Measurement of physical activity using a questionnaire

Measuring the physical activity by a questionnaire may have several uses in the PAC project. A physical activity questionnaire provides an estimate of the physical activity level of a person before any physiological data have been gathered. Physiological data can also be completed by questionnaire data that give a broader view of all types of physical activity: not only for sports, but also for other leisure time activities, travel, work etcetera.

Feedback of the PAC must be adjusted to the level of physical activity that a person shows. So, an accurate estimate of physical activity helps to optimize the coaching feedback. A questionnaire also gives the opportunity to measure physical activity at the very beginning of the trajectory, before any coaching has taken place. This provides a useful 'reference' for later comparison, or even a basis for product advice.

The use of Computerized Adaptive Testing

When measuring physical activity using a questionnaire, we want to achieve a measurement that is as accurate and precise as possible. Most traditional paper and pencil questionnaires try to do this by asking large sets of questions that cover every bit of physical activity as much as possible. There is always a trade-off between accuracy and speed: large questionnaires will take quite some time to complete. This is something we want to avoid in the PAC. We don't want the users to spend time sitting and filling out questionnaires, we want to encourage them to exercise. Therefore, it is important to keep the number of questions to measure physical activity limited.

Computerized Adaptive Testing (CAT) is a technique that tries to combine speed and accuracy by asking only the most relevant questions for each individual. This results in a relatively short test while maintaining an acceptable accuracy.

Another important issue regarding questionnaire measurement is that people often tend to get bored while filling in a questionnaire. A smaller number of questions is already a step into the right direction, but CAT does even more. The questions that are selected by the CAT algorithm are of maximum relevance to measure the individual. These relevant questions will also be more interesting for the respondent, because each question is selected to ask for new information that is not implicitly contained in previous questions. Therefore, the respondent will not have the idea that the same information is asked for over and over again, and will not get bored

When questionnaire measurement of physical activity is repeated several times during the PAC trajectory, the CAT algorithm may select a different set of questions each time the test is taken. This is a further advantage that helps to create a more dynamic measurement and thereby prevents the boredom that questionnaire measurement is often associated with.

The principle behind CAT

How does CAT select the most relevant questions? This can be explained by a simple example. If you want to measure the math-skills of a mathematician, you don't ask a simple question like "what is 2 + 1". This question would give you almost no information; you already know that the mathematician will be able to answer this question. If you want to measure the math-skills of a 5 year old child on the other hand, the question "what is 2 + 1" would be very good: you really don't know in advance whether the child will be able to give the correct answer. The response gives you information that you don't know yet.
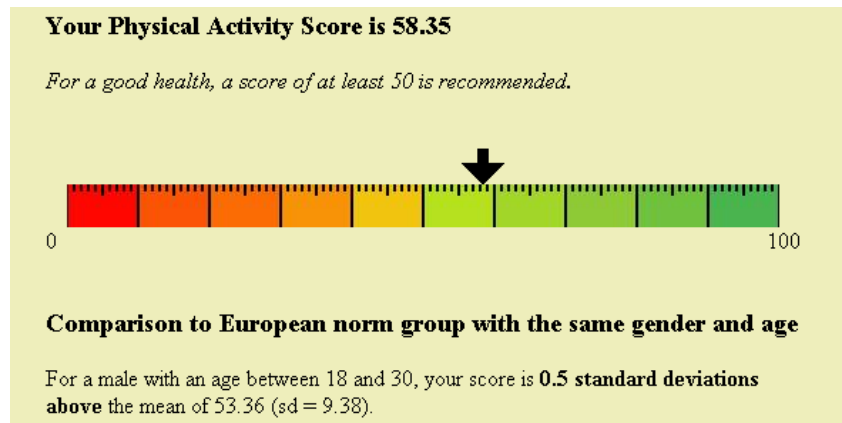
The principle of adjusting questions to the respondent also applies to the measurement of physical activity. You don't want to ask a professional sportsman "do you exercise at least once a month", because this question will give you no new information. When a respondent takes a physical activity questionnaire, we don't know in advance whether he is a professional sportsman or someone who just occasionally exercises after work. So, how are we going to select the right question? After the first randomly selected question, the CAT technique makes a provisional estimate of physical activity after every response. This estimate indicates what the physical activity level of the respondent approximately is, and can be used to find a question that maximally matches this activity level.

The provisional estimate is not a point-estimate, but rather a probability distribution that tells in what range the physical activity score is. This distribution will narrow after every response that a person gives. With other words: new information (the response) makes the estimate more precise. The above three figures give the estimated distributions after the first three questions in a CAT for physical activity. It can be seen that the distributions narrow after every response.
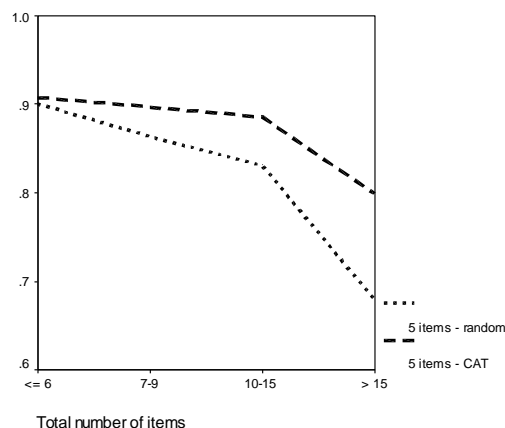
Building a CAT for physical activity

In order to use CAT in practice, we need a substantive set of questions that measure the same subject. The difficulty of each of the questions must be known in advance. The current CAT for physical activity has been based on data that were gathered during the EUPASS project (Rutten et al, 2003). This project gathered data on a large set of internationally used physical activity questions together with the International Physical Activity Questionnaire (IPAQ, www.ipaq.ki.se). The difficulty of each of the questions has been estimated through Response Conversion (Van Buuren and Tennant, 2004).

Your Physical Activity Score is 58.35

For a good health, a score of at least 50 is recommended.

0   100

Comparison to European norm group with the same gender and age

For a male with an age between 18 and 30, your score is **0.5 standard deviations above** the mean of 53.36 (sd = 9.38).

The CAT demo that was developed during phase 1 of the PAC project adaptively selects 8 questions out of 33 and computes a physical activity score that is based on the responses. The physical activity score is compared to a gender- and age specific norm score.

CAT simulation study

A simulation study has been carried out to study the advantage of using CAT to select physical activity questions. Respondents in the EUPASS dataset all responded to a set of up to 20 physical activity questions.

In our simulation study we looked at the correlation between the total score on all items and a score on a subset of 5 items. The 5 items were selected using two different methods. Once with random selection and once using CAT. The above Figure shows that CAT selection results in better correlations with the total score. This means that CAT achieves a better measurement precision with the same number of questions, by choosing the most relevant questions. When the total score is based on 16 to 20 questions, the correlation of a CAT subset of 5 is 0.80, while the correlation of a random subset of 5 is only 0.68.

References

Rutten A, Vuillemin A, Ooijendijk WT, Schena F, Sjostrom M, Stahl T, Vanden Auweele Y, Welshman J, Ziemainz H (2003). Physical activity monitoring in Europe. The European Physical Activity Surveillance System (EUPASS) approach and indicator testing. Public Health Nutr. 2003 Jun; 6(4):377-84.

Van Buuren S, Tennant A, eds. (2004). Response conversion for the Health Monitoring Program. Leiden: TNO Prevention and Health, Publ. nr. 04.145, ISBN 90-5986-082-9.