Adverse Health Effects of High-Effort/Low-Reward Conditions

Johannes Siegrist
University of Düsseldorf

In addition to the person-environment fit model (J. R. French, R. D. Caplan, & R. V. Harrison, 1982) and the demand-control model (R. A. Karasek & T. Theorell, 1990), a third theoretical concept is proposed to assess adverse health effects of stressful experience at work: the effort–reward imbalance model. The focus of this model is on reciprocity of exchange in occupational life where high-cost/low-gain conditions are considered particularly stressful. Variables measuring low reward, in terms of low status control (e.g., lack of promotion prospects, job insecurity) in association with high extrinsic (e.g., work pressure) or intrinsic (personal coping pattern, e.g., high need for control) effort independently predict new cardiovascular events in a prospective study on blue-collar men. Furthermore, these variables partly explain prevalence of cardiovascular risk factors (hypertension, atherogenic lipids) in 2 independent studies. Studying adverse health effects of high-effort/low-reward conditions seems well justified, especially in view of recent developments of the labor market.

Remarkable progress has been achieved during the past 20 years in the study of associations between work and health in general, and between psychosocial hazards at work and adverse health outcomes in particular (e.g., Cooper & Payne 1988, 1991; French, Caplan, & Harrison, 1982; Hackman & Oldham, 1980; Hall, 1990; House, 1981; Johnson & Johannson, 1991; Kahn, 1981; Karasek & Theorell, 1990; La Ferla & Levi, 1993; Sauter, Hurrell, & Cooper, 1989; Steptoe & Appels, 1989; Warr, 1987). These achievements are impressive given the many difficulties that are inherent in this type of research. However, a closer and more critical analysis of the current state of art also reveals significant limitations. These limitations can be attributed to one or several of the following challenges that still prevail in this field of scientific inquiry.

The first challenge concerns the difficulty of computing the knowledge from a wide range of disciplines dealing with the issue of work and health. In particular, material from diverse health outcomes such as social, health, and organizational psychology; occupational sociology; and epidemiology, psychosomatic, and behavioral medicine has to be integrated. Yet, integration requires more than just an additive combination of available information. Ideally, a concept is needed that links the following three types of information in a comprehensive way: (a) sociological information describing the work setting or environment; (b) psychological information describing relevant person characteristics (skills, coping processes, etc.); (c) biological information describing the immediate or long-term health consequences. This integration calls for a theoretical approach (see below), and it requires the application of adequate research designs.

The development and testing of adequate research designs in an area that is basically nonexperimental defines a second challenge: How can researchers make sure that the cause–effect associations are tested, or that the time period of studying exposure impact is adequate, or that relevant confounding conditions are assessed? At different occasions, epidemiologist Kasl has discussed these challenges in a critical and seminal way (Kasl, 1989, 1991, 1993).

As mentioned, a third challenge concerns theory. Which critical components of the working life do affect human health? To what extent are these effects attributed to the extrinsic work environment, to the individual person, or to a specific interaction between person and environment? How is the intensity and chronicity of stressful experience maintained in a person's job career, and how is it transduced into bodily dysfunction and disease (Weiner, 1992)? These are important questions that still wait for definite answers.

In this article I discuss these questions within a specific theoretical framework, the framework of high-effort/low-reward conditions at work. The experience of an imbalance between high effort and low reward received at work is assumed to be
particularly stressful as this imbalance violates core expectations about reciprocity and adequate exchange in a crucial area of social life. This notion is elaborated in more detail below (see The Model of Effort–Reward Imbalance at Work section) and empirical evidence on the effects of effort–reward imbalance on human health, with particular reference to cardiovascular health, is summarized (see Empirical Evidence section).

The theoretical questions mentioned above have been dealt with previously in other, though related conceptual frameworks. The two most important and empirically most successful conceptual approaches are the person–environment fit model developed by French et al. (1982) and the demand–control model developed by Karasek (1979) and elaborated further by Karasek and Theorell (1990) and by Johnson and Hall (1988). In the following section, I briefly discuss these models by pointing out their strengths as well as what I consider their open questions. However, this discussion is restricted to a stress–theoretical perspective dealing with links between the social environment, the psychological characteristics, and processes of a person and the organism. Therefore, the discussion does not adequately reflect the rich implications that are inherent in each one of these models.

The Stressfulness of Incongruence and Control at Work

The person–environment fit model has opened a new, important view on the role of work in human life by stressing the interplay between objective and subjective components both of the work environment and of the person. Two types of incongruence that may result from these components are of special relevance for health: the experience of an incongruence ("misfit") between a person’s abilities and the demands of his or her job, and the experience of an incongruence between a person’s goals or aspirations and the supplies offered by the work environment (French et al., 1982; Harrison, 1978; Kahn, 1981). It is important to mention that the person’s appraisal of this incongruence triggers his or her coping (or defense) mechanisms and related strain reactions. This view is in accordance with a widely prevalent psychological theory of stress (Lazarus, 1991).

The model has been elaborated to an impressive degree, and the direct and indirect interactive effects among its crucial variables were specified. However, in a stress–theoretical perspective, several questions have not yet been answered. One question concerns the relevance of job dimensions involved. Does it matter what components of the job environment contribute to the misfit experience? Is there any implicit association between the job components under study and what is commonly considered a crucial strain dimension such as “control” or “threat”? Does the impact of the work environment on strain experience vary from person to person? Another question relates to the chronicity of strain experience: If “perceived misfit” is the important condition, then why do individuals not either alter their environment or adapt their cognitions to this misfit? Moreover, one may ask why this model does not specify those characteristics of personal coping that critically enhance the intensity of strain reactions and associated disease vulnerability.

As indicated, there is no doubt that the concept of control plays a crucial role in our understanding of general relationships between stressors and strain reactions. This was convincingly demonstrated in experimental animal research (Henry & Stephens, 1977) and in psychophysiological studies in humans (Frankenhaeuser, 1979). Control is a major dimension in the second theoretical concept to be discussed here, the influential demand–control model of work stress (Karasek, 1979; Karasek & Theorell, 1990). It is not always clear, however, what “control” means in this context. Kasl (1989) and Parkes (1989) pointed out that interpretation of the control construct depends on the particular focus of respective studies and the research tradition in which they are embedded. According to Parkes, at least three different approaches to defining control in the work environment can be identified: (1.) control as an objective characteristic of the work situation, reflecting the extent to which the design of work tasks . . . allows opportunities for control; (2.) control as a subjective evaluation reflecting an individual’s judgement about the extent to which his or her work situation is amenable to control; and (3.) control as a generalized belief on the part of an individual about the extent to which important outcomes . . . are controllable” (Parkes, 1989, p. 21f).

These three aspects are often not distinguished with sufficient clarity. Even approaches that clearly favor the notion of control as an objective work characteristic tend to merge a variety of interrelated phenomena to one single conceptual scheme (for a detailed analysis of this argument, see Kasl, 1989). Moreover, such approaches often tend to disregard the range of unexplained phenomena when neglecting dimensions of personal control.

For instance, the demand–control model explicitly
restricts its notion of control to objective task characteristics in terms of decision authority and skill discretion. By doing so, variations in physiological arousal remain unexplained, which may be due to personal modes of coping with limited control. Such modes of coping include ways of changing one's level of aspiration, modifying one's degree of job involvement, reducing the amount of effort spent and distancing at the cognitive or emotional level.

In his thoughtful review of the demand–control concept, Kasl (1989) illustrated this point by quoting a number of research findings where elevated levels of strain in otherwise homogenous working groups were restricted to those individuals who exhibited high work investments or to those who failed to realize their aspirations. The role of individual differences is further substantiated in studies summarized by Parkes (1989, 1991) and by Cooper and Payne (1991). Such findings call for a conceptual clarification of the relationship between control-limiting job conditions and those personal characteristics that influence the perception of and the search for control.

In this respect, the concept "need for control" as a distinct individual pattern of coping with work demands was developed (Matschinger et al., 1986, Siegrist & Matschinger, 1989). This concept evolved from a critical analysis of the rather global pattern of Type A behavior. Need for control specifies those cognitive, emotional, and motivational components within the global concept of Type A behavior that are suspected of triggering enhanced arousal in demanding situations: Individuals who score high on measures of need for control often tend to misjudge (i.e., overestimate or underestimate) demanding stimuli in their personal perception. It seems that both types of misjudgement are instrumental in eliciting excessive efforts and in providing opportunities to experience approval, success, and dominance. Although self-rewarding and successful over a period of years in adult life, and especially so in occupational life, high levels of need for control in the long run may precipitate states of exhaustion and physiologic breakdown (see below, see also Appels & Mulder, 1989).

There is no doubt that the demand–control model, with its clear focus on the structure of task profiles and its impact on health, is of utmost importance. A large majority of empirical tests of this model were successful, especially so with respect to cardiovascular disease (Schmall, Landsbergis, & Becker, 1994), and the model proved to be helpful in implementing structural changes of work organization in a number of enterprises (Karasek & Theorell, 1990). Despite these merits a further relevant question still remains to be answered: How does exposure to a high-demand/low-control job elicit chronically stressful experience? The authors themselves give the following answer: Lack of control over how to meet the job's demands and how one can use one's skills defines a state of arousal that inhibits learning; strain-induced inhibition of learning, in turn, further increases arousal by impairing confidence and self-esteem (Karasek & Theorell, 1990). However, one may ask how inhibition of learning is associated with long-term physiologic activation. Again, one obvious reaction to this situation would be to restrict one's amount of effort, to adapt cognitively, emotionally, and motivationally to this rather unfavorable task profile.

In conclusion, despite their undisputed merits, both conceptual approaches, the person–environment fit model and the demand–control model, leave researchers with some unresolved questions, especially those concerning the chronicity of stressful experience, the meaning of "control," and the role of individual coping characteristics. I do not claim to answer these questions, but I want to demonstrate how they can be approached in a somewhat different conceptual framework.

The Model of Effort–Reward Imbalance at Work

By studying adverse health effects of high-effort/low-reward conditions at work, I shift the focus of analysis from control to reward. What does this shift mean in terms of stress theory? To answer this question, the basic arguments of my approach need to be developed. I maintain that the work role in adult life defines a crucial link between self-regulatory functions such as self-esteem and self-efficacy and the social opportunity structure. In particular, the availability of an occupational status is associated with recurrent options of contributing and performing, of being rewarded or esteemed, and of belonging to some significant group (e.g., work colleagues). Yet these potentially beneficial effects of the work role on emotional and motivational self-regulation are contingent on a basic prerequisite of exchange in social life, that is, reciprocity. Effort at work is spent as part of a socially organized exchange process to which society at large contributes in terms of rewards. Societal rewards are distributed by three transmitter systems to the working population: money, esteem, and status control (see Figure 1). The model of effort–reward
imbalance claims that lack of reciprocity between costs and gains (i.e., high-cost/low-gain conditions), define a state of emotional distress with special propensity to autonomic arousal and associated strain reactions. Before explaining why sustained emotional distress is likely to occur under such conditions, the term status control needs to be introduced in more detail.

The notion of status control evolved from my interest in those aspects of occupational life that threaten a person’s self-regulatory functions, his or her sense of mastery, efficacy, and esteem by evoking strong recurrent negative emotions of fear, anger, or irritation. According to sociological theories of self and identity (Mead, 1934; Schutz, 1962–1964) such threats are likely to occur if the continuity of crucial social roles is interrupted or lost. Under these circumstances, control over basic interpersonal rewards is restricted, and as a consequence, self-esteem and emotional well-being are impaired.

For a large part of the adult population, occupational positions provide one such crucial social role. Threats to the continuity of occupational roles are assumed to produce sustained emotional distress. Most clearly, this is the case with job termination or job instability. However, related conditions of low reward and low security in occupational life may also be identified, such as forced occupational change, downward mobility, lack of promotion prospects, or jobs held with inconsistent educational background (status inconsistency). In all these conditions of low occupational status control in combination with high effort, basic reciprocity of costs and gains is lacking. Therefore, having a demanding, but unstable job, achieving at high level without being offered any promotion prospects, are examples of a particularly stressful working context.

With respect to the notion of status control, two important differences between the effort–reward imbalance model and the demand–control model should be noted here. First, in stress–theoretical terms, a difference is likely to exist between the costs of adaptation to the two conditions of low control: It seems less costly to cognitively adapt to a low level of task control than to adapt to a low level of status control, simply because in the former condition fewer fundamental threats are involved. Second, in terms of current developments of the labor market in a global economy, the emphasis on status control reflects the growing importance of fragmented job careers, of job instability, redundancy, and forced occupational mobility. Under these conditions, concerns about status control among the labor force to some extent may override concerns about task control.

In Figure 1, three dimensions of occupational gratifications are distinguished: money, esteem or approval, and status control. Although I discussed the dimension of status control in some detail, it is nevertheless obvious that inadequate payment and lack of esteem and approval in association with high effort are similarly distressing experiences. In all these instances, high-cost/low-gain conditions are likely to elicit recurrent feelings of threat, anger, and depression or demoralization, which in turn evoke sustained autonomic arousal.

Although I have already answered the question of why high-cost/low-gain conditions at work are considered particularly stressful, two additional explications need to be given. First, in line with the concept depicted in Figure 1, I define two different sources of high effort at work, an extrinsic source, the demands on the job, and an intrinsic source, the motivations of the individual worker in a demanding situation. In this latter regard, I have already introduced the concept of need for control as a personal pattern of coping with the demands at work. It is likely that persons with high need for control spend high costs in terms of energy mobilization and job involvement even under conditions of relatively low gain. This may be explained partly by the characteristics of their perceptual and attributable style, partly by the self-gratifying experience of “being in control” of a challenging situation. Therefore, an adequate assessment of the “high cost” part of the equation requires information about either source of effort, extrinsic and intrinsic.
The second explication concerns our answer to the question of why people should engage in such unfavorable trade-offs in their working life. In fact, a well-known psychological theory predicts that effort–reward imbalance is not maintained over a longer period of time and, thus, may not be of pathophysiological importance. I briefly discuss the argument. The expectancy value theory of motivation assumes that rational choice operates in individuals to achieve and maintain a balance between energy consumption and reward experience (Schönpflug & Batmann, 1989). High-cost/low-gain conditions are likely to be avoided or dismissed to maximize one’s profit. At least, reduced expectancy operates to minimize one’s efforts. This theory may be valid in many instances. Yet it does not take into account the social constraints under which individuals must take their decisions, especially the constraints associated with low occupational status control.

For instance, blue-collar workers with reduced opportunities of changing jobs will not minimize their effort at work even if their gain is low. The reason for this behavior is obvious: The possible costs produced by disengagement (e.g., the risk of being laid off or of facing downward mobility) by far outweigh the costs of accepting inadequate benefits. Thus, I would maintain that under defined conditions of low occupational status control, effort–reward imbalance is maintained contrary to the prediction derived from the expectancy value theory of motivation.

A similar argument can be found in the classical writings of John Stuart Mill (1848/1965) who challenged Adam Smith’s theory of compensatory wage differentials by the following argument:

The really exhausting and the really repulsive labours, instead of being better paid than others, are almost invariably paid the worst of all, because performed by those who have no choice. The inequalities of wages are generally in an opposite direction to the equitable principle of compensation erroneously represented by Adam Smith as the general law of the remuneration of labour. The hardships and the earnings, instead of being directly proportional, as in any just arrangements of society they would be, are generally in an inverse ratio to one another (Mill, 1848/1965, p. 383).

To summarize: High-cost/low-gain conditions at work are likely to occur in those groups of the workforce that exhibit a low level of occupational status control.

However, among higher status groups these conditions may be prevalent as well. One example concerns persons who for strategic reasons assume extra work and additional responsibilities to compete for promotion prospects. Failed aspirations after years of excessive effort were shown to be a frequent psychosocial risk constellation among victims of premature myocardial infarction (Siegrist, Dittman, Rittner, & Weber, 1982). In summary, the model of effort–reward imbalance applies to a wide range of occupational arrangements, most markedly to groups that suffer from a growing segmentation of the labor market (Doeringer & Piore, 1971), to groups exposed to structural unemployment and rapid socioeconomic change, to some extent also to groups that are involved in highly competitive career development.

A final point of my argument concerns the pathways of affective processing in high-cost/low-gain conditions. According to a widely discussed psychological theory, the cognitive theory of emotion developed by Lazarus (1991), cognitive appraisal or evaluation of an experienced stressor precedes any form of emotional response. In this view, negative emotions are the result of a multistage appraisal process, which includes the taxing of stressor properties and of a person’s coping repertoire under exposure. Negative affect is considered a common reaction to conditions that exceed a person’s coping abilities and thus threaten her or his self. Again, this theory would predict cognitive and behavioral adjustment to a high-cost/low-gain condition as a consequence of cognitive appraisal processes.

A recent debate on cognitive theory of emotion revealed some limitations of this approach. There is growing evidence of rapid and direct pathways of affective information processing that bypass neocortical–limbic structures and, thus, are not subjected to conscious awareness (LeDoux, 1987). Moreover, affective processing in limbic structures was shown to modulate neocortically organized patterns of social cognition in humans (Adolphs, Tranel, Damasio, & Damasio, 1994).

Therefore, it is likely that affective processing is quite different from conscious computational processing. Or, as Gaillard and Wintjes (1993) argued,

In contrast to computational processing we have no control over the way in which emotional aspects of the information are processed. These processes are encapsulated and are largely unconscious. Only the results of this processing reach our consciousness. We may even feel anxious although we do not know why.... It is hardly possible to disregard the signals that are sent by our emotions. Strong negative emotions, in particular, have ‘control precedence’ relative to other signals reaching our consciousness (Gaillard & Wintjes, 1993, p. 268).

In this perspective, negative affect associated with the experience of effort–reward imbalance at work may not necessarily be subjected to conscious appraisal, especially as it is a chronically recurrent everyday experience.
Empirical Evidence

In the first, major part of the Empirical Evidence section, I summarize existing evidence in favor of the explanatory power of the effort–reward imbalance model described above. This evidence is drawn from a series of studies conducted under my responsibility as a principal investigator (Peter, Siegrist, Stork, Mann, & Labrot, 1991; Siegrist, Bernhardt, Feng, & Schettler, 1990; Siegrist, Peter, Junge, Cremer, & Seidel, 1990; Siegrist, Peter, Motz, & Strauer, 1992). Main criterion variables under study are new cardiovascular events (see Cardiovascular Disease section) and cardiovascular risk factors. In addition, associations between effort–reward imbalance and cardiovascular and hormonal reactions to an acute mental challenge are explored. In the second part (see Indirect Evidence section), some indirect evidence is put together, that is, research findings from other groups that did not explicitly measure the effort–reward imbalance model but that to some extent fit with its core notions.

Direct Evidence

Before presenting results, I indicate how the model was measured and what types of studies were performed. In social science research on work and health, three sources of information are usually available: (a) contextual information derived from independent sources such as administrative data, objective measurements; (b) descriptive information obtained from workers through structured interviews or questionnaires; and (c) evaluative information reflecting subjective appraisal (obtained through interviews and questionnaires). For two reasons, this measurement approach combines these three different sources of information. First, the theoretical argument requires a combination of evaluative and descriptive or contextual information to assess the extrinsic and intrinsic components of the model. Second, the methodological principle of triangulation is followed to secure the validity of these measures.

The following list of indicators defines the core measures of effort–reward imbalance at work.

Intrinsic effort. The coping pattern immersion is measured by a psychometric scale termed need for control (Siegrist & Matschinger, 1989). This scale contains 45 dichotomous items. By means of confirmatory factor analysis, two latent factors were repeatedly found: “vigor” and “immersion.” According to my theoretical assumption, the latter factor defines a critical style of coping with demands reflecting frustrated, but continued, efforts and associ-
activation and thus to impair the cardiovascular system in the long run.

Study samples. The two study samples from which the following results were obtained and the research designed as follows. First, we conducted a 6.5 year prospective study of a cohort of 416 male blue-collar workers (ages 25–55, \( M = 40.8 \pm 9.7 \)). All men were free from overt coronary heart disease at entry. Medical and psychosocial data were collected at entry and three times during follow-up. Baseline psychosocial measures were used to explain prevalence and change over time in major coronary risk factors and to predict new clinical events (Siegrist, Peter, Junge, et al., 1990; Siegrist, Peter, Motz, et al., 1992).

The second study to be reported was a cross-sectional analysis of associations between indicators of effort–reward imbalance at work and major coronary risk factors such as hypertension, elevated fibrinogen, elevated atherogenic lipids, and smoking in a sample of male middle managers (\( n = 179; \) ages 40–55; \( M = 48.5 \pm 4.5 \)). This sample was remarkably homogeneous in terms of age and occupational status, and it was representative of the total group of middle managers of this age group in the enterprise (Peter, Siegrist, Stork, et al., 1991). In both studies, epidemiological and clinical information was combined with psychophysiological information derived from a standardized psychomental stress test (Klein, 1990, 1995).

I restrict this review to the presentation of findings from these two studies as they represent the research that fully assessed all relevant notions of the theoretical model. It is well-known that in epidemiological studies a trade-off is needed between the social scientist’s research interests and the constraints of time, personnel, and money in conducting field studies. In this respect, parts of these measures were also included in two large-scale prospective studies, allowing only partial testing of the main research hypotheses.

One such study was conducted in a cohort of some 1,100 Chinese industrial workers in the city of Wuhan who were followed over a period of 5 years. Measurements had to be largely restricted to contextual and descriptive information. Yet, this study revealed an interesting finding in terms of high-effort/low-reward conditions: systolic blood pressure and serum cholesterol significantly increased during the 5 years in the subgroup of workers who were recently allowed to extend their working hours and were paid overtime and productivity bonuses. This augmented pressure at work was associated with an increase in job insecurity and uncertainty about further promotion prospects. The observed effects were adjusted for important confounders such as age, body weight, smoking, and alcohol consumption (Siegrist, Bernhardt, Feng, et al., 1990).

A second prospective study that includes only a part of the measures was conducted in a cohort of some 4,000 industrial workers in Germany. Final results of this study are not yet fully published, but they demonstrate an independent effect of high intrinsic effort at work (as measured by a short version of the “need for control” scale) on incidence of coronary events in a 5-year observation period (Cremer et al., 1991, p. 59).

Cardiovascular Disease

A relevant test of the explanatory power of a theoretical model concerns its ability to statistically predict disease manifestation. In prospectively designed epidemiological studies it is possible to predict incidence of clinical endpoints (e.g., as in this case, acute myocardial infarction, AMI, or sudden cardiac death, SCD; International Classification of Diseases 410–414) by using baseline information on the model and by adjusting the observed effects for important confounders such as age, smoking, blood pressure, cholesterol, or body weight. As a well-established statistical procedure, logistic regression analysis serves to estimate the odds ratios of relevant predicting variables. In these analyses, the model fit of the most parsimonious model is tested by the likelihood ratio difference test (Hennekens & Buring, 1987).

Table 1 presents a summary of findings derived from the prospective blue-collar study. The predictive power of indicators of high effort and low reward at work is indicated in terms of multivariate odds ratios and 95% confidence intervals of respective variables. I have analyzed these associations with three interrelated sets of outcome criteria: (a) AMI or SCD; (b) AMI or SCD and subclinical coronary heart disease (CHD; i.e., CHD as documented by electrocardiogram without meeting criteria of primary endpoint, total cases \( n = 42 \)); (c) AMI, SCD, or stroke. As the three groups are overlapping to some extent, findings cannot be interpreted as being independent. Rather, they underscore the relative consistency of respective results.

As can be seen from Table 1, two indicators of high effort and two indicators of low reward at work independently predict new coronary events (AMI or SCD). The magnitude of these odds ratios is comparable although the confidence intervals are
Table 1
Odds Ratios From Multivariate Logistic Regression Analysis Explaining Cardiovascular Disease by Indicators of Effort–Reward Imbalance at Work

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMI or SCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work pressure (extrinsic effort)</td>
<td>3.45</td>
<td>0.97–12.30</td>
</tr>
<tr>
<td>Immersion (intrinsic effort)</td>
<td>4.53</td>
<td>1.15–17.80</td>
</tr>
<tr>
<td>Status inconsistency (reward, status control)</td>
<td>4.40</td>
<td>1.36–14.20</td>
</tr>
<tr>
<td>Job insecurity (reward, status control)</td>
<td>3.41</td>
<td>0.81–14.50</td>
</tr>
<tr>
<td>AMI, SCD, or subclinical CHD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work pressure (extrinsic effort)</td>
<td>2.54</td>
<td>1.08–5.98</td>
</tr>
<tr>
<td>Immersion (intrinsic effort)</td>
<td>2.30</td>
<td>0.86–6.14</td>
</tr>
<tr>
<td>Status inconsistency (reward, status control)</td>
<td>2.05</td>
<td>0.90–4.70</td>
</tr>
<tr>
<td>Combined effect of effort–reward imbalance*</td>
<td>6.15</td>
<td>2.01–18.82</td>
</tr>
<tr>
<td>AMI, SCD, or stroke</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immersion (intrinsic effort)</td>
<td>3.57</td>
<td>1.22–10.47</td>
</tr>
<tr>
<td>Status inconsistency (reward, status control)</td>
<td>2.86</td>
<td>1.04–7.80</td>
</tr>
<tr>
<td>Combined effect of effort–reward imbalance*</td>
<td>8.24</td>
<td>2.34–28.43</td>
</tr>
</tbody>
</table>

Note. At entry, \( n = 416 \) male blue-collar workers. CI = confidence interval; AMI = acute myocardial infarction; SCD = sudden cardiac death; CHD = coronary heart disease.

*For detailed information and interpretation, see Direct Evidence section.

quite large. This latter fact may be due to the small number of cases in this sample. Table 1 also contains information on the odds ratios of respective psychosocial predictors of clinical and subclinical CHD as well as of CHD and stroke.

In keeping with the theoretical assumption, I explored the cumulative effect of a simultaneous manifestation of high effort and low reward in CHD patients versus people who remained free from CHD. This was done according to the following two procedures: First, I introduced respective interaction terms in multivariate logistic regression analysis. This approach, in general, failed to produce expected results. A closer inspection revealed that the relatively small number of cases contributed to the production of a series of "zero cells" in the calculations, which may have had an adverse impact on outcomes. The second procedure consisted of computing an aggregate variable that combined all relevant predicting information regarding effort–reward imbalance according to the following three categories: (a) at least one indicator of high effort and at least one indicator of low reward is present; (b) either one (or more) indicator of high effort or one (or more) indicator of low reward is present; (c) neither (a) nor (b) are observed. I expected this three-categorial variable to produce substantially elevated odds ratios in logistic regression analysis compared with the odds ratios produced by single predicting variables.

In Table 1, two examples of this latter strategy are given. First, with respect to clinical and subclinical CHD, the observed effect of the multivariate odds ratio produced by the aggregate measure (6.15) is clearly more powerful than the effects produced by respective single psychosocial variables remaining in the most parsimonious model. A second example concerns CHD and stroke. Here again, the effect of the combined variable (multivariate odds ratio = 8.24) by far exceeds the odds ratios of the two single variables remaining in the model.

Of course, these examples need to be interpreted with caution. The confidence intervals are large (which is why I did not include the results of the analysis of the aggregate measure in the third example, the AMI–SCD group, in Table 1). Moreover, the magnitude of odds ratios cannot be compared directly between the different regression models. Yet, given the relative consistency of the findings and the observation that in all instances conditions of high effort and of low reward at work predict disease outcome, I think it is worth summarizing these results in the suggested way (for details, see Siegrist, 1996; Siegrist, Peter, Junge, et al., 1990; Siegrist, Peter, Motz, & Strauer, 1992).

Cardiovascular Risk Factors

Although there is considerable evidence available on direct effects of sustained distress-induced autonomic activation on atherogenesis and its further
course (Manuck et al., 1991), theoretical models in this area should also be able to contribute to the explanation of main pathways leading to CHD, the established cardiovascular risk factors. The substance of these risk factors is quite heterogeneous, ranging from genetic to lifestyle influences. In the context of this research, I am mainly interested in effects of distress-induced autonomic activation on somatic risk factors such as atherogenic lipids, hypertension, and elevated fibrinogen. Given the fact that such pathophysiological mechanisms are currently discussed in great detail (Henry, 1992; Markovitz & Matthews, 1991), the following question is raised: Is it possible to demonstrate associations between components of the effort–reward imbalance model and these risk factors at the level of statistical analysis both cross-sectionally and longitudinally?

I answer this question by using two different statistical techniques: (a) logistic regression analysis and (b) linear structural equation modeling. I first refer to the approach introduced in the previous section where the construction of an aggregate variable as a proxy measure of high-cost/low-gain conditions at work was explained. It was argued that pronounced effects of this variable should consistently result from logistic regression analysis whatever criterion variable (disease endpoints, risk factors) was of interest.

In Table 2, results of logistic regression analyses with three different cardiovascular risk factors are summarized. These risk factors are (a) hypertension (defined according to World Health Organization criteria), (b) atherogenic lipid level (low density lipoprotein cholesterol >160 mg/dl), and (c) the comanifestation of hypertension and of high atherogenic lipids. This latter condition was shown to elevate the risk of CHD over and above the risks produced by its single components (Castelli & Anderson, 1986).

As can be seen from Table 2, the crucial predicting variables to some extent differ from analysis to analysis, but it is always the combination of at least one indicator of high effort and of at least one indicator of low reward that produces the observed strong effect, irrespective of whether data from the blue-collar study or the white-collar study are analyzed.

These findings were adjusted for a number of relevant confounders, and their robustness was further explored by additional analyses, including a partial replication using a different data set (Peter, 1991; Siegrist, 1996, Siegrist, Peter, Georg, et al.,

<p>| Table 2 |
| Odds Ratios From Multivariate Logistic Regression Analysis Explaining Cardiovascular Risk Factors by Indicators of Effort–Reward Imbalance at Work |</p>
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue-collar workers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined effect of effort–reward imbalance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overtime work (e)</td>
<td>3.29</td>
<td>1.11–9.77</td>
<td>Comanifestation of hypertension and atherogenic lipids</td>
</tr>
<tr>
<td>Cutdown of personnel (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear of job loss (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job instability (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle managers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined effect of effort–reward imbalance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work pressure (e)</td>
<td>3.33</td>
<td>1.22–9.21</td>
<td>Atherogenic lipids (LDL cholesterol &gt;160 mg/dl)</td>
</tr>
<tr>
<td>Frequent interruptions (e)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of reciprocal support (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle managers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined effect of effort–reward imbalance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequent interruptions (e)</td>
<td>6.81</td>
<td>1.70–26.60</td>
<td>Hypertension (SBP ≥160 mmHg and/or DBP ≥90 mmHg)</td>
</tr>
<tr>
<td>Forced job change (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. n = 179 male middle managers and n = 416 male blue-collar workers. CI = confidence interval; e = effort; r = reward; SBP = systolic blood pressure; DBP = diastolic blood pressure.
Moreover, components of the model successfully explained level of fibrinogen in the middle managers study (Peter, Siegrist, Cremer, et al., 1995) and were associated with cigarette smoking (Peter et al., 1991). Nevertheless, these are cross-sectional data, and the results are based on rather crude biomedical variables. Thus, in conclusion, they can be regarded as supporting the general line of research while calling for more detailed consecutive studies.

A second statistical approach was applied in search of a modeling technique that would seem more appropriate to the theoretical model outlined: linear structural equation analysis (LISREL; Jöreskog & Sörbom, 1979). LISREL represents a statistical program which by combining factor analysis and path analysis allows a confirmatory test of a set of effects between latent and manifest variables in different populations. Furthermore, one can test the model fit under certain restrictions for model parameters such as the restriction of invariance of parameters between different groups of observation.

Working with data from the blue-collar study, Siegrist and Matschinger (1989) defined two conceptually different, although empirically not totally independent conditions of low occupational status control: (a) forced piecework (defining blue-collars who for reasons of maintaining their standard of living had to continue piecework) and (b) unskilled or semiskilled job qualification. According to my hypothesis, these conditions describe two powerful social contexts that modulate the intensity of effort-related experience of distress at work. In other words, they postulated an identical structure of effects among variables measuring distress at work for the two respective subgroups of the study sample (workers with forced piecework vs. workers without forced piecework and unskilled or semiskilled workers vs. higher qualified workers, respectively), but they postulated explanatory power of the model concerning its most endogenous criterion, level of blood pressure, in the presence only of the stressful context.

Although a detailed presentation of respective findings is beyond the scope of this review (Siegrist & Matschinger, 1989), major results can be summarized as follows:

1. A model composed by the two latent factors of the construct need for control, by the variables workload, social support at work, sustained anger and hopelessness, and the confounding factors of age, body weight, and cigarette smoking was tested with respect to the amount of explained variance of systolic blood pressure. In the group of workers suffering from forced piecework, the amount of explained variance was 44% compared with 14% in the remaining group.

2. An identical model was tested in the two groups of unskilled or semiskilled workers versus higher qualified workers. Here again, the amount of explained variance of systolic blood pressure was substantially higher in the group with low occupational status control: 54% compared with 27% in the less-distressed, higher status group.

3. A third linear structural model was tested in close association with the first one described above (forced piecework). Yet, in this model, blood pressure measures from two different screenings of the prospective blue-collar cohort were integrated. Again, an acceptable model fit was observed, and unstandardized beta and gamma coefficients were always significant in the stressed group while insignificant in the nonstressed group.

Cardiovascular and Hormonal Reactions

One of the three crucial questions raised in the introductory section was stated as follows: How is stressful experience transduced into bodily dysfunction and disease? Traditionally, epidemiologic studies in the field of psychosocial occupational health are restricted to the analysis of statistical associations between predicting variables measuring work stress and outcome criteria measuring health. Yet, it is crucial to obtain additional information on the biological mechanisms underlying these associations. Ambulatory monitoring techniques and standardized mental stress tests define two methodological traditions in this regard, despite the many restrictions involved in either approach. As mentioned before, we combined mental stress testing with epidemiologic explorations in the blue-collar study as well as in the middle managers study. This combination was given high priority for theoretical reasons, which are briefly summarized as follows (for details, see Klein, 1995; Siegrist, 1996).

1. In the long run, recurrent autonomic activation following the experience of effort–reward imbalance at work is expected to tax the cardiovascular and hormonal systems involved in these responses.

2. As a consequence of long-term taxing, cardiovascular and hormonal reactions to acute challenges may be compromised (i.e., reduced rather than elevated maximal responsiveness is expected to occur). Reduced responsiveness may be modulated peripher-
ally (e.g., via down regulation of beta receptors) or centrally at control sites within the brain.

3. In a cohort of workers of roughly the same age, where additional confounders are controlled for, exposure to high level of chronic effort–reward imbalance at work is associated with reduced maximal cardiovascular and hormonal responsiveness to acute mental stress. This association gives some limited (cross-sectional study design) indirect evidence on (at least one of the) biological processes that may mediate chronically stressful experience to bodily dysfunction.

In fact, Klein (1990, 1995) and Siegrist (1996) found support for systematic associations between measures of chronic stress at work (in terms of high effort–low reward) and reduced responsiveness to a standardized mental stressor (a modified version of the Stroop color–word conflict task) in either study.

Measures of responsiveness in the blue-collar study were restricted to heart rate and blood pressure responses (differences between baseline and maximal stress). After carefully adjusting for a number of confounding factors, significant effects were observed for the following indicators of chronic work stress: (a) high demand in combination with low job security; (b) worsening of job conditions; (c) cumulative workload (Klein, 1990).

In the middle managers study, hormonal measures could be assessed in addition to measures of cardiovascular reactivity. Moreover, it was possible to compute a summary index of high-effort/low-reward experience at work (not including, however, the scale of need for control). Again, results showed the same trend: Middle managers in the upper tertile of chronic work stress were found to react with significantly reduced heart rate, adrenalin, and cortisol reactions to mental challenge if compared with managers defined by lower levels of work stress. Main effects of analysis of variance were adjusted for age, body weight, medication, cigarette smoking, physical inactivity, baseline level of reactivity measure, coffee consumption before test, and diurnal time. Results could not be attributed to test performance or different test evaluation (Klein, 1995).

Although the pathophysiological significance of these results is far from clear, it nevertheless adds to the consistency of reported findings: In addition to predicting cardiovascular events and to explaining prevalence of selected cardiovascular risk factors, measures of high-effort/low-reward conditions at work are consistently associated with reduced reactivity to acute mental challenge. These latter observations may be interpreted in the framework of long-term effects of exposure to chronic stressors on autonomic nervous system regulation.

**Indirect Evidence**

A considerable part of published findings in social-epidemiologic studies on work stress and health to some extent can be interpreted in the frame of the described model. The problem inherent in a respective reinterpretation consists in the selectivity of available information and, of course, in the ex-post nature of respective arguments. Therefore, I restrict this section to a very limited number of studies showing both the limits and the gains of such an enterprise.

In his excellent review, Theorell (1992) recently quoted a number of studies in the field of social epidemiology of CHD whose results at least partly fit with the above mentioned model. For instance, Kornitzer, Kittel, Dromaix, and de Backer (1982) found increased CHD incidence among clerks of a private bank where increase of workload in combination with reduced job security was obvious. In a more recent study, Mattiasson, Lindgärde, Nilsson, and Theorell (1990) observed elevated levels of atherogenic lipids in shipyard employees who were threatened by unemployment. Another example is drawn from a recent study of Johnson and Stewart (1993). When putting together demand–control characteristics in a lifetime perspective in a cohort of workers, Johnson and Stewart observed a decrease in decision latitude–control levels 2–3 years before CHD manifestation in future victims but not in workers who remained free from manifest CHD. This decrease may be analyzed in terms of threats to or loss of occupational status control.

Some but not all earlier studies on occupational downward mobility and forced job change found an elevated risk of CHD (for overview, see Siegrist, 1996). The particular case of job termination and unemployment is difficult to evaluate in this context. First, until recently, there were few studies that conformed to the methodological requirements in this area of research. This has changed recently where at least two independent prospective studies documented an adverse effect of (involuntary) unemployment on cardiovascular health (Martiikainen, 1990; Moser, Goldblatt, Fox, & Jones, 1987). Second, however, virtually no information is available from these studies concerning the effort component of the pre-unemployment job career. Again, I hypothesize
that the predictive power in these studies could be improved if this information was available.

Concluding Remarks

In this review, three relevant questions concerning the links between psychosocial occupational stress and health were discussed.

The first question asked how to identify those components within the global psychosocial occupational environment that are of critical importance to health. It was argued that the help of theory is needed to analytically define those critical components. In this regard, two theoretical models were briefly reviewed, the person–environment fit model and the demand–control model.

The second question asked how chronically stressful experience is maintained in individuals who are exposed to the psychosocial stressors identified in theoretical models. To answer this question, the notions of threat, status control, and reciprocity of exchange in occupational life were introduced. Based on these notions, a third theoretical concept was introduced: the model of effort–reward imbalance at work.

The third question related to the adverse health effects of chronically stressful experience in terms of high effort and low reward. Here the focus was put on cardiovascular health, and the three answers given to this question summarized major findings from two social epidemiologic and psychophysio logic studies on middle aged men conducted in the group. First, the predictive power of components of the model with respect to the incidence of cardiovascular events was demonstrated in terms of multivariate odds ratios (see Cardiovascular Disease section and Table 1). Second, associations of indicators of high effort and low reward with cardiovascular risk factors were presented (see Cardiovascular Risk Factors section and Table 2). Finally, within the limits of cross-sectional psychophysio logic studies, I explored the effects of chronic exposure to high-cost/low-gain conditions at work on cardiovascular and hormonal reactions to a standardized acute mental challenge (see Cardiovascular and Hormonal Reactions section).

In conclusion, high-cost/low-gain conditions at work must be considered a risk constellation for cardiovascular health. Indirect support of this notion came from several related studies that were performed without explicit reference to the model (see Indirect Evidence section).

This analysis also revealed a number of restrictions and open questions, and these should direct future research efforts in this area. In these final remarks, I discuss those restrictions and open questions that seem most important and urgent.

A first set of open questions concerns the operational status of the effort–reward imbalance model. To justify the use of the term model, the relations between the three sets of variables delineated in Figure 1 (extrinsic effort, intrinsic effort, reward) need to be specified further. The summary indices of (or ratios between) variables measuring high effort and low reward, or constructed variables containing combined information to estimate the postulated effects of imbalance on health, was computed. The multiplicative interaction terms in logistic regression analysis was also tested, but this failed to find consistent trends in the relatively small samples. A further statistical approach, linear structural equation modeling, was discussed, and respective findings were summarized.

The relative importance of extrinsic versus intrinsic effort was not specified a priori in this model. It was argued that either component was capable of elevating the risk of stress-related disease if combined with low occupational reward. Similarly, the relative importance of the three reward components was not specified in advance although threats to status control were assumed to produce highest intensity of stressful experience. This assumption was supported by the findings in the blue-collar study (Siegrist, Peter, Junge, et al., 1990). However, in the middle managers study, under conditions of relatively high status control, the two remaining reward components produced similarly powerful statistical effects (Peter et al., 1991; Peter et al., 1995).

The small number of empirical tests performed so far with the effort–reward imbalance model must be considered a second limitation. Evidence so far is restricted to middle-aged working men in advanced western societies. One single study so far has analyzed in part its applicability to a different sociocultural context of Chinese blue-collar workers (Siegrist, Bernhard, Feng, et al., 1990). Several studies are now under way that include employed women and occupational groups belonging to the service sector (e.g., bus drivers, hospital nurses, computer specialists). Specification of measurements according to these occupational contexts in general has not proven to be particularly difficult, and a series of preliminary findings indicate that, in principle, the model is working under these conditions as well.

A discussion of the contribution of the effort–reward imbalance model to explanations of the social gradient of CHD (Marmot, Shipley, & Rose, 1984) is
beyond the scope of this article (see Siegrist 1991, 1996). Yet, it is interesting to note that prediction based on this model, although particularly strong within a blue-collar sample, is not restricted to blue-collar workers but can be extended at least to one group with a higher socioeconomic standing (middle managers). Thus, this offers the opportunity of analyzing differences in CHD both between and within socioeconomic status groups.

A third limitation of this work is its disease specificity. On one hand, having an objective and clear-cut measurement of a relevant endpoint is important because it rules out some of the problems of causal sequence that are prevalent in the area of work stress and mental health. Moreover, research based on the effort–reward imbalance model is expected to contribute to the development of transdisciplinary theories of health and disease because it may offer opportunities to combine biological, psychological, and sociological information within a comprehensive approach. On the other hand, I may have missed information on the predictive power of the model due to the narrow focus of the defined endpoints. In fact, a series of currently unpublished findings indicate that a high ratio of effort–reward imbalance is associated with high level of symptom reporting in bus drivers and with high scores on two out of three “burn out” measures in hospital nurses (Siegrist, 1996). I also have results from the middle managers study showing that some measures of sickness absence are associated with low reward but not with high effort at work (Siegrist, 1996).

This latter observation is of special interest because it may indicate differential predictive power of the model: The specific intensity of negative affect resulting from high effort in combination with low reward may directly result in autonomic arousal and stress-related physiological responses, whereas negative affect associated with low reward only may influence mood, motivation, and behavioral decision making (e.g., the decision to stay away from work). In terms of scientific development, there exists a debate on whether it is desirable for a concept to extend its scope of application to a wide range of phenomena or whether it is more promising to deepen its explanatory potential by restricting the range of phenomena under study.

Some strong arguments are now evolving in support of the latter strategy as new insights into basic regulatory processes of the human organism are becoming available (Weiner, 1992). For instance, Williams (1994) recently pointed to the heuristically fruitful links that start developing between cellular and molecular biology and psychosomatic medicine. More specifically, he pointed to the alteration of macrophage activation following the neurohormonal and immunological changes that are induced by excessive hostility (Williams, 1994). Altered macrophage activation, in turn, may play a role in the pathogenesis of atherosclerosis and in some forms of cancer development.

Another example concerns possible links between stress and endogenous oxidative DNA damage in the human organism. There is now evidence that such damage contributes to atherosclerosis, cancer development, and aging (Ames & Shigenaga, 1993).

These examples challenge the traditional clinical taxonomies and ask for new strategies to define meaningful sets of outcome variables in stress-related studies. Moreover, new markers and mediating processes need to be considered. The potential of these innovations for stress-related research on work and health to researchers' knowledge has not yet been explored.

Finally, coming back to the core question of chronicity of stressful experience raised earlier, more information on the cumulative effects of work-related and extra-work related distress in terms of experiencing high-cost/lowl gain conditions is clearly needed. In this context, beneficial effects of rewarding experiences in private and social life on work-related distress should be explored. In a salutogenic perspective, theories on social support (e.g., Berkman & Syme, 1979; House, Landis, & Umberson, 1988; Johnson & Johansson, 1991) and theories on health-promoting aspects of psychosocial working conditions (Karasek & Theorell, 1990; Siegrist, 1996) still wait for a cross-fertilization.

References

Cooper, C. L., & Payne, R. (Eds.). (1988). Causes, coping...


does not apply.