

Measuring Health and Disease

Suggested answer: Classroom Exercise on Latent Variables and Factor Analysis

In this exercise, four papers were presented, all of which used factor analysis. The exercise was to read the paper and decide the following:

- what was the reason for doing factor analysis,
- how they did the authors decide the number of factors,
- how did they extract them,
- what did they find.

Paper 1: Preferences of patients for patient centred approach to consultation in primary care: observational study

This is the most straightforward of the papers. The authors say, at the end of their introduction ‘We report on a study of patient preferences for patient centredness in the context of an impending consultation with a doctor in primary care. We report the principal domains of patient centredness from the patients’ perspective and compare preferences for these components with preferences for other more conventional “biomedical” aspects of the consultation such as providing a prescription or physical examination.’ First, they attempt to investigate the components of patient preference. They then look at the relationships of these with other variables.

The reason for doing factor analysis was to get a description of patient preferences. Hence, this was an exploratory factor analysis, without preconceived ideas about the structure. Indeed, they mention a five domain model as their inspiration, but get only three factors.

The authors do not say how they decided the number of factors. They say only that ‘Factor analysis suggested a three component solution, which explained 91% of the variance’. My guess is that this is based on the ‘Eigenvalues greater than 1.0’ principle, as this is completely automatic and often built into software.

The method used to extract the factors is stated to be varimax rotation, the simplest method.

They found three factors, which they labelled ‘Factor 1 communication: illness experience, communication, and doctor-patient relationship’, ‘Factor 2 partnership: interest in beliefs, expectations, and negotiating common ground’, and ‘Factor 3: health promotion’. There were also several variables which did not contribute much to any factor and were treated separately.

Paper 2: Psychometric performance of an assessment scale for strain in nursing care: The M-NCAS

This study is clearly stated to be a post-hoc assessment of reliability and validity. They have added 11 items to an established 21 item questionnaire. Each of these 32 items had two aspects, strain and attitude, which appear to have been analysed separately.

The reason for doing factor analysis was to evaluate the underlying subscale structure of the M-NCAS.

To extract the factors, they started with a completely exploratory analysis and say that 'with the number of factors unspecified, no clear factor pattern was discernible'. They then imposed a six-factor solution for both dimensions (strain and attitude), based on the suitability of six factors for the original instrument with 21 items. They say that 'examination of the oblique rotated factor loadings demonstrated substantial item overlap. Results of subsequent factor analyses indicated that the best solution was a three-factor solution for the attitude domain, and a five-factor solution for the strain domain'. So really we have no idea how they decided on the number of factors. It looks dangerously subjective.

They decided that there were three factors for the attitude domain, which they labelled 'Attention Seeking', 'Autonomy' and 'Difficulty'. They decided that there were five factors for the strain domain, which they labelled 'Affect', 'Job Satisfaction', 'Neediness', 'Predictability', and 'Self Direction'.

Paper 3: Cholesterol control, medication adherence and illness cognition

The authors state the purpose of the study as being to examine the role of psychological variables suggested by the self-regulation model in predicting LDL cholesterol control and cholesterol-lowering medication compliance. They carried out a survey with 10 questions which they wanted to use as predictors in regression. They reported that several of the variables created by averaging responses to survey items were moderately correlated. To eliminate the problem of colinearity, they conducted an exploratory factor analysis of the 10 illness cognition items used in the study.

They used principal components analysis with a varimax rotation to extract the factors, what we might call bog standard.

They say 'After examining the factor loadings and the eigenvalues, we chose a 5-factor solution for theoretical reasons based on the selfregulation model'. This is another example where we not know exactly how they did it. Don't the referees read these papers? Based on eigenvalues, they would have stopped at 4.

They labelled their five factors 'Symptoms', 'Consequences', 'Timeline', 'Cause', and 'Cure'. They were then able to look at the effect of these factors on measured cholesterol using regression.