in this manner requires continuous, or at least frequent, measurement of heart rate and blood pressure. Intermittent monitoring, with measurements taken only every few minutes, is unlikely to produce data useful for estimating curve-fitting parameters. However, whereas the continuous blood pressure monitor used here (a Finapres) is of severely limited availability, there are other devices soon to be available that will make this sort of curve fitting possible for other researchers. Second, the sort of recovery that has been described in this study is fairly rapid, occurring within a few minutes of the end of the stressor. It is possible that there are important individual differences that were not captured by the laboratory tasks we used, but that have implications for recovery. For example, people may differ in their tendency to ruminate about stressful episodes, and such rumination might extend stress responses far past the actual event. It may be that other techniques will be required to measure this sort of delayed recovery. However, it is clear that short-term cardiovascular recovery from stress can be measured reliably, that such measurements can add some information that is independent of that provided by the assessment of only reactivity, and that, whereas various simple approaches cannot provide additional information, curve fitting can.

REFERENCES

- Arcangeli, G., Giovinozzo, G., Saracino, B., D'Angelo, L., Giannarelli, D., Arcangeli, G., & Micheli, A. (1998). Radiation therapy in the management of symptomatic bone metastases: The effect of total dose and histology on pain relief and response duration. *International Journal of Radiation Oncology, Biology, Physics*, *42*, 1119–1126.
- Borghii, C., Costa, F., Boschi, S., Mussi, A., & Ambrosini, E. (1986). Predictors of stable hypertension in young borderline subjects: A five-year follow-up study. *Journal of Cardiovascular Pharmacology*, *8*, S138–S141.
- Earle, T. L., Linden, W., & Weinberg, J. (1999). Differential effects of harassment on cardiovascular and salivary cortisol stress reactivity and recovery in women and men. *Journal of Psychosomatic Research*, *46*, 125–141.
- Faber, S. D., & Burns, J. W. (1996). Anger management style, degree of expressed anger, and gender influence cardiovascular recovery from interpersonal harassment. *Journal of Behavioral Medicine*, *19*, 31–53.
- Fletcher, R. (June 1972). FORTRAN subroutines for minimization by quasi-Newton methods. United Kingdom Atomic Energy Authority, Harwell, UK.
- Franz, I. W. (1986). *Ergometry in hypertensive patients*. Berlin: Springer.
- Gerin, W., & Pickering, T. G. (1995). Association between delayed recovery of blood pressure after acute mental stress and parental history of hypertension. *Journal of Hypertension*, *13*, 603–610.
- Gerin, W., Rosofsky, M., Pieper, C., & Pickering, T. G. (1993). A test of reproducibility of blood pressure and heart rate variability using a controlled ambulatory procedure. *Journal of Hypertension*, *11*, 1127–1131.
- Glynn, L. M., Christenfeld, N., & Gerin, W. (1997). Implications of alternative methods of computing blood pressure means. *Blood Pressure Monitoring*, *2*, 175–178.
- Gorback, M. S., Quill, T. J., & Lavine, M. L. (1991). The relative accuracies of two automated noninvasive arterial pressure measurement devices. *Journal of Clinical Monitoring*, *7*, 13–22.
- Hankin, J. R., Sloan, J. J., Firestone, I. J., Ager, J. W., et al. (1996). Has awareness of the alcohol warning label reached its upper limit? *Alcoholism: Clinical and Experimental Research*, *20*, 440–444.
- Haynes, S. N., Gannon, L. R., Orimoto, L., O'Brien, W. H., & Brandt, M. (1991). Psychophysiological assessment of poststress recovery. *Psychophysiological Assessment*, *3*, 1–10.
- Hays, W. L. (1973). *Statistics for the social sciences* (2nd ed.). New York: Holt, Rinehart and Winston.
- Hocking Schuler, J. L., & O'Brien, W. H. (1997). Cardiovascular recovery from stress and hypertension risk factors: A meta-analytic review. *Psychophysiology*, *34*, 649–659.
- Imholz, B. P. M., van Montfrans, G. A., Settels, J. J., van der Hoeven, G. M. A., Karemaker, M., & Wieling, W. (1988). Continuous noninvasive blood pressure monitoring: Reliability of Finapres device during the Valsalva manoeuvre. *Cardiovascular Research*, *22*, 390–397.
- Imholz, B. P. M., Settels, J. J., van der Meiracker, A. H., Wesseling, K. H., & Wieling, W. (1990). Non-invasive continuous finger blood pressure measurement during orthostatic stress compared to intra-arterial pressure. *Cardiovascular Research*, *24*, 214–221.
- Imholz, B. P. M., Wieling, W., Langewouters, G. J., & van Montfrans, G. A. (1991). Continuous finger arterial pressure: Utility in the cardiovascular laboratory. *Clinical Autonomic Research*, *1*, 43–53.
- Jamieson, J. L., & Lavoie, N. F. (1987). Type A behavior, aerobic power, and cardiovascular recovery from a psychosocial stressor. *Health Psychology*, *6*, 361–371.
- Kemper, S., Rice, K., & Chen, Y. (1995). Complexity metrics and growth curves for measuring grammatical development from five to ten. *First Language*, *15*, 151–166.
- Lai, J. Y., & Linden, W. (1992). Gender, anger expression style, and opportunity for anger release determine cardiovascular reaction to and recovery from anger provocation. *Psychosomatic Medicine*, *54*, 297–310.
- Linden, W., Earle, T. L., Gerin, W., & Christenfeld, N. (1997). Physiological stress reactivity and recovery: Conceptual siblings separated at birth? *Journal of Psychosomatic Research*, *42*, 117–35.
- Manuck, S. B., Kasprovicz, A. L., Monroe, S. M., Larking, K. T., & Kaplan, J. R. (1989). Psychophysiological reactivity as a dimension of individual differences. In N. Schneiderman, S. M. Weiss, & P. G. Kaufmann (Eds.), *Handbook of research methods in cardiovascular behavioral medicine* (pp. 365–382). New York: Plenum Press.

- McCubbin, J. A., Cheung, R., Montgomery, T. B., Bulbulian, R., & Wilson, J. F. (1992). Aerobic fitness and opioidergic inhibition of cardiovascular stress reactivity. *Psychophysiology*, *29*, 687–697.
- McDowell, J. E., Clementz, B. A., & Wixted, J. T. (1996). Timing and amplitude of saccades during predictive saccadic tracking in schizophrenia. *Psychophysiology*, *33*, 93–101.
- Parati, G., Casadei, R., Groppelli, A., Di Rienzo, M., & Mancia, G. (1989). Comparison of finger and intra-arterial blood pressure monitoring at rest and during laboratory testing. *Hypertension*, *13*, 647–655.
- Suhnel, J. (1998). Parallel dose–response curves in combination experiments. *Bulletin of Mathematical Biology*, *60*, 197–213.
- Van Egeren, L. F., & Sparrow, A. W. (1989). Laboratory stress testing to assess real life cardiovascular reactivity. *Psychosomatic Medicine*, *5*, 1–9.
- Wesseling, K. H., Settels, J. J., van der Hoeven, G. M. A., Nijboer, J. A., Butijn, M. W. T., & Dorlas, J. C. (1985). Effects of peripheral vasoconstriction on the measurement of blood pressure in a finger. *Cardiovascular Research*, *19*, 139–145.
- Wieling, W., ten Harkel, A. D. J., & van Lieshout, J. J. (1991). Spectrum of orthostatic disorders: Classification based on an analysis of the short-term circulatory response upon standing. *Clinical Science*, *81*, 241–248.
- Wilson, P. W. (1997). An epidemiological perspective on systemic hypertension, ischemic heart disease, and heart failure. *American Journal of Cardiology*, *80*, 3J–8J.

(RECEIVED December 14, 1998; ACCEPTED August 31, 1999)