Injustice at work and incidence of psychiatric morbidity: the Whitehall II study


Background: Previous studies of organisational justice and mental health have mostly examined women and have not examined the effect of change in justice.

Aim: To examine effects of change in the treatment of employees by supervisors (the relational component of organisational justice) on minor psychiatric morbidity, using a cohort with a large proportion of men.

Methods: Data are from the Whitehall II study, a prospective cohort of 10 308 white-collar British civil servants (3143 women and 6895 men, aged 35–55 at baseline) (Phase 1, 1985–88). Employment grade, relational justice, job demands, job control, social support at work, effort–reward imbalance, physical illness, and psychiatric morbidity were measured at baseline. Relational justice was assessed again at Phase 2 (1989–90). The outcome was cases of psychiatric morbidity by Phases 2 and 3 (1991–93) among participants case-free at baseline.

Results: In analyses adjusted for age, grade, and baseline physical illness, women and men exposed to low relational justice at Phase 1 were at higher risk of psychiatric morbidity by Phases 2 and 3. Adjustment for other psychosocial work characteristics, particularly social support and effort–reward imbalance, partially attenuated these associations. A favourable change in justice between Phase 1 and Phase 2 reduced the immediate risk (Phase 2) of psychiatric morbidity, while an adverse change increased the immediate and longer term risk (Phase 3).

Conclusion: This study shows that unfair treatment by supervisors increases risk of poor mental health. It appears that the employers’ duty to ensure that employees are treated fairly at work also has benefits for health.
The prevalence of mental ill-health, including minor psychiatric morbidity, is not equally distributed between the sexes; the prevalence being greater among women than men. However, although minor psychiatric morbidity is a stronger predictor of short spells of psychiatric sickness absence in women, it is a stronger predictor of long spells of absence in men. Reflecting this finding, work from the Renfrew and Paisley study has shown that the long term suicide risk associated with minor psychiatric morbidity is much stronger in men than women. Minor psychiatric morbidity has also been shown to be consistently associated with an increased risk of coronary heart disease in men, but not in women in both the Whitehall II and Renfrew and Paisley studies, and analyses of population data from the Netherlands (the NEMESIS study) concluded that mental disorders are a more important risk factor for sickness absence among men than among women.

In this study we examine associations of relational justice and change in relational justice with mental health, measured as minor psychiatric morbidity, using longitudinal data from the Whitehall II cohort of British civil servants, over two thirds of whom are men.

METHODS

The target population for the Whitehall II study was all London based office staff working in 20 civil service departments between 1985 and 1988 and aged 35–55 on entry to the study. With a response rate of 73%, the final cohort consisted of 10 308 participants (3413 women and 6895 men). The true response rate was higher, however, because around 4% of those invited were not eligible for inclusion. Although mostly white-collar, respondents covered a wide range of grades from office support to permanent administrative staff. The final cohort involved 9 934 respondents (3161 women and 6773 men) who returned a signed consent form.

Baseline screening (Phase 1) took place between late 1985 and early 1988. This involved a clinical examination in which height, weight, blood pressure, and serum cholesterol were determined, among other anthropometric and biomedical measures. A self-administered questionnaire containing sections on demographic characteristics, health, lifestyle factors, work characteristics, social support, life events, and chronic difficulties was completed by each respondent at home and checked for completion at the clinic. Non-responders to the initial invitation, which was delivered to potential participants at their place of work, were followed up to two reminder letters. From Phase 1 onwards participants could indicate whether they wanted to be contacted at work or at home, and non-responders were followed up by telephone as well as by letter. In 1989/90 (Phase 2, response rate 79% of participants at Phase 1), the same questionnaire data were collected by post. The Phase 3 (1992–93, response rate 81% of participants at Phase 1) data collection included a repeat of the clinical examination in addition to a questionnaire. Full details of all contact procedures and all methods are contained in the Whitehall II Phase 3 manual.

Measures

Relational justice

Organisational justice has not been measured directly in the Whitehall II study. However, it has been possible to construct an indicator of relational justice with face validity from questions available in the Phase 1 and 2 surveys. Of the 19 items that deal with relational justice, five items that deal with relational justice were identified using factor analysis.

- Do you get consistent information from line management (your superior)?
Injustice and psychiatric morbidity

- Do you get sufficient information from line management (your superior)?
- When you are having difficulties at work, how often is your superior willing to listen to your problems?
- Do you ever get criticised unfairly? (reverse scored)
- Do you ever get praised for your work?

These five items formed an internally consistent scale of relational justice (Cronbach’s alpha 0.71 at Phase 1, 0.73 at Phase 2). Responses were scored on a four point scale from never (1) to often (4). The mean of scores were scaled from 25 to 100. For example, a participant who selected “often” in answer to every question would score 100 while a participant who selected “never” would score 25. The resulting distribution was divided into tertiles (25–70 = “low”, 71–88 = “intermediate”, 89–100 = “high”). The bottom tertile indicated a low level of relational justice, the top tertile a high level of relational justice, and the middle tertile an intermediate level. Change in relational justice was calculated by deducting the Phase 1 score from the Phase 2 score. Participants were classified into one of three groups: “no change”, “adverse change” (a decline of 10 score points or more), and “favourable change” (an increase of 10 score points or more).

Other psychosocial characteristics of the work environment

The demand–control and the effort–reward imbalance model represent the two main models of organisational stress. Subsequent to its initial conception, social support was added to the demand–control model. Forty self-report questions based on the constructs of the demand–control–support model, derived from well known questionnaires used in the United States and Sweden and the Job Content Questionnaire, were included in the Phase 1 questionnaire.13,35 Reliability analysis and exploratory principal components analysis at Phase 1 showed that some of the items were not reliable measures of the underlying construct. Furthermore, some items were deleted from the scale as they were left unanswered by too many participants. For such reasons 15 of the 40 questions were dropped from successive questionnaires. Of the remaining 25 questions, 15 measure job control or decision authority, 9 of the 15 items cover decision latitude, and 6 cover skill discretion. These subscales were equally weighted (internal consistency, Cronbach’s $\alpha = 0.84$). Job demands were measured using 4 items (internal consistency, Cronbach’s $\alpha = 0.67$), and social support at work, which comprised three components (support from colleagues, support from supervisors, and clarity and consistency of information from supervisors), was measured using 6 items (internal consistency, Cronbach’s $\alpha = 0.79$). A list of all 25 questions has been published previously.37 All questions were answered on a four point scale from “often” to “never/almost never”. Responses on a four point scale from “often” to “never/almost never” were combined into summary scales and then divided into tertiles. In all measures, the bottom tertile indicated a low level, the top tertile a high level, and the middle tertile an intermediate level for each of these indicators.38 The effort–reward model of work stress in its current form was not available at Phase 1, so a measure derived from existing questions was used. Details of the derivation and testing of this measure have been published previously.39

Minor psychiatric morbidity

The 30 item General Health Questionnaire (GHQ) was used to assess psychiatric morbidity.23 GHQ items consist of statements about behavioural and psychological functioning. The respondent is asked to say how well the statement applies to them during “the past few weeks” in comparison...
to their “usual” behaviour or state of mind. For example: “Have you recently felt you are playing a useful part in things?”; “Have you recently felt you couldn’t overcome your difficulties?”.

The response alternatives are: “not at all”, the “same as usual”, “rather more than usual”, and “much more than usual”. As little is gained in terms of case identification by discriminating between the severity of symptoms, questions were scored as follows: “not at all” = 0, “same as usual” = 0, “rather more than usual” = 1, and “much more than usual” = 1. The GHQ-30 has been validated in a number of diverse populations and has been validated specifically against the Clinical Interview Schedule in Whitehall II data, giving a cut-off point of 4/5 for dividing “non-cases” from “cases”. GHQ cases was thus defined as a score of 5 or more. Cases are participants who are at higher risk of minor psychiatric morbidity, largely depression, and anxiety disorders.

**Covariates**

Age and employment grade were derived from the Phase 1 questionnaire. Employment grade was determined from the participant’s Civil Service grade title. For analysis, employment grade titles were divided into three categories in order of decreasing salary: administrative, professional/executive, and clerical/support. A composite indicator of physical illness (physical illness indicator) was comprised of diabetes, diagnosed heart trouble, ECG abnormalities, hypertension, and/or respiratory illness. The category “diabetes” included all diabetics. Data on past medical history of doctor diagnosed coronary heart disease (CHD) were derived from the Phase 1 questionnaire. ECG abnormalities were probable/possible ischaemia identified on ECG during the baseline screening examination. The category “hypertension” included all participants on antihypertensive medication or with a systolic or diastolic blood pressure greater than 160 or 95 mm Hg respectively. Presence of a respiratory illness was detected using the Medical Research Council chronic bronchitis questionnaire.

**Statistical analysis**

Associations between relational justice at Phase 1 and GHQ caseness by Phases 2 and 3 were determined using logistic regression analyses among participants who were not GHQ cases at baseline. The outcome of interest in these analyses at Phase 2 was prevalence of GHQ caseness by Phase 2, while at Phase 3 it was GHQ caseness by Phase 3 (GHQ case at Phase 2 or at Phase 3). Results are presented as odds ratios and 95% confidence intervals adjusted for age in five year categories, employment grade, and physical illness at Phase 1. The physical illness indicator was included as a covariate in order to control for physical illness and reduce the effects of reverse causality. A linear trend between relational justice and GHQ caseness was tested by entering the continuous justice score in the model.

Three of the questions used to construct our relational justice measure are from the social support scale and one from the effort–reward imbalance model. In order to determine whether our measure of relational justice had explanatory power beyond that provided by social support at work and the existing well known work stress models, the next step in our analyses was to adjust for these work characteristics, firstly separately and then all together. The effect of these adjustments was summarised by calculating the percentage change in the linear trend term for relational justice. Only participants with no missing data for any of the work characteristics were included in these models.

The next step in these analyses tested whether favourable or adverse change in relational justice between Phase 1 and Phase 2, compared with no change, predicted GHQ caseness by Phase 2 and Phase 3 after adjustment for age, employment grade, physical illness, and relational justice at Phase 1. Linear trend was tested by entering the change score in the model. All analyses were conducted separately for women and men using the SAS statistical program (SAS Institute, Cary, NC, USA).

Ethical approval for the Whitehall II study was obtained from the University College London Medical School Committee on the ethics of human research.

**RESULTS**

At Phase 1, 2744 participants had psychiatric disorder as indicated by GHQ caseness, and 119 had missing data on the GHQ. Both these groups were excluded from the main analyses. Of the 7445 participants free of psychiatric disorder at Phase 1 (non-GHQ cases), 734 responded to justice items at Phase 1, 5810 also responded to the GHQ at Phase 2, and 9172 additionally responded to the GHQ at Phase 3 (table 1). In the analyses of justice at Phase 1 and GHQ incidence by Phase 2, those 5736 participants who additionally had no missing data in any of the Phase 1 covariates were included (table 2). In the analyses of GHQ caseness by Phase 3, the number of included participants was 5109 (table 3). Analyses of change in justice between Phase 1 and Phase 2 were adjusted for age, grade, physical illness, and relational justice at Phase 1 (but not for the other baseline covariates); among these, 5700 participants had complete data for incident GHQ at Phase 2 and 5086 for incident GHQ by Phase 3 (table 4). The latter population (the group with the greatest attrition) compared with all non-GHQ cases at Phase 1 contained a slightly smaller proportion of women (29% v 31%) and manual workers (19% v 23%), but any differences in the prevalence of physical illness (9% v 10%) and the level of job demands (56.5 v 56.0) and job control (66.4 v 65.5), support at work (77.4 v 76.7), and effort–reward imbalance (1.0 v 1.0) were minimal or non-existent. Participants who were GHQ cases at baseline had lower relational justice score (75.4) than included participants.

As can be seen in table 1, in contrast to the other psychosocial characteristics of the work environment, women and men reported the same mean levels of organisational justice at Phase 1. Furthermore, there were few differences in relational justice scores by age or by grade, with the exception of men in the middle grades who had a slightly lower relational justice score than men in the highest and lowest grades. However, relational justice scores for both sexes were lower at Phase 2 than at baseline but again did not differ between women and men. While statistically significant (p < 0.001), this decline in relational justice was relatively small (0.12 SD).

Tables 2 and 3 show strong associations (p < 0.001) between relational justice at Phase 1 and GHQ caseness by Phases 2 and 3 in both sexes. Among women the association seems to be driven mainly by the significantly higher risk of GHQ caseness associated with low relational justice. Similar results are seen in men, apart from the intermediate relational justice group in which there appears to be a slightly higher risk of GHQ caseness by both phases. With regard to low relational justice, adjustment for age, grade, and the physical illness indicator produces figures little different from those seen in the unadjusted associations.

The association between relational justice and GHQ caseness in women is attenuated by 12% for caseness by Phase 2 and 9% for caseness by Phase 3 when job control is included in the model adjusted for age, grade, and the physical illness indicator. Social support produces little attenuation of the association at Phase 2, but the association with caseness by Phase 3 is attenuated by 30%. When all the other characteristics of the work environment are added into the
model together, the association between relational justice and GHQ caseness in women is attenuated by 16% at Phase 2 and 38% for caseness by Phase 3. Among men, job demands, job control, and social support produce a modest attenuation (6–16%) of the association between relational justice and GHQ caseness by Phase 2 and by Phase 3, whereas effort–reward imbalance attenuates the association by 41% at Phase 2 and by 38% for caseness by Phase 3. At Phase 2 the greatest attenuation, 46%, is seen when all the work characteristics are entered into the model together. However, the test for trend remains statistically significant for the fully adjusted model across both phases and sexes.

Table 4 shows that in both sexes a favourable change in relational justice between Phase 1 and Phase 2 is associated with a decreased risk of incident GHQ caseness at Phase 2 compared with the group that experienced no change. Conversely, an adverse change is associated with an increased risk. In both sexes the effect of a favourable change in relational justice between Phase 1 and Phase 2 on GHQ caseness appears to have weakened by Phase 3. However, the effects of an adverse change remain strong by Phase 3 and the test for linear trend between change in justice and GHQ caseness remains unchanged at both phases (p < 0.001).

**DISCUSSION**

Women and men who reported at baseline that they were treated unfairly by their supervisors had a significantly higher risk of incident psychiatric morbidity three years and six years later. Less than half of this effect appeared to be explained by other psychosocial characteristics of the work environment, the most important explanatory factor being social support at work among women and effort–reward imbalance among men. Overall there was a small decline in the level of relational justice among Whitehall II participants of both sexes over the three years following the baseline survey. Within this overall decline, some participants experienced a favourable change in justice, while others experienced an adverse change. A favourable change was associated with a significant reduction in the immediate risk of psychiatric morbidity, while an adverse change significantly increased both the immediate risk and the longer term risk.

This study benefits from using data from the Whitehall II study, a well-characterised cohort with sufficient power to detect effects in both sexes. It is also, to our knowledge, the first study to examine the effects of change in relational justice on mental health in the workplace. By analysing change in a sub-population comprised of participants who were not GHQ cases at baseline, the study addressed the possibility of reverse causation, that is the situation in which participants with higher GHQ scores are treated less fairly by supervisors, rather than GHQ caseness being the result of such unfair treatment. In the absence of controlled experiments, the examination of change in an exposure over time, while removing prevalent cases of the outcome at baseline, provides the best means of tackling such reverse causation in addition to other problems, such as bias and confounding, which can be controlled in the experimental situation. Adjusting for a measure of physical illness at baseline also reduced the likelihood that our findings could be attributable to pre-existing physical morbidity.

Although using a sub-population of non-GHQ cases at baseline reduces the possibility of reverse causation, it remains possible that borderline GHQ caseness also elicited less fair treatment. To control for the effects of sub-clinical psychiatric morbidity and to determine the effects of relational justice and change in relational justice in the whole study population, we repeated the analyses in a
population that included all participants irrespective of GHQ status at baseline and excluded only the 1247 with missing data. In addition to adjustment for age, grade, and the physical illness indicator at baseline, these analyses were also adjusted for baseline GHQ score. The outcome of interest in these analyses at Phase 2 was prevalence of GHQ caseness at Phase 2, while at Phase 3 it was GHQ caseness at both Phase 2 and Phase 3, an outcome chosen to capture the effect of relational justice on longer term psychiatric morbidity. On adjustment for GHQ score at baseline, in general, these analyses produced findings little different from those presented (tables available from the authors on request). The two main differences were that the association between low relational justice and caseness among women at Phase 2 was attenuated by 32% rather than 16% on adjustment for all the work characteristics, and the test for trend for the change score and GHQ caseness at Phase 3 in women was \( p = 0.02 \) instead of \( p < 0.001 \). Similarly, although GHQ caseness at both Phase 2 and Phase 3 among the total cohort is not strictly comparable to the outcomes used in the main analyses, the findings followed the same pattern seen for GHQ caseness by Phase 3 (GHQ case at Phase 2 or at Phase 3) among participants case-free at baseline. The only deviation was that the association between low relational justice and GHQ caseness among all men at Phase 3 was attenuated by 63% instead of 33% on adjustment for all the work characteristics.

In the present study findings were very similar in both sexes. However, while low relational justice at baseline was a significant predictor of psychiatric morbidity at follow up in both sexes, there was an indication that both low and intermediate relational justice were predictive in men. These findings reflect those of previous work on relational justice. Studies have demonstrated cross-sectional associations with prevalent psychiatric morbidity, and longitudinal association in both sexes between low relational justice at baseline and psychiatric morbidity at follow up, two years later, although relational justice appears not to predict doctor diagnosed psychiatric disorders in women. The findings also follow a similar pattern to those seen between relational justice and self-rated health in the Whitehall II cohort, although the size of the effect of low relational justice and adverse change in relational justice on GHQ caseness is greater than the effect on self-rated health. Given the self-appraisal element involved in both measures one might expect to see some overlap in the findings, but it is important to stress that self-rated health and the GHQ measure different health domains.

With the exception of social support in women and effort–reward imbalance in men, only a modest fraction of the association between relational justice and psychiatric morbidity appeared to be explained by any other psychosocial work characteristic measured in the present study. The finding for social support at work among women is unsurprising given the considerable overlap between our relational justice and social support scales. Previous work that has examined whether associations between relational justice and psychiatric morbidity are independent of other psychosocial work characteristics has reported mixed findings. In results similar to those of the present study, the cross-sectional association between relational justice and psychiatric morbidity was attenuated, but survived, adjustment for job control and social support at work in a cohort of Finnish hospital personnel, mostly women. However, later findings from the same study found that the prospective association between relational justice and psychiatric morbidity did not survive adjustment for decision authority.

No previous study appears to have examined the effect of adjusting the relational justice–health association for effort–reward imbalance. The effort–reward imbalance model is based on the concept of reciprocity, where effort at work is reciprocated by rewards that include salary, esteem, promotion opportunities, and job security. An imbalance between efforts and rewards has been shown adversely to affect self-esteem and predict a range of illnesses in employees. One of the items comprising our relational justice scale comes from the reward construct of the effort–reward model and an examination of the two constructs indicate that both are likely to have similar effects on self-esteem. In addition to being a measure of stress at work, the effort–reward concept could also be viewed as a tolerable proxy for distributive justice. As distributive justice relates to the perceived fairness of formal decision making procedures, this is likely to be highly correlated with perceptions of just treatment by supervisors. Effort–reward imbalance has been examined prospectively in relation to GHQ caseness in Whitehall II and high effort and low reward separately and in combination were found to be predictive of GHQ caseness in men, but not in women.

In the new labour market, characterised by increasing competition, downsizing, and tough, “macho” management styles, workplace bullying, recently described as “the silent epidemic”, appears to be on the increase. Although undocumented as yet, organisational justice is also likely to prove a casualty under such circumstances. During the baseline survey of the Whitehall II study the jobs of British
Main messages

- Employees who perceive that they are treated unfairly by their supervisors are at increased risk of poor mental health.
- An increase in unfair treatment increases the risk of poor mental health.
- An increase in fair treatment reduces the risk of poor mental health.
- Effects are only partially explained by other established occupational stressors.

Policy implications

- Results indicate that the duty of employers to ensure employees are treated fairly at work also has benefits for mental health and wellbeing.

...can be important sources of justice evaluations, because they work together, and share opinions and emotions with each other. Thus, the characteristics and opinions of other members of the work group may have an impact on attitudes and behaviours of the individual employee. Only a few studies have been able to test the effect of the work group. These have reported that employees within the same group share common perceptions of justice in the organisation, as indicated by a significant variation in justice perceptions both within and between work groups.31-33

Injustice and psychiatric morbidity

In such a situation fairness is important to people, because it helps them to deal with uncertainty, in that solid, firmly constructed fairness judgments can help remove uncertainty or at least alleviate much of the discomfort that it would otherwise generate.

Study limitations

The main limitation of this study is the absence of standard measures of organisational justice. Other limitations include loss to follow up, the lack of an externally assessed measure of mental health, and inability to include work group level measures into the analyses.

Standard measures of organisational justice were not available at the time of the baseline survey of the Whitehall II study. Using existing questions we were able to derive a five item measure of relational justice. The Cronbach's alphas for our relational justice scale at Phases 1 and 2 were 0.71 and 0.73 respectively. A Cronbach's alpha of 0.70 is seen as acceptable, and this is particularly true for a short measure assessing a relatively broad concept, as in this study. Although three of the items were those that also form the work social support scale and one that also forms the effort-reward imbalance scale, in both women and men the association between low relational justice and psychiatric morbidity was only partially attenuated by adjustment for social support at work and effort-reward imbalance. This indicates that our measure of relational justice has effects that are independent of both measures. Although we feel that effort-reward imbalance represents a tolerable proxy for distributive justice, there is nothing in the Whitehall II dataset that could be used to construct a measure of procedural justice, the third component of organisational justice. This is unfortunate because procedural justice is an important determinant of health as relational justice.14-16

The main analyses, those presented in tables 2-4, dealt with participants free of psychiatric disorders at baseline, and of this population up to 33% were lost to follow up due to missing data. However, any differences in baseline characteristics between included and excluded civil servants were small, suggesting that a major bias is unlikely.

Although the GHQ is not an externally assessed measure of mental health, it is widely used, has been validated in a number of studies, and has been shown to be at least as important a determinant of psychiatric morbidity as procedural justice. This is unfortunate because procedural justice has been shown to be at least as important a determinant of psychiatric morbidity as relational justice.14

...circulate, providing a possible explanation for the overall decline in perceived relational justice between Phases 1 and 2. In such a situation fairness is important to people, because it helps them to deal with uncertainty, in that solid, firmly constructed fairness judgments can help remove uncertainty or at least alleviate much of the discomfort that it would otherwise generate.

**References**

working at the University of Helsinki, Finland, and JV were supported by the Academy of Finland (projects 77560, 104891, and 105195) and the Finnish Environment Fund, JEF is supported by the MRC (Grant number G8802774), MJS by a grant from the British Heart Foundation, and MGM by an MRC Research Professorship.

Competing interests: none

REFERENCES