

May1, 2007

To: Persons interested in the WHO HPQ absenteeism and presenteeism questions
Fr: Ron Kessler, Maria Petukhova, Keith McInnes, Harvard Medical School
T Bedirhan Üstün, World Health Organization

Re: Content and scoring rules for the WHO HPQ absenteeism and presenteeism questions

I. Overview

A number of researchers have asked whether they can abstract the absenteeism and presenteeism questions from the full HPQ and use these questions alone or in conjunction with another interview. This memo lists the questions that are needed to do this and describes the scoring rules for the absenteeism and presenteeism measures based on this core set of HPQ questions. In describing the use of these questions in publications, authors should refer to the questions as the “absenteeism and presenteeism questions of the World Health Organization’s Health and Work Performance Questionnaire (HPQ)” and should cite one or both of the following two HPQ methodology papers:

Kessler, R.C., Barber, C., Beck, A.L., Berglund, P.A., Cleary, P.D., McKenas, D., Pronk, N.P., Simon, G.E., Stang, P.E., Üstün, T.B., Wang, P.S. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). Journal of Occupational and Environmental Medicine, 45 (2), 156-174.

Kessler, R.C., Ames, M., Hymel, P.A., Loeppke, R., McKenas, D.K., Richling, D., Stang, P.E., Üstün, T.B. (2004). Using the WHO Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. Journal of Occupational and Environmental Medicine, 46(Suppl. 6), S23-S37.

II. The absenteeism and presenteeism questions

**B3. About how many hours altogether did you work in the past 7 days?
(If more than 97, enter 97.)**

Number of hours (00-97)

**B4. How many hours does your employer expect you to work in a typical 7-day week?
(If it varies, estimate the average. If more than 97, enter 97.)**

Number of hours (00-97)

B5. Now please think of your work experiences over the past 4 weeks (28 days). In the spaces provided below, write the number of days you spent in each of the following work situations.

In the past 4 weeks (28 days), how many days did you...

	Number of days (00-28)
B5a. ...miss an entire work day because of problems with your physical or mental health? (Please include only days missed for your own health, not someone else's health.)	<input type="text"/> <input type="text"/>
B5b. ...miss an entire work day for any other reason (including vacation)?	<input type="text"/> <input type="text"/>
B5c. ...miss part of a work day because of problems with your physical or mental health? (Please include only days missed for your own health, not someone else's health.)	<input type="text"/> <input type="text"/>
B5d. ...miss part of a work day for any other reason (including vacation)?	<input type="text"/> <input type="text"/>
B5e. ...come in early, go home late, or work on your day off?	<input type="text"/> <input type="text"/>

B6. About how many hours altogether did you work in the past 4 weeks (28 days)? (See examples below.)

Number of hours in the past 4 weeks (28 days)

Examples for Calculating Hours Worked in the Past 4 Weeks

40 hours per week for 4 weeks = 160 hours
 35 hours per week for 4 weeks = 140 hours
 40 hours per week for 4 weeks with 2 8-hour days missed = 144 hours
 40 hours per week for 4 weeks with 3 4-hour partial days missed = 148 hours
 35 hours per week for 4 weeks with 2 8-hour days missed and 3 4-hour partial days missed = 112 hours

B9. On a scale from 0 to 10 where 0 is the worst job performance anyone could have at your job and 10 is the performance of a top worker, how would you rate the usual performance of most workers in a job similar to yours?

<i>Worst</i> Performance											<i>Top</i> Performance
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B10. Using the same 0-to-10 scale, how would you rate your usual job performance over the past year or two?

<i>Worst</i> Performance											<i>Top</i> Performance
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B11. Using the same 0-to-10 scale, how would you rate your overall job performance on the days you worked during the past 4 weeks (28 days)?

<i>Worst</i> Performance											<i>Top</i> Performance
0	1	2	3	4	5	6	7	8	9	10	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

III. Scoring absenteeism

We have two ways of measuring and scoring absenteeism. One relies on the respondent estimating how many hours he/she worked over a four-week period. The other asks the respondent to estimate how many hours he/she worked in the past 7 days. We recommend using the four week estimates when available, as they will tend to smooth out any up or down spikes that might have occurred in a particular week (e.g. a respondent missing several hours of work because of a doctor's appointment). Both are included, however, because some employees find it hard to estimate the number of hours they worked for a four week period. Here we describe the scoring for both the 7-day and the 4-week estimates. Finally, when both the 7-day and 4-week estimates are available for most workers, some employers compare the results to get a sense of the reliability of the data.

Absenteeism is scored in terms of hours lost per month, which is to say that a high score indicates a higher amount of absenteeism. The measure of absolute absenteeism is expressed in raw hours, with a negative lower bound (if the person works more than expected) and an upper bound equal to the number of hours the respondent is expected to work. The measure of relative absenteeism is expressed as a percentage of expected hours and ranges between a negative number (works more than expected) and 1.0 (always absent).

a) Using 4-Week Estimates

Absolute absenteeism: $4xB4 - B6$

Relative absenteeism: $(4xB4 - B6)/(4xB4)$

Relative hours of work: $B6/(4xB4)$

b) Using 7-day Estimates

Absolute absenteeism: $4xB4 - 4xB3$

Relative absenteeism: $(4xB4 - 4xB3)/4xB4$

Relative hours of work: $B3/B4$

IV. Missing values, extreme values, unused questions in scoring absenteeism

The researcher has to make decisions about imputation of missing values based on other data available either in the larger survey or from administrative records. When ancillary data are not available to make imputations, cases should be deleted from analysis, ideally with a weight imposed on the remaining cases to restore representativeness of the sample.

In some cases, respondents tell us in response to B4 that their employer expects them to work “24 hours a day 7 days a week” or “more than is humanly possible” or etc. In cases of this sort, we make a rational imputation to the 90th percentile of the distribution of the acceptable responses to this question in the sub-sample of respondents in the organization under study of the same age, sex, and broad occupational type as the respondent who gave this extreme value.

In other cases, respondents tell us in response to B4 that their employer doesn’t care how many hours they work so long as they complete their work. In cases of this sort, we make a rational imputation that the actual hours the respondent reports working equals the number of hours the employer expects them to work.

B6 is consciously designed to be a difficult question to answer. The goal is to force respondents into thinking carefully, as it is difficult to answer this question without pausing and giving it some thought. Note that B5 is not used in calculating absenteeism. Instead, B5 is designed to help the respondent review the past four weeks before coming to the question that requires effort (B6). B6 is, in this regard, what survey methodologists call a “prequest” – a question designed to prime the respondent to give more accurate information than otherwise to the following question.

Comparison of responses to B5 and B6 can be used to detect respondents who give superficial responses to B6. The best way to do this is to compare $(4xB4 - 8xB5a - 8xB5b - 4xB5c - 4xB5d + 4xB5e)$ to B6. Large discrepancies are indicative of superficial responding and more detailed examination of case-by-case responses can be used to make rational decisions about case deletion.

V. Scoring presenteeism

Presenteeism is conceptualized as a measure of actual performance in relation to possible performance. Unlike absenteeism, a higher score indicates a lower amount of lost performance. Simple scoring, which is the only approach available in the absence of objective benchmark data, assumes that responses on the 0-10 response scales in questions B9-B11 indicate percent of performance. With this assumption made, **absolute presenteeism** has a lower bound of 0 (total lack of performance during time on the job) and an upper bound of 100 (no lack of performance during time on the job).

Relative presenteeism is a ratio of actual performance (B11) to the performance of most workers at the same job (B9, possible performance). We recommend restricting the distribution of relative presenteeism to the range of 0.25 to 2.0, where 0.25 is the worst relative performance (25% or less of other workers' performance) and 2.0 is the best performance (200% or more of other workers' performance). For example, if the respondent rates his own performance as 1, and the average worker's performance at 8, he is assigned a presenteeism score of 0.25 (1 divided by 8 is 0.125, restricted to the lower bound of 0.25). If another respondent rates his performance at 9, but rates the average worker's performance at 3, he gets the score of 2 (9 divided by 3 is 3, restricted to upper bound of 2).

In the rare cases where the performance of the average worker is rated zero, any non-zero rating of the respondent's own performance would be assigned a relative presenteeism score of 2.0 (e.g. 7 for own performance divided by 0 for other workers, theoretically an infinitely large number, would be recoded to the upper bound of 2). If the respondent rates both the average worker and his own performance as zero, a relative presenteeism score of 1.0 would be assigned. Note that the 0.25-2.0 recommended range is asymmetric by design due to the fact that objective ratings show that even the best workers seldom have more than twice the productivity of the average worker, while the worst workers often have productivity less than half the average and sometimes have productivity as low as one-fourth the average. The scoring rules are:

Absolute presenteeism scoring rule: $10 \times B11$

Relative presenteeism scoring rule: $B11/B9$ (restricted to the range of 0.25 to 2.0)

VI. Missing values, extreme values, unused questions in assessing presenteeism

The researcher has to make decisions about imputation of missing values based on other data available either in the larger survey or from administrative records. When ancillary data are not available to make imputations, cases should be deleted from analysis, ideally with a weight imposed on the remaining cases to restore representativeness of the sample.

In some cases, respondents tell us in response to B4 that their employer does not care how many hours they work so long as they complete their work. In cases of this sort, we make a rational imputation that the actual hours the respondent reports working equals the number of hours the employer expects them to work.

Note that B10 is not used in calculating presenteeism. Instead, B10 is designed to help the respondent focus response to B11 on the past four weeks by asking separately about earlier times. This approach is known in the

methodological literature as a synthetic bounded recall question -- a question designed to prime the respondent to give more accurate information than otherwise to the following question.

VII. Combining absenteeism and presenteeism

Relative absenteeism and *relative* presenteeism can be combined into one score. To do so, the measure of relative hours of work should be created first. It is equal to one minus relative absenteeism. For example, if the respondent reports expected hours of work per week to be 40 (i.e. 160 hours in 4 weeks) and actually worked 120 hours in the past 4 weeks, his relative absenteeism is $(40 \times 4 - 120) / 40 \times 4 = 0.25$. Then his relative hours of work are $1 - 0.25 = 0.75$ (or $120 / 40 \times 4$). If this respondent's measure of relative presenteeism is 1.6, the combined score is $0.75 \times 1.6 = 1.2$. Notice that this hypothetical respondent compensates lost hours with higher productivity as the total score is greater than 1.

Absolute presenteeism could be combined with relative hours worked to obtain a total score that discounts relative hours worked by the amount of lost performance. For example, if absolute presenteeism is 80%, the total score will be $0.8 \times 0.75 = 0.6$. This measure will always be below the total score obtained from the combination of relative presenteeism and relative absenteeism because absolute presenteeism cannot be higher than 100%.

Combining *absolute* absenteeism with absolute or relative presenteeism leads to difficulties. Let us assume that this combined score measures lost hours of work just like absolute absenteeism. One could argue that hours of work lost during 4 weeks should be discounted (or inflated) by the measure of performance (absolute or relative presenteeism). The rationale here is that if the person was especially productive during that period, the lost hours should count for more, and if the respondent was not productive, the lost hours should be discounted. In this case, the total score increases with productivity. On the other hand, as we have seen in the example above, respondents may compensate lost hours with higher performance. From this point of view, the total score combining absolute absenteeism and presenteeism should decrease with productivity. However, actual and expected working hours are necessary to compute the total score under these assumptions, which brings us back to relative absenteeism. In any case, given that the same two questions (about actual and expected working hours) are used to obtain both absolute and relative absenteeism, it is always advisable to use relative absenteeism when creating the combined score.

VIII. When to use absolute versus relative absenteeism and presenteeism measures

The decision to use absolute versus relative absenteeism and presenteeism measures should be made based on substantive considerations that vary from one research question to another. For example, we use relative absenteeism as the dependent variable in the regression equations to estimate the effects of health problems on work absence because we believe that the structural effects of health problems across workers who vary in FTE are best conceptualized as proportional rather than as absolute. For example, we believe that workers with half-time jobs are likely to miss only half as many hours because of the flu as workers with full-time jobs. However, we convert to absolute absenteeism when we carry out simulations that estimate the overall number of lost work days due to particular health problems because users of the HPQ reports find it more useful to know the absolute number of lost work days due to health problems than the ratio of such days. The situation can be

different for presenteeism, where absolute scores might make more substantive sense both as outcomes in regression analyses and in simulations.

IX. Limitations of not using the entire HPQ

Three limitations of not using the full HPQ have to be mentioned. First, the full HPQ includes additional questions that allow imputations and consistency checks to be made for questions that are most often left missing in the short absenteeism and presenteeism question series (e.g., B3, B12). Second, the full HPQ includes additional memory priming questions that improve the accuracy of report in the absenteeism and presenteeism questions (e.g., B8a-g). Third, use of the full HPQ makes it possible to have the data included in the HPQ master dataset, in which case more complex scoring rules can be used to code presenteeism than the simple scaling methods described here. These more complex rules use regression-based methods to calibrate scores on the presenteeism scales to objective measures obtained in several archival calibration studies. Nonetheless, despite these limitations, the short HPQ absenteeism and presenteeism question can be quite useful in providing a quick assessment of lost work performance in a workplace sample.