

MULTILEVEL ANALYSIS: PHYSIOLOGICAL AND BIOCHEMICAL MEASURES

Gary G. Berntson and John T. Cacioppo

The National Institute on Aging commissioned the National Academies of Science to organize scientific discussion that culminated in a workshop volume whose title queried, “*Cells and surveys: Should biological measures be included in social science research?*” (Committee on Population, 2001). The short answer to that question was yes.

Although psychologists have long appreciated the value of converging operations using multi-method approaches, the NAS report found that psychologists are increasingly engaged in research entailing multilevel analyses that extend well beyond the traditional disciplinary boundaries. Multilevel analyses represent a subset of multi-method approaches in which the measures, constructs, and theories extend across levels of organization—from the psychological to the physiological to the cellular and ultimately to the gene and beyond. Efforts to integrate information across levels of analyses are especially challenging, but this is precisely what is necessary for the ultimate interdisciplinary convergence on mind–body issues.

Multilevel analyses can be problematic, as the terms, constructs, and measures are often diverse, and the concepts and theories at different levels of analyses may develop largely independently of those of another level. This fosters what has been termed the category error, wherein seemingly parallel concepts from different levels of analysis may reflect only partially overlapping domains, rather than representing a one-to-one isomorphism. The

ultimate goal of multilevel analysis is to mutually calibrate concepts, relate measures, and integrate information across levels, so as to inform processes and constrain theories at multiple levels of analysis.

More important, this process entails *reductionism*, but not in the pernicious sense of *substitutionism*. Although it may be conceivable to explicate a motivational state in terms of the interactions of atomic elements, there are several important limitations to this approach. The first is the matter of efficiency and scale. The atomic underpinnings of motivational states are so extraordinarily complex that the language and constructs pertaining to atoms may not be the most efficient or feasible way to conceptualize motivation.

A second problem is the likelihood of a category error. Even if we could identify a set of atomic events that correspond to the motivational state, this does not imply an isomorphism. Motivation is a construct that has developed to account for variations in behavior of organisms; in the absence of behavior there would be no need for such a concept. Not only would there not be an agent to conjure up such a notion, but there would be no applicability at the atomic level. Although motivation certainly has causal relations to processes at the atomic level of analysis, there is not an identity across these vast levels, and it is patently silly to apply motivational constructs to atoms. Motivation applies to functional properties of more complex living organisms.

One might argue that motivational phenomena may be explicable ultimately in terms of the properties of atomic particles, and that the problem is simply one of the intricacy of mapping across such distal levels. This is a specious perspective, however. The third and most important limitation to substitutionism is that it *begs the question*¹ if the properties imputed to lower level elements to account for higher level phenomena are knowable only by observations from the higher level of organization. This is a logical fallacy (begging the question or circular reasoning) because the “explanatory” properties are derived from the phenomena to be explained. These properties cannot be said to be proper to the elements, but only derivable from a higher level of analysis that studies the elements in relation to others. Some properties of atoms may be knowable by the study of individual atoms, but others (e.g., atomic behavior in crystals) may become known only in interactions with other atoms. Similarly, although atomic or subatomic events ultimately underlie all our thoughts, feelings, and actions, the latter phenomena could not be said to be proper characteristics of the atomic elements. If they were, then all principles and properties would be assigned to quantum particles, which would be patently senseless because these properties and principles would *not* be of the particles, but of their configurations into aggregates, which may be meaningfully explained by constructs at different levels of organization.

Multilevel analysis is not about *substitutionism*, but about the ability of information derived from distinct levels of analysis to mutually inform others. *Reductionism* refers to the ability of events at lower levels of analysis to inform or explicate events at higher levels of analysis. Multilevel analysis is a two-way process, however, as higher level analyses can also elucidate or inform lower level processes (*extensionism*). Important in this effort is the development and refinement of meaningful *theories* of the relations between levels. Also central to this reductionism–extensionism process is the mutual tuning and calibration of concepts to enhance cross-

level mappings and minimize category errors. This is especially important because of the intricacies and multiple mappings across distinct levels and the associated need for model constraints. This chapter highlights some features of multilevel analysis, provides a reductionism–extensionism framework for conceptualizing and implementing such analyses, and offers illustrative examples. A major theme is the mutual benefit that multilevel analyses offers for both the higher (e.g., psychological) and lower (e.g., physiological) levels of analysis.

PRINCIPLES OF MULTILEVEL ANALYSIS

Some principles pertaining to multilevel analysis have been articulated by Cacioppo and Berntson (1992; see also Cacioppo, Berntson, et al., 2000), which serve to frame issues and organize research perspectives. They are enumerated following.

The principle of *multiple determinism* stipulates that a target event at one level of organization, especially at more molar levels, will have multiple antecedents within and across levels of analysis. Parenting, for example, has both social and genetic determinants (Meaney, 2001). Because of the multiple antecedents across even proximal levels, the mappings across more divergent levels of analysis become increasingly complex. This is captured by an important corollary to the principle of multiple determinism. Although the ultimate goal of multilevel analysis is to bridge distal levels, the *corollary of proximity* suggests that this effort may be more straightforward for more proximal levels. As bridges are built among adjacent levels, those integrations will facilitate the superordinate mappings across progressively more disparate levels. This is not to say that bridging across broader levels of analysis is not possible or desirable. There are examples of programmatic research efforts that span multiple levels, such as the collaborative effort of Michael Meaney to map from the gene to maternal behavior and back again (Meaney, 2001). This was accomplished, however, through a systematic series of interdisciplinary collaborative efforts, which individually cut across a more limited span of levels.

¹Originally *petitio principii* from Aristotle (350 B.C.) Posterior Analytics, translated by G. R. G. Mure, MIT Internet Classics Archive: <http://classics.mit.edu/Aristotle/posterior.mb.txt>.

The principle of *reciprocal determinism* asserts that there may be mutual, reciprocal influences among levels of organization—that is, the direction of causation is not one way. To continue with our example of gene–maternal interaction, there is a clear genetic bias in the pattern of maternal behavior in rats, but the pattern of maternal behavior has also been shown to impact specific gene expression in the offspring (Meaney, 2001). Moreover, this experience-dependent influence on gene regulation can extend beyond the subsequent generation, through nongenomic inheritance (Meaney, 2001). The principle of reciprocal determinism also has a guiding research corollary. Because causal influences among levels can be bi-directional, the *corollary of interdependence* states that a single level of analysis cannot yield a comprehensive account of multilevel phenomena, and that no single, preferred level of analysis applies uniformly. This is not to say that researchers should not do single-level research, as important phenomena for multilevel analyses derive from research and theory within a single level of analysis. Moreover, the selection of the most optimal level of analysis for single-level research depends on the experimental question and the theoretical interest (e.g., genetic vs. maternal determinants). The corollary indicates, however, that a comprehensive understanding of multilevel phenomena will require multilevel analysis.

Finally, the principle of *nonadditive determinism* reflects the fact that the properties of the whole cannot always be predicted by knowledge of properties of the parts. The sources of variance from higher level processes are often broader than those for lower levels of organization, so higher level systems tend to be more complex. Following the preceding example, the mere knowledge of a genotype may be uninformative as to phenotype, which in critical ways depends on multiple interactions with the social/maternal context (Meaney, 2001). Consequently, understanding genetics would not be complete if the study were restricted to the cellular domain. This principle reflects the increase in relational complexity with higher levels of organization and introduces the final corollary. The *corollary of asymmetry* states that the definition of a phenomena of interest should include observations at the highest

level of organization at which it manifests, as it may not be understood by appeal exclusively to lower levels of analysis. That is, higher level analyses can identify and characterize phenomena that may be explicated in part by lower level organizations, but these phenomena may never be known from analyses limited to the lower level processes. This corollary would not preclude strictly lower level (e.g., molecular) analyses, but would apply at the point those molecular analyses were invoked to account for higher level phenomena (e.g., behavior).

The principles and corollaries just outlined are conceptual guidelines rather than prescriptions. Moreover, we wish to emphasize that merely mapping concepts from one level to another, although informative, does not in itself constitute an explanation of those relations. The latter will require well-developed theories that can foster predictions, allow experimental control, and permit hypothesis testing and theoretical refinements.

APPLICATIONS TO MULTILEVEL ANALYSES

Psychophysiological measures offer a unique vantage for multilevel analysis as they index physiological processes and events that may intervene between psychological processes and health or behavioral outcomes. Because they represent the operations of integrated physiological systems rather than isolated molecular events, these measures are more proximal to psychological processes than are molecular events. This is in keeping with the *corollary of proximity*, and the intermediate level of psychophysiological processes may provide important bridges between psychological and more molecular levels of organization.

Heart Rate Measures of Psychological States and Processes: Multiple Determinism

There is now an extensive history of theory and research on the potential links between psychological states, autonomic regulation, and disease processes. A common measure in this literature has been heart rate. The electrical signature of the heart beat is readily recorded as the electrocardiogram (EKG) by noninvasive surface electrodes, and heart rate has been known for centuries to be sensitive to

psychological states. It is theorized, for example, that decreases in heart rate are triggered by an external direction of attention, a decrease in arousal, passive coping, or an orienting response; whereas increases in heart rate have been said to reflect inwardly directed attention, an increase in arousal, effort, active coping, or a startle or defensive response (Graham, 1984; Lacey & Lacey, 1980; Obrist, 1981). A potential advantage of heart rate is the fact that it may reflect implicit psychological states in the absence of verbal or other behavioral actions and thus may provide a metric of psychological processes that may otherwise not be apparent.

The principle of multiple determinism, however, cautions against an overly simplistic interpretation of heart rate and heart rate change. Not only is there a wide range of psychological states or processes that influence heart rate, physical (e.g., temperature, posture) and physiological (e.g., activity, blood pressure) variables also impact heart rate. Hence, the utility of heart rate as an index of psychological processes is dependent on the rigor of the experimental design and the interpretive logic to be applied. This is underscored by the high error rates (both hits and misses) in the misapplication of physiological measures to the detection of deception (see Committee report, 2003; Lykken, 1998).

Part of the difficulty in this area relates to the multiple mappings across levels of organization and analysis. Although a fear stimulus may alter heart rate, there are many translations in this cascade: from the stimulus to percept, from percept to emotion, and from emotion to autonomic outflows. There is one further translation involved as the heart is not an autonomic organ, per se, but is merely regulated by the autonomic nervous system. As each translation likely entails multiple mappings from one stage of processing to the next, the overall intricacy in psychophysiological relations can be staggering. The *corollary of proximity* emphasizes the advantages of bridging across more proximal levels. A major goal of multilevel research is to progressively elucidate the mapping between disparate levels by building a series of local bridges among more adjacent levels.

The measurement model: heart versus autonomic outflow. The heart is dually innervated by the sympathetic and parasympathetic divisions of the autonomic nervous system, with the sympathetic system exerting a positive chronotropic effect (increasing heart rate) and the parasympathetic system exerting a negative chronotropic influence (decreasing heart rate). Changes in heart rate represent at best an indirect reflection of autonomic control. One legacy from the Walter Cannon era is that the autonomic branches are subject to reciprocal central control, with increases in activity of one branch associated with decreases in the activity of the other (see Berntson & Cacioppo, 2000; Berntson, Cacioppo, & Quigley, 1991). Within this conceptual framework, heart rate should reflect the state of sympathetic–parasympathetic balance, and this appears to hold for many autonomic reflexes that are organized at lower levels of the brain stem. Higher neurobehavioral substrates, however, can inhibit, modulate, or bypass lower reflex substrates and thereby exert broader and more flexible control over the autonomic branches (Berntson & Cacioppo, 2000; Berntson et al., 1991).

In behavioral contexts, one can see not only the classical reciprocal mode of control, but also independent changes of the autonomic branches, or even the concurrent coactivation or coinhibition of both branches. This clearly necessitates an expansion of the theoretical model, and hence the measurement model, from the classical bipolar continuum from sympathetic to parasympathetic dominance, to a bivariate autonomic space that more appropriately characterizes the multiple modes of control. As illustrated in Figure 12.1, the bivariate model subsumes the bipolar model for a reciprocal mode of control, but also expands this model to capture independent or coactive changes that cannot be represented in the bipolar model. This in turn raises serious questions about the utility of heart rate measures as an index of autonomic outflow, as increases in heart rate, for example, could result from an independent increase in sympathetic control, an independent decrease in parasympathetic control, a sympathetically domi-

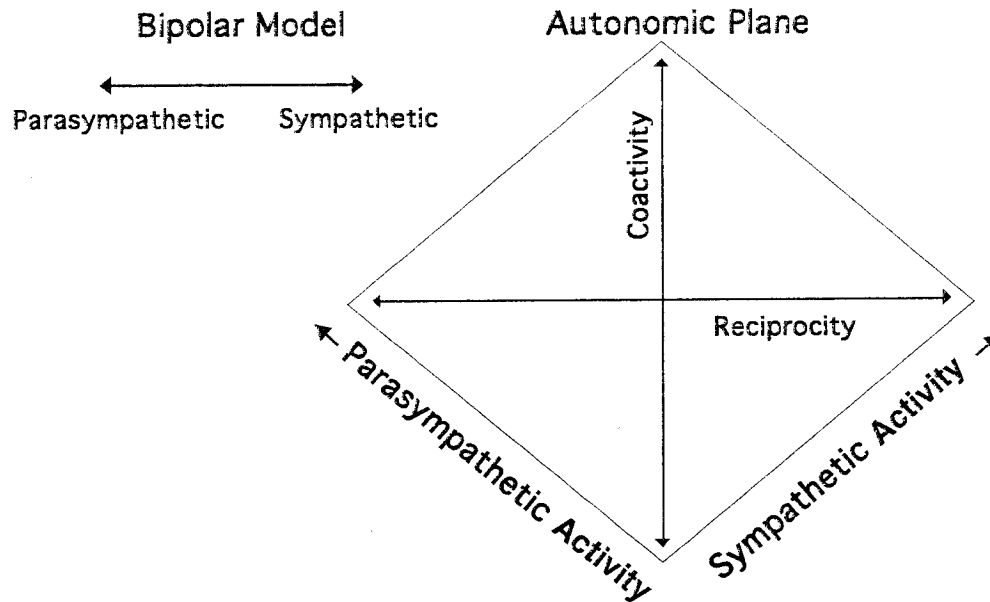


FIGURE 12.1. Conceptual models of autonomic control. Left: Bipolar model of reciprocal sympathetic/parasympathetic control. Right: Bivariate model of sympathetic and parasympathetic control that allows independent and coactive as well as reciprocal modes of autonomic response.

nated coactivation, or a parasympathetically dominated coinhibition. This ambiguity is illustrated by the isofunctional contour lines in the three-dimensional map of Figure 12.2, which illustrates the chronotropic state of the heart as a function of location within the autonomic plane.² These contour lines illustrate loci within the autonomic plane (i.e., different combinations of sympathetic and parasympathetic activities) that translate into equivalent chronotropic states. Consequently, the chronotropic state of the heart does not map simply on patterns of autonomic outflow, as a given chronotropic state is ambiguous with regard to its autonomic origins. Because neurobehavioral substrates control autonomic outflows, not the heart directly, measures of the chronotropic state of the heart necessarily entail a loss of fidelity in psychophysiological mappings.

Metrics of autonomic space. Differences in the modes of cardiac control for physiological reflexes

and psychological contexts are illustrated by a human study of autonomic responses to an orthostatic stressor (assumption of an upright posture) and to psychological stressors (mental arithmetic, speech stressor, and speeded reaction time task). Before considering those results, however, a measurement issue must be addressed. The change in measurement model from a bipolar to a bivariate representation has obvious implications for experimental dependent measures. If heart rate or heart period are not adequate, how does one measure autonomic outflows? That is, what constitutes a valid measure of sympathetic and parasympathetic activities? In anesthetized animal studies, direct recordings have been made of neural firing in sympathetic and parasympathetic cardiac nerves. This is not feasible in human subjects, however, and has limited applicability even in animals as the requirement for anesthesia precludes meaningful psychophysiological investigations. Microneurography (using a fine microelectrode) has been applied to

²From here on, the chronotropic state of the heart will be designated in the metric of heart period, or the reciprocal of heart rate. The former has advantages as heart period is more linearly related to neural activity within the autonomic branches.

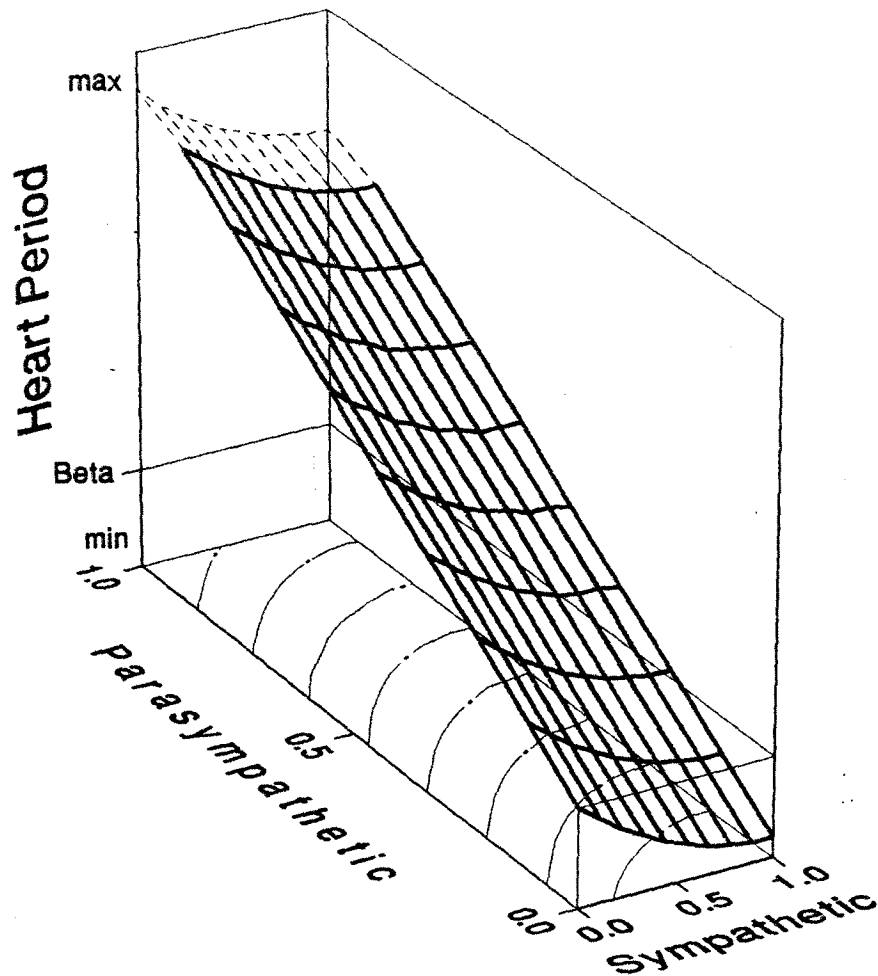


FIGURE 12.2. Three-dimensional autonomic space representation of chronotropic control of the heart. The effector surface depicts the heart period level for all possible loci within the autonomic plane. Parasympathetic and sympathetic axes are scaled in proportion to the extent of their functional range of control, and the curvature in the surface reflects nonlinearities in these controls. Beta (on the abscissa) illustrates the heart period in the absence of autonomic control. The curved lines on the autonomic plane are isofunctional contour lines, which represent varying combinations of sympathetic and parasympathetic control that yield comparable heart period effects. Reprinted from *Behavioral Brain Research*, 94, Berntson, Sarter, and Cacioppo "Anxiety and Cardiovascular Reactivity: The Basal Forebrain Cholinergic Link," 225–248. Copyright (1998), Elsevier.

measure autonomic neural activity in conscious humans, but this technique is only applicable for rather superficial autonomic nerves (e.g., Macefield, Elam, & Wallin, 2002).

Another approach to measuring the separate contributions of the autonomic branches to cardiac control entails pharmacological blockade of the branches. Blockade of the parasympathetic branch, for example, will prevent the action of that branch

and reveal the isolated contribution of the sympathetic branch, and vice versa. This has been problematic, however, as blocking one branch may indirectly alter the other (e.g., by reflex adjustments). Moreover, although drugs may be highly specific to a receptor type and can thus differentiate sympathetic and parasympathetic effector synapses, they are not specific as to the target organ and may exert actions at some remote site, including the

brain. Such remote actions could alter the psychological states of interest or otherwise bias reactivity. The complications with pharmacological blockades have been sufficiently serious as to question their validity and limit their application. A new measurement methodology was clearly needed.

A more extensive pharmacological protocol and a more comprehensive analytical approach provided that methodology (Bertson, Cacioppo, & Quigley, 1994). Consider the observed heart period response (\emptyset) to some evocative stimulus occurring at the vertical line in Figure 12.3. As depicted, blockade of the parasympathetic branch would reveal the isolated sympathetic response $\emptyset Pblk$, which provides an estimate of the sympathetic contribution (termed the residual estimate or s'). At the same time, the response decrement from the unblocked condition ($\emptyset - \emptyset Pblk$) offers an estimate of the normal contribution of the parasympathetic branch (termed the subtractive estimate, or p'). Conversely, blockade of the sympathetic branch ($\emptyset Sblk$) provides a residual index of the isolated parasympathetic response (p') and the response decrement from the unblocked

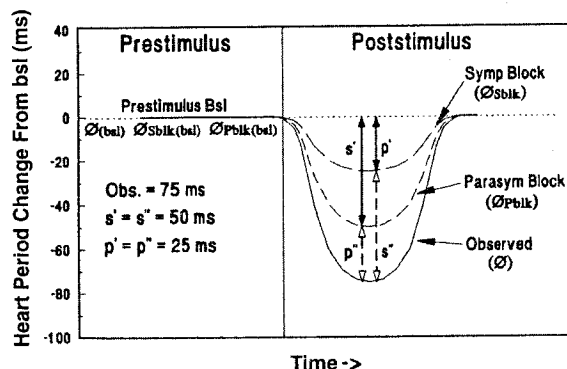


Figure 12.3. Illustration of heart period response in pharmacological blockade analyses. Solid line illustrates the observed response in the absence of blockade (under saline control conditions). Dashed lines illustrate the response under selective sympathetic and parasympathetic blockades. Arrows illustrate the residual (s' and p') and subtractive (s'' and p'') estimates of sympathetic and parasympathetic control. From "Autonomic Cardiac Control. I. Estimation and Validation from Pharmacological Blockades," by G. G. Bertson, J. T. Cacioppo, and K. S. Quigley, 1994, *Psychophysiology*, 31, 572-585. Copyright 1994 by Blackwell Publishing, Ltd. Reprinted with permission.

condition ($\emptyset - \emptyset Sblk$) offers an estimate of the normal contribution of the sympathetic branch (s'').

The preceding analyses provide two estimates of the functional contributions of each autonomic branch, and an overall estimate can be derived as the means:

Estimate of sympathetic response (at time t) =

$$\Delta S_t = (\Delta s_t' + \Delta s_t'') / 2$$

Estimate of parasympathetic response (at time t) =

$$\Delta P_t = (\Delta p_t' + \Delta p_t'') / 2$$

More important, because the residual and subtractive estimates are derived from distinct pharmacological blockers (muscarinic cholinergic antagonists for the parasympathetic branch and β_1 adrenergic antagonists for the sympathetic branch), their side effects and remote actions would be different. If the estimates agree, despite these differences, one would have increased confidence in the estimates of autonomic control. Moreover, any discrepancy in the independent estimates could be indexed by an error term (ϵblk), which is the difference between the two estimates at a given point in time. This value can be formally shown to be equivalent for the two branches. Thus

$$\Delta \epsilon blk_t = (\Delta s_t' - \Delta p_t'') = (\Delta p_t' + \Delta p_t'')$$

As the discrepancy between the two estimates becomes larger, ϵblk_t increases, and one would have lower confidence in the estimate. This is formalized in a validity coefficient:

$$v_\delta = (|\text{effect size}| / |\text{effect size}| + \epsilon blk)$$

The validity coefficient can range from 0 when the error is very large relative to the estimated response, to 1.0 when the error term is negligible. An example of this analysis is shown in Figure 12.4 for orienting responses of rats to auditory stimuli. The top panel illustrates the observed responses under the control condition and after sympathetic (atenolol) and parasympathetic (scopolamine) blockade. The lower panels illustrate the overall as well as the residual and subtractive estimates of the contributions of the branches to the observed

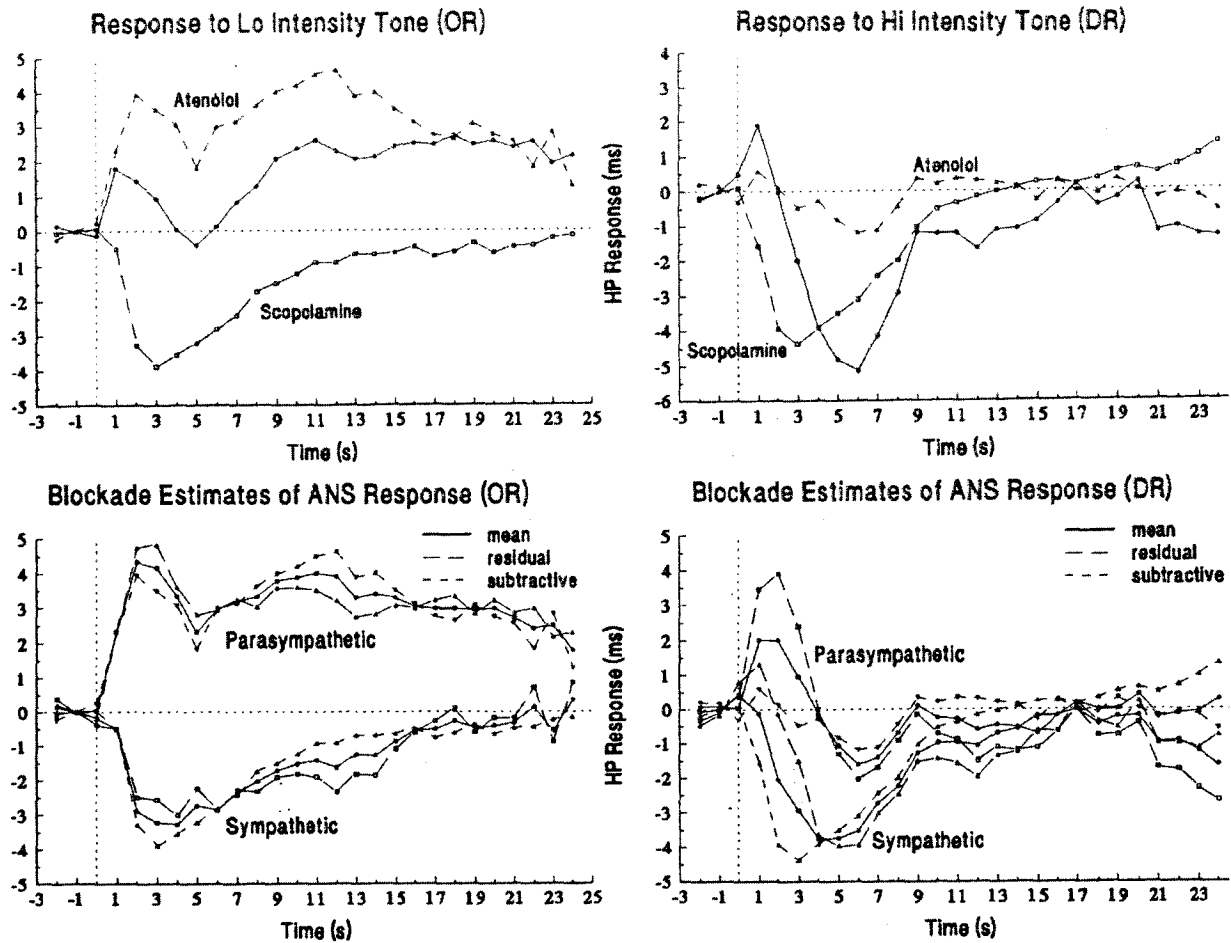


FIGURE 12.4. Pharmacological analysis of sympathetic and parasympathetic responses of orienting (OR) and defensive (DR) responses in the rat. Upper panels illustrate responses to a discrete auditory stimulus of low or high intensity under the saline control condition, and after sympathetic (atenolol) and parasympathetic (scopolamine) blockades. Bottom panels show residual, subtractive, and overall estimates of sympathetic and parasympathetic response. The stimuli occurred at the time of the vertical dotted line in each panel, and responses are expressed as a change from prestimulus baseline. From "Autonomic Cardiac Control. I. Estimation and Validation from Pharmacological Blockades," by G. G. Berntson, J. T. Cacioppo, and K. S. Quigley, 1994, *Psychophysiology*, 31, 572-585. Copyright 1994 by Blackwell Publishing, Ltd. Reprinted with permission.

response. As can be seen, there was relatively good agreement between the residual and subtractive estimates, yielding a small error term and a high validity coefficient. The response to the orienting stimulus revealed autonomic coactivation, as the increased heart period due to parasympathetic control indicates parasympathetic activation, and the decrease in heart period under sympathetic control similarly revealed sympathetic activation. Because activation of the two branches tends to oppose one another, the observed response in the unblocked

condition was smaller than under either blockade condition.

With the refined measurement method outlined, we now return to the human study of physical and psychological stress. Nine human subjects were tested for the autonomic response to the orthostatic stressor and the psychological stressors after intravenous infusions of saline, Metoprolol (a sympathetic β_1 blocker), and atropine (a parasympathetic blocker). Estimates were derived as outlined, and response vectors were derived on the autonomic plane, based on

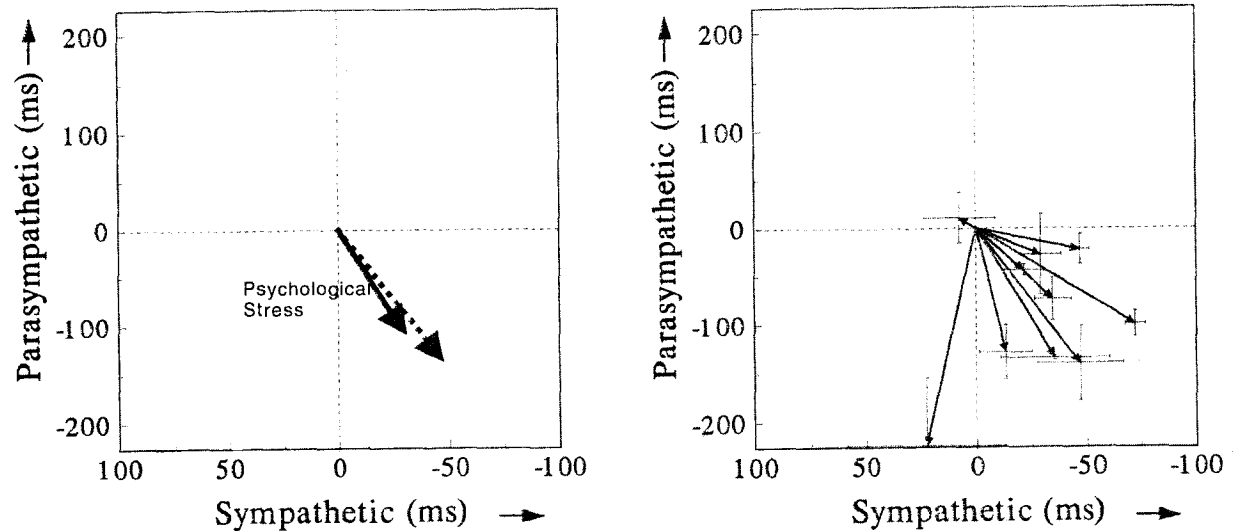


FIGURE 12.5. Orthostatic versus psychological stress. Left: Group mean responses to orthostatic and psychological stressors depicted as response vectors on the autonomic plane, from prestress baseline (intersection of horizontal and vertical dotted lines) to the stress conditions (arrowheads). Axes depict *ms* of heart period change related to sympathetic and parasympathetic control. Right: Individual response vectors ($N = 9$) to the psychological stressors revealing individual differences in the direction of response. Individual differences were stable, as evidenced by standard deviation bars at the arrowheads, reflecting deviations across the three psychological stressors (mental arithmetic, speech stress, and reaction time). Note that responses for a given individual were generally consistent across stressors. From "Autonomic Cardiac Control. III. Psychological Stress and Cardiac Response in Autonomic Space as Revealed by Pharmacological Blockades," by G. G. Berntson, J. T. Cacioppo, and K. S. Quigley, 1994, *Psychophysiology*, 31, 599–608. Copyright 1994 by Blackwell Publishing, Ltd. Reprinted with permission.

the change score along the sympathetic and parasympathetic axes. Results are illustrated in Figure 12.5, which displays response vectors from baseline (intersecting dotted lines). Both classes of stressors yielded an overall reciprocal pattern of sympathetic activation and parasympathetic withdrawal.

This similarity at the group level, however, belies a fundamental difference between the two classes of stressors. In accord with the reciprocal pattern of control in many autonomic reflexes, there was a significant negative correlation between the responses of the autonomic branches across subjects with the orthostatic stressor. Greater increases in sympathetic control were associated with larger decreases in parasympathetic control. All subjects showed similar response vectors, differing only in magnitude. In contrast, there was no correlation between the autonomic branches to the psychological stressors. Rather, as illustrated in Figure 12.5 (right), there were notable individual dif-

ferences in responses to psychological stress. Some subjects showed primarily parasympathetic withdrawal, others reciprocal sympathetic activation and parasympathetic withdrawal, and still others primarily sympathetic activation. This was not attributable simply to an increase in error variance to psychological stress, as individual response vectors were stable across the psychological stressors. This can be seen in the error bars at the arrowheads of Figure 12.5 (right), which depict the standard errors for the response vectors under the different psychological stressors.

Why does it matter? Without independent measures of sympathetic and parasympathetic control, lawful differences between orthostatic and psychological stressors would not have been apparent, and individual differences in the response to psychological stress would not have been discerned. In accord with the corollary of proximity, psychophysiological

mapping in this case was improved by the deployment of a more appropriate analytical method that assessed autonomic control at a more proximal level than could be derived from the end organ response. This is important not only for basic studies of psychophysiological relations, but also because different modes of autonomic control may have distinct health implications.

There have been reports of a relation between cardiac reactivity to stressors and negative health status, including diminished immune functions, although the predictive power of heart rate is small and not always significant (see Cacioppo, 1994). This is likely attributable to the use of heart rate measures, as Cacioppo (1994) found no relation between overall heart rate reactivity and the immune response to vaccine, but did find a significant relation between immune status and the sympathetically mediated component of heart rate reactivity. Multilevel analysis, capitalizing on more proximal mappings, revealed order in psychosomatic relations where none was apparent with more distal mappings.

Glucocorticoids and Behavioral States: Reciprocal Determinism

Glucocorticoids (cortisol in humans, corticosterone in rats) are steroid hormones of the adrenal cortex that have potent effects on glucose metabolism and immune function, as well as on psychological processes (Gore & Roberts, 2003; Lovallo & Thomas, 2000; Schimmer & Parker, 1996). Glucocorticoids are classic stress hormones and have been commonly used as biochemical markers of stress reactions (McEwen, 2000). As illustrated in Figure 12.6, the secretion of glucocorticoids is regulated by the anterior pituitary hormone adrenocorticotropic hormone (ACTH), which in turn is controlled by the hypothalamic peptide corticotropin releasing hormone (CRH). CRH is released in a pulsatile fashion (see Veldhuis et al., 2001), regulated by pituitary, hypothalamic, and hippocampal circuits that bear glucocorticoid receptors and are sensitive to glucocorticoid levels. These circuits exert a feedback inhibitory influence on CRH release. The hypothalamic and pituitary negative feedback mechanisms represent the traditionally

recognized routes responsible for short-term regulation of glucocorticoid secretion, whereas the hippocampus appears to be involved in stress reactions and longer term glucocorticoid regulation.

In addition to the short-term pulsatile patterns of release, the hypothalamic-pituitary-adrenal-cortical axis (HPAC) displays a circadian rhythm, with plasma glucocorticoid levels peaking in the early morning hours and showing a nadir in the late afternoon and minor peaks around mealtimes (see Lovallo & Thomas, 2000). Glucocorticoids bind to both glucocorticoid (GR) and mineralocorticoid (MR) receptors, and the steroid/receptor complex is translocated to the nucleus, where it can serve as a transcription factor to regulate gene expression (Gore & Roberts, 2003). More rapid actions may be exerted by glucocorticoid binding to membrane bound receptors (see Lupien & McEwen, 1997).

Measurement issues: reliability and validity. The pulsatile nature of ACTH and cortisol release poses a problem of reliability, as the level of hormone in plasma will vary depending on the time relation of the sampling to the pulsatile pattern of release. One approach to improving reliability has been to take multiple samples (e.g., Veldhuis et al., 2001) and then aggregating over samples if the interest is in tonic levels or preserve the temporal samples if the interest is in time-varying patterns of secretion. An additional measurement complication is the notable circadian rhythm in cortisol release. The measurement of the diurnal rhythm, by repeated cortisol measurements across the day, has been used to assess the status of the HPAC. If a more limited sample of cortisol is desired (e.g., as a stress marker), the diurnal rhythm not only imposes the restriction that samples be taken at the same time of the day, but also raises a question concerning the optimal time for sampling.

Measures derived late in the day or at night generally are not optimal for studies of chronic stress, because of the low levels of secretion and sensitivity limits. Consequently, for measures of chronic stress, samples are commonly taken during peak levels in the morning. The change in cortisol over 30 minutes (or so) from waking, for example, has been suggested to be a sensitive measure of adreno-

HPAC System and Hormonal Secretion

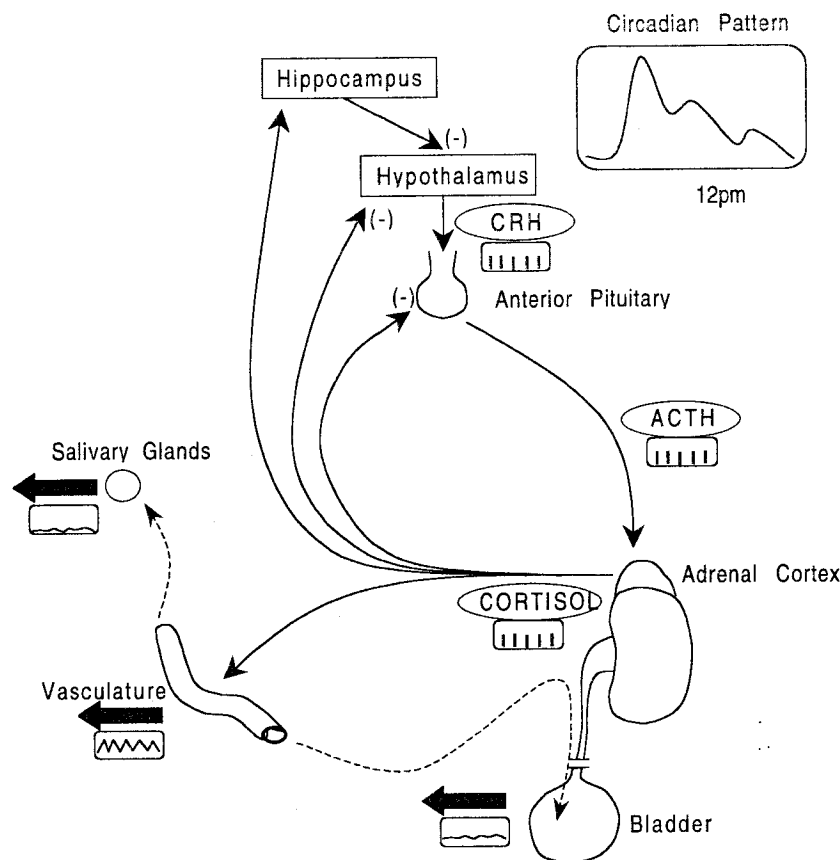


Figure 12.6. Structures and secretions of the glucocorticoid system. Hormones are listed in oval text boxes, and the rectangular inserts illustrate time-varying patterns of secretion or local concentrations, from pulsatile to more steady state, and over the circadian cycle. Solid arrows illustrate sample methods, salivary, vascular, and urinary. As illustrated, plasma cortisol levels are the highest and most variable and include both bound and unbound hormones, whereas salivary and urinary are more time stable and are considerably lower, as they represent unbound hormones.

cortical reactivity (Schmidt-Reinwald et al., 1999). Conversely, phasic reactivity to stress may be more appropriately assessed in the afternoon, when basal levels are lower and more stable. Some conditions such as chronic stress or depression may be associated with blunted negative feedback regulation and hence a diminished circadian pattern. This pattern may be more readily identified by evening measures or by indices of circadian fluctuations. Circadian fluctuations are often thought to arise in large part from changes in feedback regulation, but other factors could also impact these rhythms, including

altered or disrupted sleep/waking cycles, and should be considered when interpreting differences in circadian fluctuations (Spath-Schwalbe et al., 1993).

A more direct test of feedback control, having high construct validity, is the dexamethasone-suppression test. This procedure entails the administration of a standardized dose of the synthetic steroid dexamethasone, along with pre- and postadministration measurements of ACTH or cortisol. Secretion of these endogenous hormones will be suppressed in proportion to the potency of

steroid feedback inhibition. A subset of depressed patients (50%–60%) show elevated cortisol levels, an attenuated circadian rhythm, and a blunted response to dexamethasone (see Parker, Schatzberg, & Lyons, 2003). This may reflect conditions within brain feedback circuits or changes in glucocorticoid receptor sensitivity. We will return to these possibilities later, in the context of stress effects.

Additional measurement issues arise over the fact that typically less than 10% of plasma cortisol is in a free, unbound, biologically active state. The rest is reversibly bound to plasma proteins (corticosteroid-binding globulin, CBG), which decrease bioavailability and metabolic clearance (Breuner & Orchinik, 2002). Adding further complexity is the fact that the proportion of bound cortisol may vary with cortisol levels or other physiological conditions. Moreover, CBG binding may enhance bioavailability under some conditions, as it represents a releasable cortisol reservoir (Breuner & Orchinik, 2002). Because plasma cortisol reflects both the free and bound fractions, this measure may not provide the most valid index of cortisol tissue bioactivity under all conditions, despite the fact that plasma levels are often considered the gold standard of cortisol measures.

As illustrated by the solid arrows in Figure 12.6, plasma cortisol represents only one metric of HPAC activity. With regard to the issue of bound vs. unbound cortisol, salivary cortisol levels offer the advantage of indexing only the unbound fraction. This is because CBG and other proteins do not readily diffuse across cellular membranes. Consequently, protein-bound cortisol does not readily enter the salivary glands, and salivary cortisol levels reflect primarily the unbound fraction of plasma cortisol. Salivary cortisol levels are also noninvasive and can be obtained under a wider range of experimental conditions, including ambulatory studies. Although the time constant of cortisol diffusion into salivary glands tends to dampen pulsatile patterns somewhat, salivary cortisol can still show short-term pulse-related fluctuations. Time required to acquire an assayable saliva sample also tends to blunt, but does not eliminate, short-term fluctuations. Consequently, the time sampling issues raised

earlier for plasma cortisol levels also apply to salivary cortisol measures.

A measure of cortisol can be derived also from urine. Urinary cortisol reflects the free, unbound fraction of plasma cortisol, as protein-bound cortisol does not readily enter the renal tubular system. Cortisol accumulates in the urine in proportion to plasma free-cortisol levels, and because of the general stability of this molecule, collection of urinary output can provide an integral index of cortisol over extended (including daily) periods.

There is considerable debate as to what constitutes the best measure of HPAC activity, and there may be no single answer to this question. Rather, the validity of a measure may be defined by the problem under study. Urinary measures integrated over a day or more may be most relevant for studies of chronic stress. In contrast, shorter term measures such as those from plasma or saliva may be more useful for studies of acute stress or circadian rhythms.

Reciprocal influences between neurobehavioral and HPAC systems. In accord with the principle of Reciprocal Determinism, the HPAC system offers an illustration of the multiple interactions between neuroendocrine systems, neurobehavioral substrates, and psychological processes. Although physical stressors are known activators of the HPAC system (Selye, 1956), psychological stressors are among the most potent (Mason, 1968, 1975; McEwen, 2000). It is also clear that HPAC activity can impact both cognitive and emotional processes (e.g., Lupien & McEwen, 1997; Parker et al., 2003). Psychological states can alter HPAC activity, and HPAC activity can modulate the psychological states that gave rise to this activity. These reciprocal actions can be dose and context dependent. Glucocorticoid administration can either enhance or impair cognitive processes, as a function of dose, context, and the specific receptor populations activated (Lupien & McEwen, 1997).

Some of the complexity of these effects relate to the multiple reciprocal interactions within the HPAC system (e.g., CRH and cortisol feedback) and between the HPAC and psychological processes (e.g., stress and cortisol). Reciprocally interacting systems are difficult to study and characterize in isolation, as

their functional outputs represent a close interplay across levels of organization. Consequently, manipulations at one point may have diverse effects throughout these circuits. The central CRH system, in addition to its regulation of pituitary ACTH release, is considered to be a general orchestrator of the cognitive, affective, behavioral, autonomic, and neuroendocrine aspects of stress. Local intracerebroventricular infusions of CRH in primates results in an activation of stress-related brain circuits, induces anxiety- and depressive-like reactions, and decreases social interactions (Strome et al., 2002). Because glucocorticoid administration alters central CRH activity, it is not immediately apparent whether the effects of this manipulation reveal the direct actions of glucocorticoids or indirect effects on CRH systems. Dissecting reciprocally interacting systems requires multiple experimental approaches and converging data that can provide a more comprehensive perspective than a more restricted analyses. Because interactions may never be known by studying hormonal systems in isolation, the *corollary of interdependence* asserts that the most meaningful studies will entail manipulations and observations of both CRH and cortisol, involving a combination of methods.

Social psychological influences and the corollary of interdependence. Relations across levels are particularly difficult to conceptualize when the reciprocally interacting nodes extend across broad spans of organization or analysis, as the complexity of mappings tends to increase across more distal levels. A recent line of research in psychoneuroimmunology illustrates this. It has long been recognized that psychological stressors are potent activators of the HPAC system (Mason, 1968, 1975; McEwen, 2000). In contrast to the general adaptation model of Selye (1956), it further appears that there may be fundamental differences in kind among physical and social-psychological stressors. Social reorganization stress in mice (rotation of alpha males among housing colonies) can lead to reactivation of herpes simplex Type 1 virus (HSV1), similar to the stress-related HSV1 reactivation that causes cold sores in humans (Padgett et al., 1998). In contrast, physical stressors (e.g., restraint-stress or shock) are ineffective despite producing comparable glucocor-

ticoid levels. Subsequent work has revealed further unique characteristics of social stressors, highlighting the need for multilevel research and mandating expansion and refinements in the concept of stress and the nature of stressors.

Subsequent studies revealed that social stressors in mice are associated with an exaggerated and often lethal inflammatory response to influenza virus, compared to restraint stress (Sheridan, Stark, Avitsur, & Padgett, 2000). The difference between the social and the physical stressors could not be accounted for by differences in secretion of anti-inflammatory glucocorticoids, because both classes of stressors again yielded comparable glucocorticoid levels. Rather, it appears that social stress induced a state of glucocorticoid resistance or receptor insensitivity attributable to an impairment in nuclear translocation of the glucocorticoid/receptor complex in specific macrophages of socially stressed animals (Quan et al., 2003). As a result, glucocorticoids failed to suppress the actions of a transcription factor (NF-kappaB), which promotes the production of pro-inflammatory cytokines (interleukin 1 and tumor-necrotizing factor alpha). In this research, a bridge was established between social processes and gene expression in the health effects of stress.

These examples illustrate the principle of Reciprocal Determinism and its Corollary of Interdependence. Multilevel studies can elucidate influences across levels of organization and clarify relations that could not be known by studies limited to a single level of analysis.

Loneliness and Health: Nonadditive Determinism

The utility of multilevel analysis to understanding psychological processes and psychosomatic relations is illustrated by our recent work on loneliness. Social isolation and loneliness are potent but little understood risk factors for broad-based morbidity and mortality (Seeman, 2000). Although loneliness has a heritable component, differences in social cognition provide a better explanation for the physiological characteristics of lonely versus nonlonely individuals than does a model based on invariant traits or genetic determinism. Lonely individuals tend to construe their world, including the behavior

of others, as punitive or potentially punitive. Consequently, lonely individuals are more likely to be socially anxious and to adopt a prevention focus rather than a promotion focus in their social interactions (Ernst & Cacioppo, 1999). Lonely individuals are more likely to appraise stressors as threats rather than challenges and to cope in a passive, isolative fashion rather than an active fashion that includes seeking the help and support of others. These differences in social cognition predictably result in an increased likelihood of lonely individuals acting in self-protective and, paradoxically, self-defeating ways (Cacioppo, Berntson, et al., 2000).

From the dual clues that isolation and loneliness are associated with broad-based mortality and with a higher death rate across the adult life span, one can surmise that the underlying mechanism operates on a wide range of bodily systems, or that there is more than one mechanism through which loneliness influences health. In the absence of lower level analyses, however, the relations between loneliness and health remain mere empirical associations. Recent evidence suggests that different transduction pathways account for acute effects of loneliness on morbidity and mortality (e.g., suicide) and chronic effects (e.g., heart diseases, cancers). Two of the neurobehavioral mechanisms that contribute to the association between loneliness and chronic disease are (a) catabolic processes—lonely individuals perceive more hassles and stressors in daily life and are characterized by higher tonic levels of peripheral resistance in the cardiovascular system, which over time may have damaging effects on body organs and systems; and (b) anabolic processes—lonely individuals show physiological repair and maintenance processes (e.g., wound healing, sleep) that are less efficient than nonlonely individuals (Cacioppo, Hawkley, & Berntson, 2003).

A variety of autonomic differences have been found to distinguish lonely and nonlonely individuals. The bivariate model of autonomic control outlined previously represents a significant advance in our understanding of higher neural influences on autonomic substrates, considerably clarifies psychophysiological relations, and increases the fidelity of mappings from psychological processes to autonomic cardiac control.

Characterization of response modes. As discussed earlier, pharmacological blockade analyses represent a gold standard for the quantification of patterns of sympathetic and parasympathetic control of end organs such as the heart. Blockade analyses are not always possible, however, so noninvasive measures are desirable. Noninvasive measures of parasympathetic and sympathetic control have now been validated, at least for the heart.

Respiratory sinus arrhythmia (RSA) is a fluctuation in heart rate in phase with respiration, inspiration being associated with an increase in heart rate and expiration with a decrease. RSA arises from pulmonary and thoracic stretch receptor afferents to brain stem reflex substrates that trigger inhibition of vagal outflow and excitation of sympathetic outflow (Berntson, Cacioppo, & Quigley, 1993). Both of these changes synergistically act to increase heart rate, but there are differences in the time constants of the sympathetic and parasympathetic synapses at the cardiac sinoatrial node pacemaker (see Berntson et al., 1997). The consequence of these temporal dynamics is that respiratory rhythms in the parasympathetic cardiac innervation is translated into rhythmical fluctuations in heart rate, whereas the sympathetic synapses are sufficiently slow that respiratory fluctuations are filtered out. Respiratory rhythms in heart rate thus reflect vagal cardiac control, with larger fluctuations associated with higher vagal tone (Berntson et al., 1993). RSA has been repeatedly validated as an index of vagal control of the heart, with some caveats (see Berntson et al., 1997; Berntson, Cacioppo, Binkley, et al., 1994; Cacioppo et al., 1994).

An additional noninvasive measure, pre-ejection period (PEP) is available to index sympathetic control of the heart. Pre-ejection period is the time between the electrical invasion of the ventricular myocardium (Q wave of the EKG) to the opening of the aortic valve and the onset of ventricular ejection. The pre-ejection period is a standard marker of ventricular myocardial contractility, as more forceful myocardial contractions result in the more rapid rise of intraventricular pressure and hence earlier ventricular ejection. A decrease in PEP thus indicates an increase in contractility. The sympathetic innervation enhances myocardial contractil-

ity, whereas the parasympathetic system plays only a minor role. Consequently, variations in sympathetic control yield corresponding changes in contractility and inverse changes in PEP. With appropriate controls and caveats, PEP has been validated as an index of sympathetic cardiac control (Berntson, Cacioppo, Binkley, et al., 1994b; Cacioppo et al., 1994).

Through a combination of measurements and calculations, additional parameters of cardiodynamic and hemodynamic processes can be derived noninvasively. Heart rate (HR) and stroke volume (SV) each has sympathetic and parasympathetic contributions, and together these parameters determine cardiac output (CO) or the amount of blood expelled by the heart into the vascular system each minute (i.e., $CO = HR \cdot SV$). Blood pressure, which must be maintained within relatively narrow ranges to maintain adequate circulation, is a function of the cardiac output and total peripheral resistance (TPR; the resistance to blood flow through the circulatory system). Systolic and diastolic blood pressure (SBP and DBP, respectively) can be measured noninvasively, and mean arterial pressure (MAP) can be calculated from these measures (e.g., $MAP = .33 \cdot SBP + .67 \cdot DBP$). TPR, therefore, can be calculated from MAP and CO (i.e., $TPR = BP/CO$).

Psychophysiological patterns in loneliness. In the pursuit of psychophysiological mappings across levels of organization and analysis, the corollary of proximity specifies that the complexity of mapping generally increases across more disparate levels. Consequently, simple isomorphic relations between events or processes are less likely to hold across more disparate levels of organization. Rather, relations across levels may need to consider patterns of multivector mappings to achieve more isomorphic mappings. Work on loneliness highlights this issue.

The cardiovascular autonomic features of lonely individuals do not organize simply on a single sympathetic-parasympathetic dimension. Lonely and nonlonely *young* adults have comparable blood pressure, but the underlying physiology differs between these groups: lonely individuals have been found to be characterized by higher total peripheral resistance and lower cardiac output than nonlonely individuals.

This difference is equally apparent at rest (baseline) as when performing orthostatic or psychological stressors (Cacioppo, Hawkey, Crawford, et al., 2002), and ambulatory recordings further revealed that this difference is evident not only in the laboratory but also during a typical day in their lives (Hawkey, Burlison, Berntson, & Cacioppo, 2003). These physiological differences and their links to health would go unrecognized with measures of blood pressure alone, which highlights the importance of theoretical systems that aid in the selection of appropriate measures to effectively bridge across levels.

Chronic elevations in total peripheral resistance not only mean that the heart muscle must work harder to distribute the same amount of blood through the circulatory system, but the reduced diameter of the blood vessels may also increase turbulence in and potential damage to the vasculature. Both central (e.g., baroreceptor reflex) and peripheral (e.g., vascular elasticity) mechanisms may degrade over time, further diminishing the ability to maintain normotensive pressure even during rest. Consistently elevated levels of vascular resistance, coupled with age-related decreases in vascular compliance, may set the stage for the development of hypertension. A study of older adults in a south Chicago apartment development confirmed this hypothesis. Because the sample size was relatively small, participants were categorized into low or high lonely groups by a median split on their scores on the UCLA loneliness scale. Results indicated that age was positively and significantly correlated with systolic blood pressure among lonely individuals, whereas there were no age-related increases in systolic blood pressure among nonlonely individuals.

Corollary of asymmetry. The patterns of neuroendocrine and autonomic control in lonely individuals are not intuitively obvious, but may have substantial basic and health significance. Further studies will be necessary to elucidate the neurophysiological and neurobehavioral origins of these patterns, and their health implications. Both of these efforts will require multilevel analyses, through which information at multiple levels can provide converging perspectives and insights. There is no single level of analysis that would permit meaningful pur-

suit of these questions. Moreover, there is no single level that can be universally assumed to be preeminent in multilevel analyses. On the other hand, the levels of organization and analysis cannot be viewed as "equivalent," and conceptualizations of the relationships among levels are not simply a matter of preference. Rather, there is a fundamental asymmetry across levels in multilevel research.

The most optimal approach may be to conceptually guide research by the data and constructs deriving from the level that confers the greatest organization on the problem. For the question of how social stress impacts glucocorticoid resistance, for example, the most useful organizing level of organization and analysis may revolve around neuroimmune systems. In contrast, for the question as to physiological features of loneliness, the more salient level of analysis may be the psychosocial. Indeed, the organization in the literature is apparent only by parsing populations on the dimension of loneliness. In the absence of that, the lawful variance associated with loneliness would be relegated to the error term, and there would be no way of identifying this source of variance based on physiological studies alone.

This discussion is not intended to foster largely meaningless debates as to the "ultimate" level of analysis, nor is it intended to deny reductionism or support substitutionism. Rather, the organizing level of analysis is that which serves most effectively to structure knowledge and guide research and theory. In the case of loneliness, there is a natural asymmetry, with preeminence of the social psychological level that defines the primary conceptual dimension. This does not imply that the research and findings of this level of analysis are any more important than those of other levels, as all are required. Moreover, the optimal level of analysis may change over time, as constructs at the social-psychological level come to be implemented or integrated at lower levels, whereby the focus of research may shift to a lower level.

The corollary of asymmetry asserts that the optimal level of analysis for guiding research and theory

may be that at which the major conceptual dimensions are implemented or realized. Studies of the relations between physical stress and disease may not necessarily need to appeal to the social-psychological level of organization, at least not initially. Social-psychological processes are likely important modulators of such relations, however, as social relations have been shown to be important moderator variables in stress-immune relations. This is an illustration of where a shift in focus toward higher levels of organization and analysis may be as informative as a reductionistic shift toward lower levels.

SUMMARY

Cross-disciplinary, multilevel research is an increasingly salient feature of contemporary science in general and of psychology in particular. It is a trend that will undoubtedly continue. Realization of the full potential of the explosive developments within neurosciences, genetics, and molecular biology will require the integration of this information within the broader knowledge base concerning higher level behavioral, cognitive, and social psychological domains. It may appear to be a rather daunting task to integrate, for example, cellular biology with social psychology, but it is a task that must be accomplished. The principles of multiple determinism, reciprocal determinism, and nonadditive determinism, together with their corollaries, offer some strategic guidelines to organize such efforts. The ultimate goal of this enterprise is not to obliterate social sciences in a puff of *substitutionism*, but rather to promote meaningful *reductionism* and *extensionism* so that knowledge and constructs at multiple levels of organization and analysis can mutually inform, elucidate, and constrain theory and research at other levels. This goal may never be finalized, but it is already apparent that keen insights and important scientific developments can be derived from multilevel research approaches and interdisciplinary theoretical systems that can integrate information across levels of organization and analysis.