The Influence of the Work Environment on Cardiovascular Health: A Historical, Conceptual, and Methodological Perspective

Stanislav V. Kasl
Yale University School of Medicine

The framework of psychosocial epidemiology is used to examine research developments that characterize the accumulation of knowledge regarding the role of the work environment in cardiovascular health and disease. The discussion of current programs of research focuses on the work of T. Theorell and R. Karasek (1996) and J. Siegrist (1996) as exemplars of European and American studies that have contributed the most to the understanding of occupational cardiovascular health. It is argued that researchers need to maintain and nurture relatively broad conceptual models of etiology because cardiovascular disease involves multiple biomedical risk factors and because specific aspects of the work environment are embedded in a large, complex matrix of other psychosocial influences. At the same time, investigators need to push ahead with focused research strategies to clarify the precise nature of the work environmental risk factors that emerge in the broad, somewhat imprecise epidemiologic study designs.

The two articles describing programs of research linking aspects of the psychosocial work environment to coronary heart disease (Siegrist, 1996; Theorell & Karasek, 1996) summarized major developments at the intersection of occupational medicine and health psychology. This work represents exemplary accomplishments in a difficult and complex area, and signals a new maturing of research in terms of conceptual and methodological richness and sophistication. At the same time, these articles do not attempt to offer closure or to suggest that investigators are nearing the end of the story regarding the influence of work on cardiovascular health. The cliché that "this research raises more questions than it answers" is both unfair and inaccurate, in that it does provide a solid theoretical and empirical framework within which to plan with confidence future studies. But many such future studies are obviously still needed and many choices regarding direction, emphasis, specific formulation of issues, and methodology will confront investigators in this area. Hence, these articles represented a choice opportunity to look back to where the field has been and to look forward to the various directions in which it may go in order to most effectively and comprehensively advance researchers' understanding.

I am persuaded by the writings of other epidemiologists (e.g., Kaplan, 1994; Marmot, 1993; Syme, 1988) who have argued for the need to maintain, at least initially, a broad public health and social–epidemiological perspective on the work role, work environment, and on cardiovascular disease. For example, the potential usefulness of classes of variables, such as socioeconomic indicators or stable characteristics reflecting individual differences in personality traits or in skills and abilities, should not be ignored a priori just because a particular theoretical formulation does not encompass them. Obviously, investigators have to work with manageable conceptual formulations and a manageable set of variables to be assessed and analyzed, and thus have to omit, in specific studies, potentially useful classes of variables. But in this field of research, the success (or even dominance) of one particular approach or formulation seldom means that other approaches or formulations have been discredited and can be discarded because disparate formulations are seldom competing with each other so directly that success of one means the failure of the others. Rather, diverse formulations are likely to be complementary in the sense that a second formulation may complement the first by (a) providing a larger context and potential antecedent variables, (b) identifying mediating processes, and (c) offering variables with additional independent explanatory power.

This article is organized into several sections. In the first two sections, I discuss the general social epidemiological literature on coronary heart disease. The objective is to sketch out a broad framework within which the more specific research on the psychosocial work environment and cardiovascular health can be placed. In the next two sections, I
comment on the articles by Theorell and Karasek (1996) and Siegrist (1996) and try to evaluate the contribution of this program of research to the overall picture of psychosocial influences on coronary heart disease. In the last section, I discuss some of the future directions for research, with a particular focus on targeted research designs supplementing broad epidemiological studies.

It is worth noting that the articles by Theorell and Karasek (1996) and Siegrist (1996) encompass a great extent European studies with European work settings. Two recent reviews of the job strain model (Kristensen, 1995; Schnall, Landsbergis, & Baker, 1994) illustrated this European dominance, particularly for studies with clinical cardiovascular outcomes; studies of blood pressure and other cardiovascular risk factors, however, are more evenly divided between European and U.S. investigators. This European dominance is not an anomaly in these two articles because there is a strong tradition, particularly in the Scandinavian countries, of investigating the health effects of the work environment. The equally strong tradition in Great Britain tends to examine the work environment within the broader context of social class (e.g., Marmot, Rose, Shipley, & Hamilton, 1978; Marmot & Theorell, 1988; Pocock, Shaper, Cook, Phillips, & Walker, 1987). Europe is also the setting for some exciting and innovative research on general psychosocial factors in coronary heart disease, such as the Kuopio Ischemic Heart Disease Risk Factor study (Julkunen, Salonen, Kaplan, Chesney, & Salonen, 1994; Salonen, 1988).

The fact that a lot of the accumulated evidence regarding clinical outcomes is based on European studies raises a question about the extent to which the findings are fully applicable to the U.S. setting. There are two lines of evidence suggesting caution in extrapolating European findings to the U.S. setting:

1. Orth-Gomér (1979) carried out a case control study of men in New York and in Stockholm with and without ischemic heart disease. The Swedish men ascribed stress mainly to the job situation, whereas American men reported stress caused by family conflict. However, in overall quantity of stress, the excess reported by cases compared with controls was comparable in the two countries. A second study (Siegrist, Dittman, Rittner, & Weber, 1982) that compared German and American men who had heart disease with healthy controls obtained similar differences. Because these are retrospective case-control studies, the cross-national differences could reflect different content of attributions (lay concepts of etiology) rather than true differences in etiological dynamics. There is European (Marmot, 1982) and U.S. (Shekelle & Lin, 1978) evidence from studies of public beliefs about causes of heart attacks that reveals that stress (or worry, nervous tension, and pressure) is the most frequently given perceived cause.

2. Many reports from Great Britain (Morris, Cook, & Shaper, 1994), Sweden (Stefansson, 1991), Finland (Martikainen, 1990), Denmark (Iversen, Anderson, Andersen, Christoffersen, & Keiding, 1987), and Italy (Costa & Segnan, 1987) have shown that unemployment is associated with excess mortality (total as well as cardiovascular). The excess can be represented by standardized mortality ratios (SMRs) of about 150 to 200, depending on adjustments for available confounders. The most comparable U.S. study (Sortie & Rogot, 1990), which matched U.S. Census Bureau Current Population Surveys to the National Death Index, failed to obtain, for either men or women, SMRs for the unemployed that were significantly different from 100. This discrepancy with the European data is even more puzzling because it is widely held that the social net protecting the unemployed is stronger in these European countries than in the United States. In making this point, I am ignoring the studies based on business cycle analyses, involving aggregate time-series data. Although the convergence of European and U.S. results may appear greater in such studies, the methodology itself is currently seen as problematic and controversial (e.g., Catalano, 1991), and I therefore view the evidence as inconclusive. For a recent review, see Kasl, Rodriguez, and Lasch (in press).

I am also persuaded by the evidence (see Kasl, 1993) suggesting that for large segments of U.S. adults, work may not be a very meaningful human activity and that the work role may not be as important to them as to European adults. For example, one U.S. survey (Quinn & Shepard, 1974) found that in response to the question "How much do you think you can tell about a person just from knowing what he or she does for a living?", some 48% chose "nothing" or "a little" as their answers. Another study (Shepard, 1971) found that blue-collar workers in mechanized production ranked the lowest on "self-evaluative involvement," that is, the degree to which work (compared with nonwork) activity was most important to self-evaluation. However, I do realize that I am not able to cite comparable European studies that might show that the work role is more important to the European workers. Furthermore, I am also assuming that the costs of alienation and of lowering of expectations about work, coping strategies, and reactions—presumptively more characteris-
tic of U.S. workers—do not include elevated cardiovascular risk because the worker has disengaged from work pressure. Because the presumed difference in alienation between U.S. and European workers would be more likely to be observed among blue-collar workers, the following comment from Schnall et al. (1994) is relevant: "Swedish blue-collar men . . . exhibited substantially stronger associations between job strain and CVD than higher SES groups. In the Framingham Heart Study, such differences were minor for men" (p. 392–393).

Collectively, the above lines of evidence suggest the possibility that U.S. workers, or some subset of them, may not be affected by their work environment in the same way or to the same extent as are European workers because the former may be less involved in their work and the work role.

I emphasize that because Siegrist (1996) and Theorell and Karasek (1996) restricted their articles to cardiovascular outcomes, my commentary does so as well. However, this in no way implies that the formulations proposed by Theorell and Karasek and by Siegrist have little relevance for other outcomes, including psychological functioning and mental health. Clearly, the job strain model has been examined in relation to many other outcomes (Karasek & Theorell, 1990; Kristensen, 1995), but they are not considered here. The reader is cautioned not to assume that the impact of work environment on cardiovascular variables is necessarily predictive of a similar impact on psychological indicators, such as symptoms of distress or indicators of dysphoric mood.

The Psychosocial Epidemiology of Coronary Heart Disease: Etiologic Models

Psychosocial epidemiology refers to the study of the role of psychological variables in the etiology of disease. A major issue is how to understand the possible ways in which psychosocial variables can impact the health-to-disease transitions. A useful approach is to formulate, first, a disease development schema in terms of clinical and biomedical parameters, and then to graft onto this the different possible mechanisms by which psychosocial variables can influence the different steps or transitions in disease development. For coronary heart disease, the steps in such a developmental schema can be formulated as follows (Kasl, 1984): (a) asymptomatic status, risk factor or factors absent; (b) asymptomatic status, risk factor or factors present; (c) subclinical disease susceptible to detection; (d) initial symptom experience (if any); (e) first clinical event (diagnostic criteria are met); (f) course of disease (e.g., repeat episodes, residual disability), either as natural course or in the context of treatment; and (g) mortality (case fatality).

The role of psychosocial variables may be then stated generically as the influence on the transition from some particular earlier stage to some particular later stage (or, less often, as a return to an earlier stage, such as from b to a). Given this schematic approach, a number of observations can be made. For example, it is readily apparent that researchers can pinpoint the role of psychosocial variables much better if they study transitions between adjacent steps in the schema (e.g., the transition from being a nonsmoker to being a smoker), than between more distal steps (e.g., from asymptomatic status to mortality), or when the transition studied is unclear and may be variable (e.g., coronary heart disease mortality in a cohort where researchers do not know who had a previous history of a heart attack and when the event took place).

Another observation, growing out of the use of the disease development schema, is that the typical prospective epidemiologic studies that follow an initially healthy cohort do not provide information on all stage-to-stage transitions. For example, (a) at baseline some participants are already with elevated risk factors and their developmental history cannot be reconstructed (with the exception of smoking and, possibly, diabetes), or (b) information is lacking regarding extent of subclinical disease, whether at baseline or during follow-up. (The latter point may need some updating as advances in ultrasonographic techniques enable us to study the progression of extracoronary atherosclerosis in nonclinical populations; Julkenen et al., 1994.) Thus even in well-designed prospective studies, the precise role of psychosocial variables may not be determinable, particularly concerning the distinction between an influence on rate of subclinical disease progression versus an influence on time until the first clinical event for any given stage of subclinical disease.

A reading of the articles by Theorell and Karasek (1996) and by Siegrist (1996) does not reveal any explicit statements regarding the stage of disease development that is thought to be influenced by the psychosocial work variables central to the two programs of research. However, a reading of the various published studies mentioned in the articles reveals the following:

1. In studies in which the outcome is overt clinical disease, the research reports adjust for standard (or
available) risk factors before examining the role of job strain or imbalance in effort–reward. This suggests a model in which the psychosocial work variables are believed to be risk factors independent of, and not operating through, the established biomedical risk factors.

2. In studies in which the outcome is one of the biomedical risk factors, often blood pressure or lipid fractions, the model being tested is that job strain or imbalance in effort–reward affect cardiovascular disease risk factors, which then presumably mediate the development of clinical disease.

This is not a serious inconsistency, in that a clearly explicated etiological model could in fact postulate that the work variables impact on disease development at both stages. However, it does call for a careful summarization and evaluation of the evidence. Specifically, when the work environment variables fail to predict coronary heart disease, given the analysis has adjusted for the standard risk factors, this need not be a total lack of support for the etiological model if at least the baseline cross-sectional data show an association between the work variables and the standard risk factors. Thus, it is prudent and informative to build the analysis in several steps so that one can specifically note the change in the predictive power of the work variables before and after statistical adjustments for traditional risk factors.

Even more complex models could be postulated and tested. For example, Marmot (1993) noted that the high rates of smoking in Japan do not seem to translate into high rates of coronary heart disease, and that the strength of smoking as a risk factor may be related to background levels of risk represented by plasma lipids, which are quite low in Japanese men and women. By analogy, it might be suggested that job strain and imbalance in effort–reward, as risk factors, could interact differently with individual biomedical risk factors or combinations of such risk factors. Similarly, the evidence linking psychosocial work variables to fibrinogen levels (e.g., Marmot, 1986) and the recognition that platelet physiology may represent a different psychophysiological mechanism (Markowitz & Matthews, 1991), suggest that analyses should treat fibrinogen separately from the other biomedical risk factors and indicators.

The Psychosocial Epidemiology of Coronary Heart Disease: An Overview of Past Findings

The intent of this section is to provide a rather broad empirical background to the study of psychosocial influences on coronary heart disease, including the early studies of work stress. The purpose is to set the stage for a later commentary that argues that there are promising leads in this body of evidence, which can be exploited as long as researchers maintain a rather broad conceptual model of psychosocial work environment influences on coronary heart disease.

Two early, exhaustive reviews by C. D. Jenkins (1971, 1976) set the stage for later reviews to reorganize and update the evidence (e.g., Booth-Kewley & Friedman, 1987; Fletcher, 1991; C. D. Jenkins, 1982; Matthews, 1988; McQueen & Siegrist, 1982; Ostfeld & Eaker, 1985; Scheier & Bridges, 1995). The broad categories of psychosocial influences on coronary heart disease that have been examined include (a) socioeconomic factors and sociocultural variables; (b) environmental stressors, including chronic stressors and life change events; (c) aspects of the social environment reflecting social networks and social supports; and (d) stable personal characteristics, including personality traits, Type A behavior, and individual differences in physiologic reactivity (Ostfeld & Eaker, 1985).

What follow are highly selective, if not slightly idiosyncratic, comments regarding additional variables that might be relevant to the job strain and imbalance in effort–reward models of Theorell and Karasek (1996) and Siegrist (1996). It should also be kept in mind that the psychosocial etiology is quite different if one considers incidence of angina pectoris versus incidence of myocardial infarction (C. D. Jenkins, 1971, 1976); sudden death as the first manifestation of coronary heart disease presents a more mixed or confusing picture. The comments below regarding psychosocial etiology implicitly assume myocardial infarction as the primary or sole indicator of coronary heart disease.

**Socioeconomic Disadvantage**

The evidence is rather compelling that those who occupy the more disadvantaged status in an industrial society are at greater risk for coronary heart disease. Adjustments for standard risk factors generally attenuate but do not fully explain this excess risk. Social disadvantage is richly connected to other potential risk factors (Kaplan, 1994, 1995) but further adjustments for these do not eliminate the impact of social disadvantage. The role of some candidate variables, such as financial difficulties, has been minimally examined. The effect of specifically occupational status on coronary heart disease is clearly mediated by differences in aspects of the psychosocial work environment (Marmot, 1993; Marmot & Theorell, 1988), but the differential
implications of broad social status versus specific occupational status for coronary heart disease etiology are difficult to pin down. For example, the Whitehall study of British civil servants has yielded a much stronger gradient in coronary heart disease (Marmot et al., 1978) than has the British Regional Heart study, which has a somewhat broader range of status from a much greater variety of occupational settings (Pocock et al., 1987). This suggests that some unmeasured variables linked to low status in the latter study may be actually protective of coronary heart disease.

**Social Mobility and Sociocultural or Status Incongruity**

In the earlier C. D. Jenkins (1971, 1976) reviews, social mobility and status incongruity appeared to be a promising set of variables, linking higher mobility (e.g., geographical or intergenerational, but not occupational) and greater status incongruity (e.g., on status indicators between spouses) to greater risk of coronary heart disease. In a later review (C. D. Jenkins, 1982) these variables are no longer mentioned, partly because of nonreplication of findings and partly because of paucity of new studies investigating these dimensions. However, it would seem wise not to completely abandon these variables. For example, they may have important linkages to other variables that have emerged since then, such as social support. More important, some forms of status incongruity, especially between occupation status and educational attainment, may be quite important for the job strain model and are explicitly part of the effort–reward imbalance model (Siegrist & Peter, 1993). For example, in the well-known Bell Telephone System study (Hinkle et al., 1968) men without a college education had an increasing excess of coronary heart disease, compared with the college graduates, as the level of job status in the company increased. To the extent that educational level may indicate adequacy of training for the jobs they occupied, it may be argued that those without college education but in high-status jobs experienced a higher level of job demands or a lower level of skill, two of the components of the job strain model.

**Sustained Disturbing Emotions**

The phrasing “sustained disturbing emotions” by C. D. Jenkins (1982) has a certain amount of intentional vagueness in order to encompass both trait and state characteristics. That is, sometimes the measurement allows a reasonably confident interpretation that a trait is involved (e.g., neuroticism), whereas at other times (e.g., anxiety) it may not be clear if this is a trait or a more-or-less enduring state of dysphoria in response to a chronic exposure. In any case, neuroticism and anxiety appear primarily related to incidence of angina pectoris but not to first myocardial infarction; depression may also play a role in case fatality after a myocardial infarction but probably not its incidence. Two dimensions appear promising here:

**Sleep disturbance.** The label sleep disturbance by C. D. Jenkins (1982) actually includes a broader array of items predicting myocardial infarction, including being tired on awakening, being exhausted at the end of the day, and not being able to relax. The statistical evidence regarding the predictive power of these overlapping items is rather solid but the interpretation is not. The possibilities are as follows: (a) this association reflects a stable trait that is an independent risk factor for myocardial infarction; (b) this association represents a response to chronic environmental demands, including those from work; (c) this association reflects the underlying role of sleep apnea, which affects the restfulness of sleep and is also a risk factor for myocardial infarction (e.g., Ketterer, Brymer, Rhoads, Kraft, & Kenyon, 1994); (d) this association reflects the underlying role of “vital exhaustion” (van Diest & Appels, 1994), which is a strong predictor of myocardial infarction, primarily during the first year of follow-up (Appels, 1993). The linkage of vital exhaustion to the work environment, on the one hand, and to the concept of prodromal period, on the other hand, remains to be clarified. These interpretations, except the one in which sleep apnea is an unmeasured confounder, suggest that the inclusion of such items in studies of the health effects of the work environment, particularly of psychological job demands, would be quite useful.

**Hostility, cynicism, anger, irritability, suspicion.** There is a considerable convergence of evidence (e.g., Dembroski, MacDougall, Costa, & Grantid, 1989; Goldstein & Niaura, 1992; Julkunen et al., 1994; Matthews, 1988; Williams & Barefoot, 1988) that some aspect of the interrelated domain of hostility, cynicism, anger, irritability, and suspicion variables represents increased risk for coronary heart disease. It is not clear, however, whether researchers are dealing with a stable personality trait or enduring (repeated) reactions to environmental stimuli. Nor is it clear if the predominant components are cognitive or affective. Again, it would seem that this domain
should be represented in studies of the work environment because those variables mentioned have some promise of mediating the effects of the work environment and/or moderating its impact.

**Type A Behavior**

Because there has been a recent trend toward nonsignificant relationships between Type A behavior and coronary heart disease (Matthews, 1988; Miller, Turner, Tindale, Posavac, & Dugoni, 1991), and because the variables of hostility, cynicism, anger, irritability, and suspicion might be the primary pathogenic component of Type A in any significant associations obtained earlier, then it is hard to maintain a high level of interest in this area of behavior. True, something may be salvaged for the usefulness of the concept by noting that the structured interview yields more valid information than self-report measures, but this only raises the burden of work by including the Type A behavior variable in one's studies. Reviews of studies that examined the possible moderating role of Type A behavior on consequences of work stress (e.g., Parkes, 1993; Payne, 1988) suggest that (a) significant moderating effects are rather uncommon, (b) the type of moderating effect (when obtained) is often not supportive of the general hypothesis that Type A people are more adversely affected by work stress, and (c) none of the studies examined clinical disease outcomes. Type A behavior may have some negative link with social support, a variable briefly discussed next.

**Social Networks and Social Support**

This covers a range of variables from belonging to networks (vs. social isolation) to participating in various social activities to receiving instrumental and emotional support. Although the excitement over this area of research continues unabated (e.g., Berkman, 1995;Dimsdale, 1995), it is important to realize that evidence linking these variables prospectively to incidence of coronary heart disease, net of standard risk factors, is rare and only recent (e.g., Orth-Gomér, Rosengren, & Wilhelmson, 1993). Most of the studies have dealt with mortality as the outcome and many have utilized designs that do not permit an interpretation of precisely where in the disease development schema these variables have an impact. However, it appears likely that the effect is primarily on case fatality, that is, survival after serious medical events (Cohen, 1988); this would seem to apply particularly well to coronary heart disease (e.g., Berkman, Leo-Summers, & Horwitz; 1992; Oxman, Freeman, & Manheimer, 1995). Of course, there is a long tradition of examining social support in the work setting, particularly as a buffer against adverse effects of work stress, but such studies have generally not dealt with cardiovascular disease outcomes (Buunk & Peeters, 1993). The works of Johnson and Hall (1988) and Johnson, Hall, and Theorell (1989) are an exception, but this is also work done within the framework of the job strain model.

**Cardiovascular Reactivity**

The issue here is reactivity as a stable individual difference variable (Manuck, Kasprowicz, Monroe, Larkin, & Kaplan, 1989), which may predict coronary heart disease (alone or in conjunction with work environmental dimensions), rather than reactivity as a strategy for evaluating the acute impact of different work stressors (Gaillard & Wontjes, 1993). (However, data collection strategies, such as 24-h ambulatory monitoring, tend to blur this distinction.) Although there is some controversy regarding the interpretation of the body of evidence (Manuck, Kasprowicz, & Muldoon, 1990; Pickering & Gerin, 1990), even for the narrower issue of blood pressure reactivity and etiology of hypertension, the evidence suggests that the concept of a stable reactivity trait does not have sufficient promise at this time to merit routine inclusion in studies of the work environment and coronary heart disease.

**Specific Aspects of the Work Environment**

I end this section with a brief overview of the early studies of the psychosocial work environment and coronary heart disease. For this I depend on my own summaries of the evidence (Kasl, 1978, 1986, 1989, 1991; Kasl & Amick, 1995). It is interesting to note that the general reviews by C. D. Jenkins from 1971 and 1976 had very little information on the work environment, but by 1982, “excessive workload” emerges tentatively as one of the four clusters of psychosocial risk factors that organize the later review. The McQueen and Siegrist (1982) review also makes references to “subjective work load” and “psychosocial work overload” in their section on cardiovascular disease. It is also worth noting that in spite of early attempts at programmatic and conceptual integration (House, 1974; Orth-Gomér, 1974), the research on work and coronary heart disease remained relatively haphazard until the focus pro-
vided in 1981 by the article on job strain (Karasek, Baker, Marxer, Ahlbom, & Theorell, 1981).

Below, I summarize the evidence with respect to five aspects of the work environment.

1. Studies of occupational differences in cardiovascular disease prevalence and mortality: Aside from the overall impact of occupational status per se, mortality differences in occupations of comparable status were suggestive of various hypotheses; thus college professors and teachers had low rates, whereas lawyers, physicians, and insurance agents and brokers had high rates. Within specialties in medicine, those rated as stressful had higher prevalence rates than those not so rated. Among National Aeronautics and Space Administration employees, managers had higher prevalence rates than did the scientists and engineers. Bank employees of a private commercial bank had higher incidence than bank employees of a more stodgy semipublic savings bank.

2. Hours of work: Working excessive or irregular hours, working in occupations with above average overtime, and holding down two full-time jobs were found to be associated with higher rates of incidence, prevalence, and mortality. However, there was a good deal of inconsistency in these findings and it was difficult to formulate the precise circumstances under which these conditions became pathogenic.

3. Job dissatisfaction: The early suggestive and rather limited evidence (some of it based on aggregate data) that job dissatisfaction may represent another risk factor for coronary heart disease has not been supported by later studies, and most investigators seem to have lost interest in it. At best, this remains, within some occupations, an indirect indicator for work conditions that represent the more direct pathway of underlying cardiovascular risk, but is too crude an index when used across many occupations.

4. Measures of specific dimensions: Much work went into going beyond the global indicators of occupation category, hours of work, and job dissatisfaction to devise measures of specific dimensions that might represent the umbrella concepts of work load and work demands: quantitative overload, high responsibility, deadline pressures, vigilance demands, machine pacing, and so on. However, because the promise of such measures was so frequently tested in cross-sectional designs against biological risk factors (e.g., blood pressure, total cholesterol) rather than clinical outcomes in prospective studies, the evidence has failed to converge and it has been difficult to see where the greatest payoff might be. In addition, much work has been accomplished in linking these dimensions to psychological outcomes, such as symptoms of tension, distress, and dysphoria. Although the convergence of evidence has been greater here, it remains difficult to see whether such relationships reflect, in comparison to coronary heart disease, alternative outcomes, positively correlated outcomes, or unrelated outcomes. That is, observing an impact on psychological outcomes need not be promissory of impact on cardiovascular disease.

5. Other aspects: There are aspects of the work environment that influence coronary heart disease but resist being translated into "psychosocial" terms and appropriate specific dimensions within the "psychosocial" umbrella. These include, among others, shift work, heavy physical work demands, vibrations, noise, and heat (Kristensen, 1995). Shift work is a complex categorization that has characteristics of both psychosocial and physical stressors, whereas the others are part of the "physical work environment" and do not have as yet documented linkages to psychosocial processes.

The Job Strain Model of Karasek and Theorell (1990)

In my editorial accompanying the first job strain article dealing with cardiovascular health (Karasek et al., 1981), I concluded that the formulation "offers considerable promise in furthering our understanding of those aspects of the work environment which have great pathogenic potential because they make adaptation difficult" (Kasl, 1981, p. 683). It is a great pleasure to revisit this research and to comment on the great amount of work that has been accomplished since then.

I need to begin with a few disclaimers: (a) my commentary is necessarily selective and cannot do justice to the richness of the total body of research findings; (b) I refrain from standard methodological criticisms because many reports represent opportunistic analyses and investigators are typically under many practical constraints, unable to set up "ideal" research designs; (c) investigators using the job strain model are not a monolithic group who work together and are responsible for the collective research output. Thus, weaknesses in some studies do not necessarily reflect on other investigators identified with this model.

The original formulation of the job strain model (Karasek et al., 1981) emphasized that the risk for illness was not the result from "an aggregated list of 'stressors' but from the interaction of two types of job characteristics" (p. 694). Curiously, the precise nature of the interaction has never been fully spelled out.
Some of the writings seemed to treat job decision latitude as a buffer: High levels of latitude prevent variations in levels of psychological job demands from increasing the risk of illness, and the risk due to demands will be apparent at only low levels of latitude. Other times, the postulated interaction seems to be one of synergy: Both low latitude and high demands are associated with increased risk but the combination of the two increases the risk beyond mere additive effect.

Admittedly, questions about the nature of interaction and the empirical support for it have been raised before, including by those centrally identified with the model (Kristensen, 1995; Schnall et al., 1994). Karasek (1989) himself seemed to deemphasize the whole issue by arguing that the practical implications for job redesign are similar, irrespective of whether additive or interactive effects are documented. He added that "The primary 'interaction' claimed for the model is that two separate sets of outcomes are jointly predicted by two different combinations of psychological demands and decision latitude" (Karasek, 1989, p. 143); this sentence seems to add confusion by using a somewhat unusual meaning of the term interaction.

It is still possible to ask "How good is the empirical support for the presence of an interaction?" even though it is not clear whether or not a postulated interaction is important to the overall job strain model formulation. Curiously again, this is hard to answer. Theorell (1993) indicated that "true (multiplicative) interactions . . . have been observed in two studies . . . whereas in the others additive interactions were seen" (p. 248). Similarly, Schnall et al. (1994) noted that only two studies found significant multiplicative interactions. However, it is not easy to determine whether the other studies are indeed showing additive interactions. Many of the studies do not present the data in a way that would allow the reader to see the evidence for additive interactions. Typically, measures are combined into a single variable of job strain as the overall risk factor, which thus prevents one from distinguishing between additive effects only and additive interactions. In fact, sometimes one can not even tell whether both variables, or only one variable, contribute to the apparent combined effect of job strain. Theorell and Karasek (1996) asserted that the Johnson et al. (1989) study tested "the multiplicative interaction between all three of them," (p. XX), that is, demand, control, and support. However, a close reading of Johnson et al. (1989) revealed that a single index ("iso-strain") was constructed by multiplying standardized scores on the three scales and then trichotomizing the new scale. Once again, this completely destroys one's ability to see if there is an interaction, and if so, if it is additive or multiplicative. Whether the three scales are simply added or multiplied, the relative placement of individuals into tertiles is going to be extremely similar. The earlier report (Johnson & Hall, 1988) dealt with cross-sectional data and did show good evidence for interaction by displaying the full set of results based on the 27 cells that resulted from trichotomization of the three variables. Thus, it is regrettable that the prospective data in the Johnson et al. (1989) report, which are obviously of greater interest, were not presented in similar detail.

A recent article by Landsbergis, Schnall, Warren, Pickering, & Schwartz (1994) has helped me understand some of the confusion surrounding the interaction debate. On the one hand, the authors state that "the interaction between job demands and job decision latitude . . . has been operationalized primarily in four ways" (p. 350). One common way is the dichotomy of belonging versus not belonging into the job strain quadrant defined by high demands and low decision latitude. But surely such an approach destroys one's ability to detect main effects and any interactions, rather than "operationalizing" the interaction. On the other hand, the authors offer very sophisticated guidelines for analyzing data, including backward elimination of the product term from the full model with main effects and examining the F-statistic for change in $R^2$. So the term has a precise (and correct) meaning for data analysis strategies, and a loose and variable meaning for ways of combining individual variables into a superordinate dimension or categorization. But the latter procedure rules out the former, unless individual dimensions are also kept separate for analysis purposes.

Because the evidence for an interactive relationship between demands and latitude can be characterized as somewhere between limited and unclear, one may ask how this may influence the evaluation of the job strain model. Certainly to the extent that the model has stimulated a large amount of valuable and informative research, it has fulfilled well its heuristic role. However, it also suggests that the model should continue to be viewed as tentative. Instead of testing the model, investigators should still explore possible useful variations on the model and should be encouraged to include other dimensions (as additive or interactive influences) even if these are not formally part of the model or even part of some broader psychosocial formulation.

There is some uncertainty and imprecision about the fundamental concepts in the job strain model. For
example, Theorell and Karasek (1996) referred to one of the two components of decision latitude as skill utilization ("possibility to utilize and develop skills," p. 10) early on, but later the phrase "intellectual discretion" is used instead. Are these phrases meant to be synonymous and interchangeable? It might be argued that they themselves, as the authors of the theoretical model, need to provide sharp definitions of concepts and clear guidelines about their measurement. Granted that reformulations and refinements in the conceptual model are dependent on clarity and convergence of the accumulated empirical evidence. At the same time, such evidence will remain unclear unless it is guided and directed by fairly sharply defined concepts. I believe the culprit here is the fact that from the beginning, a large proportion of the studies was based on secondary analyses of already collected data—a reasonable enough strategy, given the expense of de novo prospective studies of clinical outcomes. But in the process of relying on secondary data, conceptual definitions have remained loose and imprecise as a kind of an attempt to stay flexible and open to possible inclusion of somewhat different dimensions, should the evidence warrant it. It may not be impractical to suggest, however, that because much of the new work going on involves considerably less expensive short-term reactivity studies (including ambulatory blood pressure studies), then these research strategies offer a good setting in which to advance the conceptual and operational precision of the important dimensions involved.

One area of potentially useful clarification concerns the concept of control or decision latitude (are they synonymous?) and its linkage to the other dimensions, psychological demands. It would seem that at the core of the theory is the attempt to get at those aspects of control or decision latitude that enable the individual to more effectively confront or reduce or cope with the demands. If so, then items like "my job allows me to make a lot of decisions on my own" are on target. But what about "my job requires a high level of skill," or "my job requires that I keep learning new things?" These items may reflect both job demands as well as decision latitude, particularly for individuals who do not have the requisite skills or time on the job to learn new things. And why does monotony on the job represent low control? And should one distinguish between repetitive monotony, which is viewed as pathogenic, and uneventful monotony, which is not (Johansson, 1989)? And if control is meant to buffer against negative effects of demands, should researchers not have specific control items for specific aspects of demands, such as (a) scheduling of work (pace, deadlines, hours of work, periods of rest, etc.); (b) influencing conditions at work; (c) choosing methods for carrying out tasks; (d) selecting resources that facilitate task completion; (e) influencing performance evaluation, advancement, and job security; (f) influencing interpersonal aspects of work; (g) choosing the content of one's work.

Theorell and Karasek (1996) noted that "decision latitude is determined to a great extent by the content of work in the occupation, whereas demands and social support to a greater extent local work site conditions and individual perception" (p. 18). This statement needs to be made more forcefully. In the appendix to their book (Karasek & Theorell, 1990), it can be seen that the between-occupation variance in decision latitude is about 35%, but it is only 4% for psychological demands and social support. This is a large difference that should have important implications for future developments of the job strain model.

One important question is what should be done with the concept and measurement of psychological demands. It would seem that what is needed, in part, is the development of measures of work demand that will be considerably more sensitive to differences across occupations (i.e., the between-occupation variance will approach that for decision latitude). And even though decision latitude is fairly closely aligned with occupational status, one would expect that the between-occupation variance for demands will not (should not?) be as closely aligned with status. The issue of the role of "individual perceptions" needs to be clarified as well. The job strain model thus far has maintained a strong focus on the work environment and has not concerned itself with the characteristics of individuals in those jobs. Issues that have agonized the work stress literature, such as objective versus subjective measurement of work demands and work stress (e.g., Frese & Zapf, 1988), have been neglected in the model. Although many studies have used job title averages to impute values to the individual job occupants, thereby converting the average of individual subjective perceptions into a kind of an objective measure, this hardly represents a frontal attack on developing job-based indicators that could be coordinated with the psychological descriptions of working very fast, very hard, having excessive amount of work, and so on. The recent article by Kristensen (1995) discussed some of these measurement issues.

Although Theorell and Karasek (1996) recognized that "individual traits may be associated with
systematically distorted work descriptions, and this systematic distortion may be related to illness risk" (p. 16), they did not evidence any great interest in pursuing this idea further.

As an example, we consider the item "Is your job hectic?" When we ask this question of blue-collar workers in various assembly line and machine-paced jobs it can reflect a specific aspect of pacing, plus some elements of quality control, and allowances for taking breaks. However, when the occupations also include managers, teachers, farmers, doctors, etcetera, then high–low scores on the total study population are very difficult to interpret. In fact, it is not implausible to suggest that those scoring high, but coming from a wide spectrum of jobs, are more likely to share common personal characteristics than common work-setting characteristics. So the needed dual strategy is as follows:

1. To develop measures of the work environment that are indicative of various dimensions of job demands, stated in as concrete terms as possible and avoiding words such as hectic, which capture to a large extent individual reactions rather than the work environment.

2. To develop strategies for identifying traits or person characteristics that put the individual at greater risk of feeling the job is hectic, across many different work conditions. This presumed trait is perhaps more likely to link up with the general research on sleep disturbance and vital exhaustion than with neuroticism, which is an unlikely risk factor for coronary heart disease.

Because the job strain model now incorporates social support in its basic formulation, it is useful to comment on one issue that is worth pursuing and resolving. In the original report on job strain in conjunction with social support (Johnson & Hall, 1988), the measure included five items, all dealing with social interaction with coworkers during work and outside of work. Conversely, the measures of social support from the Quality of Employment Surveys (Karasek & Theorell, 1990) included mostly items reflecting caring and help, both instrumental and emotional. Buunk and Peeters (1993) emphasized the distinction between "rewarding companionship," which appears to be associated with positive feelings, and helpful, supportive interactions, which seem associated with some negative affect, suggesting some costs of being helped. It is worth clarifying, for the job strain–cardiovascular disease association, which component (coworker interaction or support) may be the more important moderator, and whether the two components are differently involved in moderating the effects of high demands versus the effects of low latitude. Because social support effects appear differentially important by gender and for white-collar versus blue-collar workers, the distinction between interaction versus support may also be helpful here. Finally, the dynamics of the spillover effects of work and the process of unwinding after work (Frankenhaeuser, 1991) may be clarified by the distinction between interaction and support.

The High-Effort/Low-Reward Model of Siegrist (1996)

My comments on the Siegrist (1996) article are considerably briefer. The primary reason for this is that the model is of fairly recent origin and the empirical evidence in its support is not as yet extensive. At the same time, the model encompasses a much broader social context than does the job strain model, and it is not quite clear to me what parts of the model are intended for empirical testing and what parts are intended to provide the broad theoretical context for the testable formulations, which themselves are not targeted for empirical challenge. In any case, the model offers exciting and innovative formulations, and I simply feel that it would be premature second-guessing if one questioned, at this stage, the general direction in which this theoretical formulation is heading.

In contrast to the job strain model, which works within a theoretical niche that is somewhat narrower than that of occupational status, the Siegrist (1996) model is ambitiously broad. In encompassing occupational mobility, job insecurity, and unemployment it is, in fact, a broader formulation than the context of occupational status can provide; it also addresses, in part, the health impact of the work role itself, not just the dimensions of the work environment.

One challenge in such a broad formulation is language itself: how to label various measures and combinations of measures and what words to attach to concepts, especially higher order constructs. Readers will inevitably search for familiar labels and assume they are used in familiar ways in order to begin to master the understanding of a new complex theoretical position. I found my understanding of the model hampered by some of the labels. (I trust this commentary would be applicable even if I were able to read all the articles in German.) For example, "immersion" seems similar to older concepts such as job involvement or work commitment, but it also includes need for approval, competitiveness, and disproportional irritability. The conceptual and/or linguistic linkage between "immersion" and "need
for control” is not clear to me. And because Siegrist (1996) developed a new, broad concept of “status control,” then the term need for control acquires a broad new meaning rather than the one with which many readers are familiar. The distinction between extrinsic and intrinsic effort is novel, but I found no similar distinction between extrinsic and intrinsic rewards, one which has been part of the work psychology literature for quite a while. In fact, the Siegrist model works with extrinsic rewards only.

The labeling of some of the measures may also be questioned. For example, one of the two items measuring esteem reward is about “receiving help in difficult conditions” (Siegrist, 1996, p. 32). Most others would label that (instrumental) support and, as the discussion by Buunk and Peeters (1993) suggested, it may relate to lower esteem rather than higher esteem. Shift work is included as one of the contextual indicators of extrinsic effort; but, in fact, shift work is such a complex and unexplained categorical variable, with highly variable meaning depending on the industrial or occupational setting (e.g., much lower effort for air traffic controllers or bus drivers), that its coordination to the dimension of effort seems unwarranted.

Status inconsistency is another interesting label. The Siegrist et al. (1992) article described occupational–educational discrepancy in either direction. It is not clear why any type of such discrepancy exemplifies low reward or low status control. In the traditional occupational psychology literature, particularly that which dealt with the equity concept (e.g., Vroom, 1964), excess of occupational status over educational attainment was viewed as a positive effort–reward imbalance, whereas excess of education over occupational status was seen as negative imbalance, as insufficient upward mobility. So only the latter would seem to fit the model, but it is the former that is more likely to relate to coronary heart disease (e.g., Hinkle et al., 1968). In fact, it might be argued that the broad theoretical model includes some variables (e.g., shift work, noise, sustained anger–irritability, sleep disturbance) that are known to be cardiovascular risk factors but that make the development of a single, tightly integrated conceptual model rather difficult.

In short, the model is an ambitious attempt to integrate a good deal of the sociological literature on heart disease (Siegrist, Siegrist, & Weber, 1986). The research findings reported by Siegrist (1996), and in his earlier publications (Siegrist, 1984; Siegrist, Matschinger, Cremer, & Seidel, 1988; Siegrist, Peter, Georg, Cremer, & Seidel, 1991; Siegrist, Peter, Junge, Cremer, & Seidel, 1990; Siegrist et al., 1992) are impressive, indeed, and suggested the potential power of the theoretical approach. However, because the prospective data so far are based on a selected sample (blue-collar male workers in steel and metal plants, some of which were undergoing reductions in work force), it is not clear in what ways the model will be eventually revised, amplified, or pruned. The very wording of the model, effort–reward imbalance, suggests that reducing the imbalance by reducing effort or increasing reward will reduce the presumed pathogenicity of the imbalance. However, the whole thrust of the model seems to be to emphasize those aspects of effort and reward that are difficult to modify, particularly in low-status occupations. Thus it would be of considerable interest to formulate additional hypotheses regarding the possible effective ways of redressing the imbalance.

Some Comments on Research Design Strategies

The job strain and effort–reward imbalance research programs represent considerable achievements in a difficult domain of research, typically marked by fragmentary studies and uneven progress. Unfortunately, it is the tradition of those who provide overviews and commentary to keep raising the goals and aspirations and never stop with praise alone. In that vein, I express regret about the neglect of two important areas:

1. The measurement of “objective” work demands and work characteristics (e.g., Hacker, 1993; G. D. Jenkins, Nadler, Lawler, & Cammann, 1975): The issue here is not just to get around the possible problems of biased or contaminated self-reports; researchers also need to get more detail about the nature of tasks, the frequency and distribution of the tasks, ergonomic aspects of work, and so on. For example, to capture various psychosocial dimensions of work, the “standardized observations” approach focuses on variety, autonomy, external feedback, task feedback, rigidity, certainty, conflicting demands, interruptions, required skills and abilities, worker pace control, required interdependence, required cooperation, work pressure, employee effort, meaningfulness, resource adequacy, comfort, and task identity (G. D. Jenkins et al., 1975).

2. The measurement of stable personal characteristics: The focus here should not be on those general traits that are often used in the existing literature, such as neuroticism, hardness, locus of control, and Type A behavior. Rather, the emphasis is on specific
dimensions of skills and abilities that are coordinated to the specific analysis of tasks and job demands. In addition, biobehavioral traits may also be important; for example, the effect of shift work on complaints such as sleep depends, in part, on individual differences in diurnal variation in levels of activity (Torsvall & Akerstedt, 1980).

Research that might identify dimensions of the objective environment and of the person’s skills and abilities relevant to cardiovascular health, and then refine its measurement, cannot be carried out in the context of expensive epidemiologic studies of disease incidence. Such epidemiologic designs are also too unwieldy for targeted research testing the usefulness of modifying or expanding the job strain and effort–reward imbalance models. The new field of molecular epidemiology (McMichael, 1994) helps researchers to clarify what they are trying to accomplish: Without studying long-term disease outcomes, researchers are trying to identify biomarkers that (a) can help with the measurement of internal exposure (including the “psychologically effective” dose), (b) can identify an early state of response to the exposure, and (c) can identify individual susceptibilities and effect-modifying host characteristics.

One class of biomarkers often used are cortisol and catecholamines (e.g., Frankenhaeuser, 1991; Gaillard & Wiensjes, 1993), particularly if their joint reactivity is assessed. Although this strategy has been quite useful in identifying effects of such aspects of work as short and repetitious work cycles, lack of control over work pace, and rigid postures, it is not yet clear if changes in cortisol and catecholamines can be viewed as “early warnings of long-term health risk” (Frankenhaeuser, 1991, p. 197). In addition, patterns of reactivity in this class of biomarkers are complex. As Baum and Grunberg (1995) noted:

Some changes appear to be alerting or alarm-oriented and serve to facilitate the initiation of stress responding. Other neuroendocrine changes are supportive of general systemic response, by increasing availability of energy, potentiating response, and/or by facilitating mobilization or recovery (p. 176).

Another class of biomarkers often used are cardiovascular variables, above all blood pressure. As Theorell and Karasek (1996) revealed, an extensive program of research is being carried out involving job strain and acute blood pressure changes. As with the neuroendocrine variables, the blood pressure studies are helpful but not definitive. For example, Schnall et al. (1994) reviewed a variety of studies in which different comparisons were made, such as (a) at work versus at home; (b) at the beginning versus at the end of the work day; (c) acute response to challenge (versus resting level) during work. In addition they reviewed ambulatory blood pressure studies in which high-strain workers were found to have higher blood pressure at work, at home, and during sleep, thus documenting a “carryover” effect. It is worth noting, however, that some comparisons may carry some ambiguity of interpretation. For example, the effect of high-strain jobs (vs. low-strain jobs) is seen often in large differences in blood pressure values at work but small differences at home. But Marmot and Theorell (1988) reported that the impact of having low (vs. high) civil grade jobs was evident in substantial blood pressure differences at home but not at work. Thus a comparison of work versus home can reflect stress at work, degree of unwinding after work, or stress at home. Similarly, large increases in blood pressure during the work day have been linked to high-strain jobs; however; anticipatory reactions at the start of the work day on a high-strain job could lead to initially high values with no further increases. The hypothesized higher reactivity to challenge among workers on high-strain jobs need not be observed; Siegrist (1996) noted that chronic work-related stress may actually attenuate, rather than enhance, acute reactivity.

It thus appears that relevant biomarkers, in the spirit of the new molecular epidemiology, are difficult to come by when we study exposure to high strain or effort–reward imbalance and cardiovascular health. There are no p53 gene mutations (McMichael, 1994) in this difficult field! Given that neuroendocrine and blood pressure reactivity do not represent definitive strategies for studying proximate or intermediate impact of work exposure, more use should be made of other variables including lipids and lipid fractions, fibrinogen, cigarette smoking, physical inactivity, symptoms of tension and distress, and illness absences.

Conclusion

The articles by Theorell and Karasek (1996) and by Siegrist (1996) reviewed an impressive body of research in a difficult and complex area. Although it may be easy to begin a debate about the exact nature of the evidence and its interpretation or about where to go next, there is no question that the work represents a major contribution to the occupational health literature. And because this is primarily European work, my hope is that the U.S. readers of this U.S. journal will be inspired to join in this research effort.
References


Received June 13, 1995
Revision received July 14, 1995
Accepted July 14, 1995