

# Methodological issues in assessing job stress and burnout in psychosocial research

PAOLO CAMPANINI

**KEY WORDS:** Job stress; burnout; methodological issues

## SUMMARY

*In recent years, researchers identified a “reproducibility crisis” of scientific studies. In assessing job stress and burnout in psychosocial research two biases that prevent the results from being generalized are common: sample bias (included web survey) and common method bias using questionnaires. These issues are commented and remedies are proposed to prevent or contain biases.*

Workers employed in several sectors may experience burnout, a condition specific for occupational settings, and therefore classified as an occupational disease (syndrome). WHO included such a condition in ICD-11 (1), providing the following definition: “Burn-out is a syndrome conceptualized as resulting from chronic workplace stress that has not been successfully managed. It is characterized by three dimensions: (i) feelings of energy depletion or exhaustion; (ii) increased mental distance from one’s job, or feelings of negativism or cynicism related to one’s job; and (iii) reduced professional efficacy. Burn-out refers specifically to phenomena in the occupational context and should not be applied to describe experiences in other areas of life.” Two out of three criteria are subjective, whereas the last one could be objectivated, though it will always be dependent on a subjective attitude to cope with the job. It is not surprising that the occurrence of such a syndrome is assessed relying on questionnaires. It is also not surprising that the increasing pressure of modern life makes stress at work a prominent issue in Occupational Medicine,

and this is why about 30 papers on this subject have been published in our journal, four of which in the last two years (2-5).

Studies on burnout and stress, or more simply psychosocial studies, almost exclusively use questionnaires distributed to the working population. This approach contains biases that can contribute to unexpected results.

Take, for example, the study “Burnout syndrome and its determinants among healthcare workers during the first wave of the Covid-19 outbreak in Italy: a cross-sectional study to identify sex-related differences” that reports counterintuitive results: as the exposure of job demands increases, there is a decrease in burnout in women. At the same time, resilience is a burnout risk factor in men (2). To explain these results, the authors advance explanations and describe limitations in the manuscript. Among others, two important limits are correctly reported: “the convenience sample” and “the ad hoc developed scales could lack in psychometric propriety”. These limitations are extremely widespread in psychosocial research and offer an opportunity to broaden the commen-

---

Editorial Commentary

Corresponding author: Paolo Campanini, La Medicina del Lavoro|Work, Environment and Health.

E.mail: redazione@lamedicinadellavoro.com

 open access [www.lamedicinadellavoro.it](http://www.lamedicinadellavoro.it)

tary to psychosocial research in general, including those recently published in our journal (2-5).

Science results from the critical assessment of the validity of studies, reports, and reviews. The most frequent warnings concern a small sample size, small effect sizes, sampling errors or convenience samples, imprecise data collection techniques and systematic errors which could increase significantly the probability that a published result is incorrect.

In Web-surveys (as in the study mentioned above, the survey was distributed via email and social network, with snowball sampling method), the selection of participants is a major factor limiting the generalizability of results and the sampling bias may be more likely. Using the web and social media to distribute the questionnaires does not allow to understand the target population and there is a risk of having very low response rates due to the high potential number of people reached. Social media in particular preclude non-members or those who do not follow them, and easily some members of the population are less likely to be included than others.

This 'self-selection bias' originates from the fact that people are more likely to respond to questionnaires for some reason, such as if they see items which interest them. This bias can also affect the observed associations between predictor and outcome. In the example study, we do not know if the respondents to the e-survey were mainly people enthusiastic about their role during the pandemic while most of those invited, with the same job demands but distressed, declined to respond; we have no idea how many people have been asked to fill out the questionnaire and in what context the survey was conducted.

At work, the context refers to the work organization and important psycho-social factors which also includes organizational culture and interpersonal relationship between colleagues and superiors. If the organizational reference is missing, it is impossible to develop inferences on the organization of work, thus excluding an interpretation to understanding the psychosocial factors at work.

The critical importance of carefully describing the method of the web-surveys and the sample who chose to respond has been evident at least since 2004, when Eysenbach started to elaborate a check list to

ensure the quality of reports in the medical literature (6), that it would be appropriate to use.

The second point of frequent bias is the tool used to collect data. Questionnaires are the most popular and frequently used technique for data collection in psychosocial sciences and related fields, owing to low cost, broad potential reach, and ease of administration (even via the internet), but both development and use of a valid and reliable questionnaire imply several steps needing considerable attention.

Using questionnaires is not without problems. The most frequent is the common method variance bias (i.e., the variance attributable to the measurement method rather than to the constructs the measures represent). Research has illustrated a variety of ways in which data obtained using questionnaires may be compromised in this way (7). Such bias must be carefully considered in interpreting research data.

When the predictor and the outcome are measured simultaneously in the same person, with the same method (the questionnaire), and without knowing the context of when the questionnaire is filled out (e.g. online survey), the bias becomes more evident.

The effect of the common method bias is still debated: examining the amount of common method variance present in measures across 70 psychosocial, marketing and educational studies, Cote and Buckley (8) found that approximately one-quarter of the variance might be due to systematic sources of measurement error like common method biases, whereas Lance et al. (9) conclude that "In contrast to conventional wisdom, common method effects do not appear to be so large as to pose a serious threat to organizational research, especially when the counteracting effects of measurement error are considered".

It is essential to consider that usually questionnaires measure predictor and outcome variables at the same time and same place and may share systematic covariation because this common measurement context may increase the likelihood that responses to measures of the predictor and outcome variables will co-exist in short-term memory, influencing the retrieval of information from long-term memory.

The questionnaires for psychosocial studies ask, on the one hand, to remember the frequency of

some events at work (in the past or recently), on the other hand, to indicate how the responding subject feels or the presence of any symptoms.

It is well known that the human memory system can be fallacious and may fail to remember the frequency of some event. The studies of Kahneman and Tversky in the early 1970s (10), that have been very influential on understanding of the mechanisms of human judgment, have focused on the judgmental error (cognitive bias) to reveal the heuristics and basic principles that govern human reasoning. The authors propose evidence that the evaluation of frequencies can lead to erroneous considerations and be influenced by the context (11).

The effects of the respondent's personality are also still debated. The main topic is positive and negative affectivity, defined as a mood-dispositional dimension that reflects pervasive individual differences in positive or negative emotionality and self-concept (12). For some authors, people who express high negative affectivity view themselves and various aspects of the world around them in generally negative terms; other authors found little support to this hypothesis (13).

As there is not yet strong evidence, such bias must be carefully considered in interpreting research data. Podsakoff et al. (7) suggest some procedural and statistical techniques for controlling biases, and the main remedy is to obtain measures of the predictor and outcome variables from different sources.

It would be better to use the questionnaire to assess the outcomes and other techniques (for example, observational) to define exposure. The above considerations lead one to assume that asking how people feel produces more accurate answers than asking the frequency of a past event. When it is impossible to obtain data from different sources, it is helpful to divide the exposure measures from the outcome measures temporally or with psychological tricks.

Questionnaires should be administered in contexts where it is possible to check for elements that can produce bias. The questionnaires themselves must be well validated (better with an external criterion) and have good scientific literature to document their validity.

Using these precautions is highly recommended because recent studies suggest reproducibility is

lower than is desirable (14-16). In a 2016 survey in *Nature*, 90% of respondents consider a "reproducibility crisis" (17). In particular, in psycho-social science empirical assessment of a random sample of articles published between 2014 and 2017 suggests a "serious neglect of transparency and reproducibility" (18), whereas in a random sample of 250 psychology articles, the transparency and reproducibility-related research practices were "far from routine" (19).

The reproducibility question can also arise from the over-interpretation of noise, facilitated by the extent to which data analysis is rapid, flexible and automated. The data collected during cross-sectional studies using ad hoc questionnaires that assess both exposure and outcome, which do not control the compilation context, should be interpreted with greater caution and carefully evaluating the effects of biases.

Using sources other than questionnaires, or questionnaire well validated by external criteria, to gather data assessing exposures can increase the validity of the results.

**CONFLICT OF INTEREST:** NO POTENTIAL CONFLICT OF INTEREST RELEVANT TO THIS ARTICLE WAS REPORTED BY THE AUTHOR

## REFERENCES

1. WHO. International Statistical Classification of Diseases and Related Health Problems (ICD). WHO, Geneva. <https://www.who.int/standards/classifications/classification-of-diseases> (last accessed 4.8.2021)
2. Brera AS, Arrigoni C, Dellafiore F, et al. Burnout syndrome and its determinants among healthcare workers during the first wave of the Covid-19 outbreak in Italy: a cross-sectional study to identify sex-related differences. *Med Lav.* 2021; 112 (4): 306-319
3. Converso D, Sottimano I, Balducci C. Violence exposure and burnout in healthcare sector: mediating role of work ability. *Med Lav.* 2021;112(1):58-67
4. Vignoli M, Mazzetti G, Converso D, Guglielmi D. How workers' emotional dissonance explains the association between customers' relations, burnout and health in an Italian supermarket chain. *Med Lav.* 2021;112(3): 200-208
5. Viotti S, Guglielmetti C, Gilardi S, Guidetti G. The role of colleague incivility in linking work-related stressors and job burnout. A cross-sectional study in a sample of faculty administrative employees. *Med Lav* 2021;112(3):209-218

6. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004;6(3):e34
7. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *J Applied Psychology*. 2003;88(5):879-903.
8. Cote JA, Buckley R. Estimating trait, method, and error variance: Generalizing across 70 construct validation studies. *J Marketing Research*. 1987;24:315-318.
9. Lance CE, Dawson B, Birkelbach D, Hoffman BJ. Method Effects, Measurement Error, and Substantive Conclusions. *Organizational Research Methods*. 2000;13:3:435-455
10. Kahneman D, Slovic P, Tversky A: *Judgment under uncertainty: Heuristics and biases*: Cambridge University Press, 1982
11. Kahneman D, Tversky A. On the Reality of Cognitive Illusions. *Psychological Review*. 1996;103(3):582-591
12. Watson D, Clark LA. Negative affectivity: The disposition to experience negative aversive emotional states. *Psychological Bulletin*. 1984;96:465-490
13. Spector PE, Chen PY, O'Connell BJ. A longitudinal study of relations between job stressors and job strains while controlling for prior negative affectivity and strains. *Journal of Applied Psychology*. 2000;85: 211-218
14. Wicherts JM, Borsboom D, Kats J, Molenaar D. The poor availability of psychological research data for reanalysis. *Am. Psychol*. 2006;61:726-728
15. Al-Shahi Salman R, Beller E, Kagan J, et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet*. 2014;383(9912): 176-185
16. Chan AW, Song F, Vickers A, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet*. 2014;383(9913):257-266. doi:10.1016/S0140-6736(13)62296-5
17. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016;533:452-454.
18. Hardwicke TE, Wallach JD, Kidwell MC, Bendixen T, Crüwell S, Ioannidis JPA. An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014-2017). *R Soc Open Sci*. 2020;7(2):190806
19. Hardwicke TE, Thibault RT, Kosie JE, Wallach JD, Kidwell MC, Ioannidis JPA. Estimating the Prevalence of Transparency and Reproducibility-Related Research Practices in Psychology (2014-2017). *Perspect Psychol Sci*. 2021;1745691620979806