Assessing Stage of Change for physical activity: how congruent are parallel methods?

J. Marttila and R. Nupponen

Abstract

The single-question self-classification Stages of Change scales (SAS) for two modes of physical activity were compared with parallel staging methods. In Study 1, the participants (#N = 50) completed SAS in a questionnaire and were then personally interviewed on their physical activity. In four fifths of the cases, SAS indicated the same stage as the interviewer’s judgment. In Study 2, a representative survey sample (#N = 600) completed both SAS and, in another questionnaire, a three-question algorithm staging instrument (TSQ) constructed for the same target behaviors. About 50% of all participants were placed in the same stage with both instruments. The compatibility rate rose to 80% when the number of stages was reduced from the original eight to five. However, TSQ also accumulated a higher share of cases in the stages with regular action. In both studies, the most obvious sources for incompatible staging were the regularity and time frame of the targeted behavior. Thus, neither SAS nor TSQ is on its own a sufficiently accurate instrument for use in personalized stage-based interventions. TSQ shows no obvious advantages over SAS. In counseling, SAS seems useful in combination with a personal interview.

Introduction: approaches to the measurement of stage

A few psychological models of health behavior address the temporal dimension of behavior change. In the Stages of Change model a distinction is most typically made between four or five stages of the adoption of a certain behavior: Precontemplation with no consideration of initiation of the behavior concerned; Contemplation with intention or serious consideration, but without any behavior; Preparation or Decision with strong intention and irregular or tentative performance of the behavior; Action with recent initiation of regular full scale behavior; and Maintenance, i.e. the establishment of permanent behavior. The ultimate stage of Termination with no temptation to return to past behavior is sometimes included as well (Prochaska and Velicer, 1997).

The growing popularity of the Stages of Change model has created increasing demand for feasible, non-expensive staging tools, but no standardized instruments are available. The need is particularly obvious in health behavior interventions where repeated measurements are necessary to verify the impact.

Several approaches can be discerned in the development of the methods. One of these is based on structured interviews carried out either face-to-face in a clinical setting, or over the phone for research or intervention purposes. The method thus consists of the interview questions and, less regularly, the guidelines for the staging decision. The judgment on the person’s proper stage is left to
the interviewer. The properties of these methods are not sufficiently known (Carey et al., 1999).

Another approach presumes a multi-dimensional continuum underlying the stages. The constitutive dimensions are operationalized through a series of questions. It is assumed that each stage shows a special constellation of dimensions. Accordingly, the person will receive one or more scores for each of the stages or a placement based on a cluster score (Kerns et al., 1997; Carey et al., 1999; Kerns and Rosenberg, 2000). Although good psychometric properties can be achieved, the number of items, the scoring procedures needed and the use of several scores or profiles mean that these scales are less suitable in population-based interventions.

Short questionnaire inventories consisting of several separate questions are also commonly used (e.g. (Lechner and DeVries, 1995; Ommundsen and Aaro, 1995; Nguyen et al., 1997; Povey et al., 1999)) as well as scales with marked stages, so called ‘ladders’ [for exercise, e.g. (Marcus and Owen, 1992)]. A more advanced methodological approach has also been gaining ground, i.e. that of staging algorithms. Here, the route from a certain response to the stage indicated by this response is built into the question or questions and the response scales. In a single-question algorithm, each response option indicates one definite stage and the respondent has to classify him/herself [e.g. (Courneya, 1995; Haire-Joshu et al., 1999)]. The options are differentiated from one another by several critical components: awareness of the target behavior, actual behavior, time frame of the behavior and intentions to change the behavior. When the algorithm consists of several questions, they are aimed at tracking these components (Glanz et al., 1994; Reed et al., 1997) and response scales with true–false dichotomies [e.g. (Curry et al., 1992; Marcus and Simkin, 1993)] or Likert scales [e.g. (Marcus et al., 1992)] are used.

Self-assessment algorithms are rather easy to use in questionnaire surveys with normal populations. However, the evaluation of these measures is complicated by inconsistencies in uses and only limited information is available on their adequacy (Reed et al., 1997; Carey et al., 1999). In the case of single-question self-assessment algorithms for Stages of Change of exercise, only three studies have reported test–retest reliability values (Marcus et al., 1992; Courneya, 1995; Donovan et al., 1998). We found no studies comparing single-question and multi-question self-assessment algorithm scales in the area of physical activity and exercise. Further, the validity of staging algorithms is a complicated issue as no external criterion covers the total array of stages.

The basic demand for accurate staging is that the target behavior is explicit and discrete enough to allow for an unequivocal definition. In order to define non-overlapping stages, we need to have clearly stated criteria for the behavior. In the case of exercise, Reed et al. (Reed et al., 1997) emphasize three criteria: frequency, time frame (i.e. duration per interval) and intensity of the behavior in question. In addition, all this has to be expressed in a lucid, understandable wording to avoid any confusion.

Comparing staging measures for physical activity

Aim of the studies

We carried out two studies on the single-question self-classification algorithm scales developed in a project concerned with health-related physical activity (Miilunpalo et al., 2000). Our aim was to find out (1) whether the respondents were confused by the definition given of the target behavior or by the multi-component structure of the stage–response alternatives in the scales, (2) and we wanted to explore the multi-component structure of these single-question scales by means of a parallel questionnaire measure based on a multi-question staging algorithm.

The two studies employed the same definitions of the target behaviors and the same one-question staging scales, but the other methodological solutions were different. For each study, a different sample was selected from the population of the Region of Tampere, Southern Finland, including the city itself. The following describes the partici-
pants, methods and implementation as well as the results obtained first for Study 1, then for Study 2.

**Target behaviors**

The target behaviors were selected on the basis of (1) our theoretical analysis of physical activity as a context-bound psychological phenomenon (Marttila *et al.*, 1998), (2) the health-enhancing potential of the behaviors (American College of Sports Medicine, 1998) and (3) the practical demands of the behaviors in question, such as minor requirements of skills as well as social acceptance.

We defined two specific modes of physical activity, *outdoor aerobic exercise* (OAE) and *everyday commuting activity* (ECA). Both include a variety of simple, easy to perform behaviors. OAE comprises walking, jogging, running, bicycling and cross-country skiing for the purposes of health, fitness or recreation, causing at least slight shortness of breath and sweating, and lasting at least 20 min. ECA covers behaviors such as going to work, going shopping, taking a child to day care or to school, or running other errands done wholly or partially on foot or by bicycle lasting at least 10 min each time; for walking this means at least a distance of 1 km; for cycling 1.5–3 km (Miilunpalo *et al.*, 2000). Both OAE and ECA are commonly practiced among Finnish adults, and both are certainly familiar to all Finns regardless of personal practice. In the Finnish language both activities can be expressed with one simple word: *lenkkeily* for OAE and *asiointiliikunta* for ECA.

**Self-assessment scales (SAS) of Stages of Change**

Each person’s OAE and ECA stages were measured separately. In the eight-item SAS the respondents were asked to mark the stage that most closely corresponded with their current status. The stages were: Negative Precontemplation (the behavior is out of question); Neutral Precontemplation (no current behavior and no consideration to begin); Contemplation (no current behavior but considering to adopt regular behavior within the next 6 months); Ongoing Occasional Activity (subcriterion behavior without intention to change) and Preparation to Change (subcriterion behavior and intention to make it more regular within the next month); and Action, Short- and Long-Term Maintenance (regular behavior meeting the criterion during the past 6 months, between 6 months and 2 years, and for more than 2 years, respectively). The definition of the target behaviors was given at the top of the page above the scale, first for OAE and then, on the next page, for ECA. Comparisons of subpooled responses in the original population survey indicated good consistency of SAS for ECA and moderate for OAE. The development of the scales is described and their translations are given in Miilunpalo *et al.* (Miilunpalo *et al.*, 2000).

**Study 1**

**Purpose**

The congruence of the responses to the SAS for OAE and ECA and staging decisions based on information gathered in semi-standardized personal interviews was examined in a middle-aged convenience sample.

**Sample and implementation**

In total, 83 individuals aged 41–43 years were invited to participate in an interview on perceptions of the determinants of OAE and ECA. Fifty of these (60%), 23 women and 27 men, agreed to take part. This number was considered sufficient for the purposes of the interview and manageable considering the demands of the semi-structured interview method. The selection of the participants is described in detail in Marttila and Nupponen (Marttila and Nupponen, 2000).

All the participants had been involved in a population survey [reported in (Miilunpalo *et al.*, 2000)] 9 months earlier. The response rate in the age group 41–43 years was 62.3% for women and 55.0% for men. The persons eligible for interviewing had reported that their current health was not poor, that they had no illness, handicap or impairment which prevented them from taking exercise.
and that their job did not require them to be away from their home or to travel a lot. Persons both practicing and not practicing the target behaviors were included.

Method
With their personal invitations the participants received a short questionnaire covering their current physical activity, the SAS scales and selected demographic facts. The interview covered preset themes in a normative order (Marttila and Nupponen, 2000). The opening theme was the general timetable and context of everyday living, with special emphasis on physical activity. OAE and ECA were then dealt with separately in closer detail. For both behaviors the interviewee’s motivational and behavioral status was checked first. As part of this theme the interviewer and the participant reviewed the SAS responses, and the interviewee was asked whether the response was congruent with his or her current situation. In the case of less than obvious congruence, the interviewer presented further questions on the frequency and definition of the targeted behaviors, the regularity of the behavior and the duration of regular behavior. The final judgment on congruence was based on the whole transcribed interview text because the interviewees sometimes revealed relevant information while discussing other themes. The procedure can be considered analogous to clinical staging methods.

Results
The interviewees had no difficulty relating themselves to the target behaviors. They spontaneously portrayed OAE as exercise, whereas they considered ECA as an integral part of everyday life. However, ECA was identified as a physical activity when this was being discussed (Marttila and Nupponen, 2000).

In OAE, there were seven mismatches between the SAS response and the interview information. In ECA, the number of mismatches was eight. For both behaviors, less than one-fifth of the self-rated stages were judged to be erroneous. The chance-corrected $k$ coefficients (Cohen, 1960) of agreement between SAS and the staging based on the interview information was about 0.80. Approximately one-third of all the mismatches were identified by the interviewees, without the interviewer drawing their attention to the discrepancy.

Most mismatches for OAE emerged between occasional and regular behavior (four cases). The main difficulty seemed to be providing a simultaneous estimate of the frequency and the time frame of continuous regular behavior (less than/more than 6 months/2 years) postulated by SAS. For ECA, the dominant cause for mismatch was related to the regularity of the behavior and caused by the exclusion of some behaviors covered by the definition of the activity. The difficulty thus lay in the recognition of relevant activities.

Study 2

Purpose
The compatibility of the results obtained from the same persons with two different staging instruments was studied in a panel survey. For both OAE and ECA, the respondents assessed their Stage of Change by means of SAS. In another questionnaire (TSQ), they responded to three separate multiple-choice questions construed for staging; by means of a special staging algorithm, the researcher placed each respondent in one of the eight stages on the basis of his or her responses.

Sample
A representative sample of 300 men and 300 women was drawn from residents aged 51–53 years in the same Region of Tampere as in the population study where the SAS scales were developed (Miilunpalo et al., 2000); the same sampling methods were used.

Questionnaires
The topics to be surveyed comprised social background data and experience of subjective health and well-being as well as leisure activities, including physical activity. Half of the questions
on each topic were presented in Questionnaire A, the other half in Questionnaire B.

In a counter-balanced AB–BA design, half of the men and half of the women in the sample were first posted Questionnaire A, whereas the other half first received Questionnaire B. The recipients were asked to respond within 14 days and indicate whether they would be prepared to complete the other questionnaire as well. If so, it was immediately mailed to them.

### Stages of Change measures

The one-question SAS for OAE and ECA were presented in Questionnaire A. An alternate questionnaire measure with three separate multiple choice questions was specially planned for use in this study. This measure, indicated as TSQ, was included in Questionnaire B; the same definitions were given of the two target behaviors as for SAS.

The three TSQ questions (Q1, Q2 and Q3) covered the components discriminating between stages: frequency of current behavior (Q1), time frame (i.e. duration of regular practice; Q2) and intention to change current behavior (Q3). The questions and the response options are presented in Tables I and II. A full range of alternatives was given for each question. However, the last response option (‘5. Regular behavior for more than 2 years’) was omitted by mistake from Q2 for both OAE and ECA. Several respondents had themselves added the missing alternative or modified option 4 to indicate their answer. In Tables I and II these responses are added to the frequency of response option 4.

By means of the algorithm, i.e. the rule specifying the stage indicated by each response, the researchers first checked the stage of the respondent separately for Q1, Q2 and Q3. Based on these, the respondents’ final placement was decided. In those cases where the responses to Q1, Q2 and Q3 did not unanimously indicate the same stage, the respondent could not be placed in any stage with TSQ.

### Analyses and results

The questionnaires that were returned with at least one staging question answered in both were included in the study material. The response rate

---

**Table I. Distributions of TSQ responses for OAE by Stage of Change**

<table>
<thead>
<tr>
<th>TSQ question for OAE</th>
<th>Stage of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Q1. Do you take OAE? n = 390</strong></td>
<td></td>
</tr>
<tr>
<td>1 No, and it would be out of the question in my case.</td>
<td>14 3 0 4 0 0 0 0</td>
</tr>
<tr>
<td>2 No, and I have never really thought about it.</td>
<td>3 16 8 4 0 0 0 0</td>
</tr>
<tr>
<td>3 No, but I might start regular OAE within the next 6 months</td>
<td>0 4 14 10 5 0 1 0</td>
</tr>
<tr>
<td>4 Yes, I occasionally take OAE</td>
<td>3 5 7 93 24 6 5 16</td>
</tr>
<tr>
<td>5 Yes, I take OAE regularly __ times a week</td>
<td>0 1 2 10 3 8 11 110</td>
</tr>
</tbody>
</table>

| **Q2. How long have you been taking OAE? n = 359** |
| 1 I have not taken OAE | 17 23 21 13 1 0 1 0 |
| 2 I have taken OAE less frequently than twice a week | 0 4 7 66 20 1 3 4 |
| 3 I have taken OAE at least twice a week, but only during the past 6 months | 1 1 0 10 6 4 2 7 |
| 4 I have taken OAE at least twice a week for more than 6 months but less than 2 years | 1 2 3 28 4 9 11 89 |

| **Q3 Do you intend to increase or decrease your OAE within the next month? n = 382** |
| 1 I’m intending to start OAE | 0 5 5 7 3 0 0 0 |
| 2 I’m intending to increase OAE | 1 0 12 23 22 5 7 22 |
| 3 I’m not intending to make any changes to my OAE | 17 21 13 88 5 9 10 102 |
| 4 I’m intending to decrease my OAE | 0 0 0 1 0 0 0 3 |
| 5 I’m intending to quit OAE | 1 0 0 0 0 0 0 0 |

Number of responses. Frequencies of responses meeting the definition of the stage in italics. Total number of respondents N = 395.
for both questionnaires was 66% (N = 395), 73% among women and 59% among men; the total sample response rate for the first questionnaire was 71%. The distributions of background variables, such as education and self-perceived health, as well as the rates of responding to SAS and TSQ questions were independent of the order (AB or BA) in which the respondent had received the two questionnaires. The response rates were 97–99% in SAS as well as in Q1 and Q3, but 90 and 92% in the case of Q2 for OAE and ECA.

The SAS distributions obtained were compared with the distributions in the population survey where SAS was developed (Miilunpalo et al., 2000). The proportions of cases in each of the eight stages were between the figures recorded for respondents aged 41–43 and 55–57 years in that study. No statistical differences with these distributions were found either with the total group or in women or men.

To get an overview of response overlap, SAS responses were cross-tabulated with questions Q1–Q3 (Tables I and II). For OAE, the percentages of corresponding responses for SAS and Q1 and Q2 were 74 and 70%, respectively. Upon closer scrutiny, a higher rate was found in Q1 (78%) than in Q2 (71%) for OAE in women, while almost equal rates (66 and 68%) were observed in men. The highest correspondence rates, 80%, were found with Q3 (intention to change) where the most alternatives are compatible with most SAS stages. For ECA, the percentages increased from Q1 (65%) and Q2 (70%) to Q3 (93%) in the total group, and also in women as well as in men.

With SAS, missing cases were those with missing response. With TSQ, cases with missing responses to any of the three questions as well as those with responses indicating discongruent stages could not be placed in any stage. The proportion of respondents placed in any stage is thus lower when staging was based on TSQ (OAE: 87.6%; ECA: 88.1%) than on SAS (OAE: 99.7%; ECA: 99.0%).

The stage distributions obtained with the two instruments are presented in Table III. The $\chi^2$ one-sample tests (goodness-of-fit) confirmed the dif-

---

**Table II. Distributions of TSQ responses for ECA by Stage of Change**

<table>
<thead>
<tr>
<th>TSQ question for ECA</th>
<th>Stage of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Do you commute on foot or by bicycle? n = 379*</td>
<td></td>
</tr>
<tr>
<td>1 No, and it would be out of the question in my case</td>
<td>43</td>
</tr>
<tr>
<td>2 No, and I have never really thought about it</td>
<td>16</td>
</tr>
<tr>
<td>3 No, but I might start within the next 6 months</td>
<td>3</td>
</tr>
<tr>
<td>4 Yes, I occasionally commute</td>
<td>6</td>
</tr>
<tr>
<td>5 Yes, I commute regularly ___ times a week</td>
<td>1</td>
</tr>
<tr>
<td>Q2. How long have you been commuting mainly on foot or by bicycle? n = 362*</td>
<td></td>
</tr>
<tr>
<td>1 I do not commute on foot or by bicycle</td>
<td>62</td>
</tr>
<tr>
<td>2 Less frequently than twice a week</td>
<td>6</td>
</tr>
<tr>
<td>3 At least twice a week, but only during the past 6 months</td>
<td>0</td>
</tr>
<tr>
<td>4 At least twice a week for more than 6 months but less than 2 years</td>
<td>2</td>
</tr>
<tr>
<td>Q3. Do you intend to change your ECA within the next month? n = 381*</td>
<td></td>
</tr>
<tr>
<td>1 I’m intending to start ECA</td>
<td>0</td>
</tr>
<tr>
<td>2 I’m intending to increase ECA</td>
<td>2</td>
</tr>
<tr>
<td>3 I’m not intending to make any changes to my ECA</td>
<td>65</td>
</tr>
<tr>
<td>4 I’m intending to decrease my ECA</td>
<td>0</td>
</tr>
<tr>
<td>5 I’m intending to quit ECA</td>
<td>0</td>
</tr>
</tbody>
</table>

*Number of responses. Frequencies of responses meeting the definition of the stage in italics. Total number of respondents N = 395.
ference of the distributions in the total group of respondents (cases without stage placement as the ninth category; $\chi^2 = 198.40$ for OAE and 91.11 for ECA, both $P < 0.001$). The main differences between the distributions are seen in the proportion of cases at Ongoing Occasional Activity and Short-Term Maintenance. Further, the three stages with regular activity (Action, and Short- and Long-Term Maintenance) together cover a considerably lower proportion of cases with SAS than with TSQ. The combined rates of these stages are 40 and 54%, respectively, for OAE, and 36 and 49% for ECA.

The overall compatibility of stage decisions based on SAS and TSQ was low: about half of the assessed cases were placed in the same stage. The proportion was only slightly higher when the two Precontemplation stages were considered as one stage (OAE: 55%, $k = 0.18$; ECA: 60%, $k = 0.33$). A satisfying correspondence (OAE: 75%, $k = 0.67$; ECA: 79%, $k = 0.73$) was hardly achieved even by further combining all three stages with regular activity (Action and the two Maintenance stages) and thus reducing the stages to five.

The highest compatibility of stage decisions, i.e. the rates of 75 and 79% correspond to 66 (OAE) and 70% (ECA) of the total of 395 respondents. In other words, for OAE, 34%, and for ECA, 30% of the respondents were placed in different stages through SAS and TSQ or not placed in any stage. Assuming that the lower response rate to Q2 caused the non-placement of about 5% of the cases, this corresponds to 30 (OAE)–25% (ECA) of the total number of respondents.

### Discussion

#### Definitions, wording and scale composition

The single-question self-classification scale SAS meets the requirements set by Reed *et al.* (Reed *et al.*, 1997) for a single-question algorithm scale. The wording and scale composition of SAS were explored in Study 1.

No obvious misunderstandings of the definitions or wordings of the SAS scales were found either among the interviewees practicing OAE or ECA, or among those not. Thus, the risk of measuring different objects at the first stages and the last stages of the scale may be smaller with SAS than with a questionnaire staging scale for ‘exercise’ or ‘physical activity’ (Donovan *et al.*, 1998). As defined in our studies, both OAE and ECA cover a more coherent set of behaviors than the broad concepts of ‘exercise’ and ‘physical activity’. However, several interviewees omitted behaviors covered by the definition of ECA when they completed the SAS. This observation is easy to understand if we bear in mind that ECA behaviors are closely related to or interwoven in daily routines and not readily perceived as physical activity (Marttila and Nupponen, 2000).
The results of our studies suggest that the inherent multi-dimensionality of the concept of stage, when expressed in a multi-component questionnaire scale, presents an obvious threat to the accuracy of self-assessment. In SAS as well as in other self-classification scales, the component criteria needed for staging are built into the response options. The task facing the respondents is quite complex: they will have to compare several alternatives which each include two simultaneous criteria. TSQ is simpler than SAS, but feedback from the respondents indicated that they did not prefer TSQ to SAS. Q1 and Q3 each focus on one criterion only, but for Q2 both time frame and frequency have to be observed. This (in addition to the omission of one option by mistake) may have contributed to the somewhat lower response rate for Q2.

Stage distributions
The results of studies concerned with exercise (Reed et al., 1997) and dietary behaviors [e.g. (Povey et al., 1999)] suggest that with a more specific behavior, a larger proportion of respondents are found in the stages without target activity and, accordingly, a smaller proportion in stages with action. However, in four representative population samples from young Finnish adults to the middle-aged, SAS stage distributions for OAE and ECA have revealed a marked accumulation of cases in Maintenance stages [Study 2 (Miilunpalo et al., 2000)].

In Study 2, the definitions of the target behaviors and the general instructions were the same for SAS and for TSQ. However, a greater proportion of cases was accumulated in Long-Term Maintenance with TSQ than with SAS. Further, as the option omitted in Q2 corresponds to Long-Term Maintenance, the possibility remains that the proportion of cases could be even higher in Long-Term Maintenance and lower in Short-Term Maintenance.

The components of regularity and time frame are implied in these differences of distributions. Regularity is again met in the lower proportion of cases in Preparation with both staging instruments and also for Ongoing Occasional Action, the stage emphasized with SAS. The role of intention is problematic. Q3 indicating intention yielded a higher percentage of overlapping responses with SAS than Q1 and Q2. It is precisely on account of Q3 that TSQ allows us to cover a wider range of intentions than SAS does. This improves the accuracy of the staging decisions for Preparation, but it may also increase discrepant information in the respondent’s responses to the three questions and, by the same token, the number of cases not placed in any stage with the TSQ algorithm.

Compatibility of stage decisions
In Study 1, all participants were staged both on the basis of SAS and the interviewer’s judgment. No more than one-fifth of the cases were placed into a different stage. For OAE, regularity and time frame were the main causes of mismatches between SAS and the interviewer’s judgments; for ECA, the behaviors omitted by the interviewee contributed to the mismatches.

In Study 2, for both target behaviors, a considerably smaller proportion of the respondents could be placed in any stage on the basis of TSQ (about 88%) than of SAS (about 99%). The stage decisions showed marked incompatibility: Only about half of the classified cases were placed in the corresponding stages by SAS and TSQ, which is about the rate of agreement obtained by chance. The total rate of cases placed in discrepant stages or in no stage was 30–35% even when the number of stages was reduced to five; the rate corresponds to about one-quarter of the number of the respondents even when the lower response rate to Q2 is taken into account.

The theoretically important component of intention deserves our attention. For OAE, SAS and TSQ gathered a different share of respondents in the stage of Preparation with intention and in the stage of Ongoing Occasional Activity without intention to initiate regular activity (Table III); at both stages the number of mismatches equals one-quarter of the responses to Q3 (Table I). For ECA, the proportions of respondents placed in
Preparation show no marked difference between SAS and TSQ.

**Conclusions**

One promising idea of staging is that persons in different phases of behavior change benefit from qualitatively different interventions (Prochaska and Velicer, 1997). Measurements of the actual stage aim to help construe stage-matched interventions to make an intervention more effective.

In interventions aimed at promoting health-enhancing physical activity, the staging instruments need to be targeted to specific, psychologically coherent categories of activity. The results from research on dietary behaviors (Haire-Joshu et al., 1999; Povey et al., 1999) as well as the experience with ECA in Study 1 show that even the identification of the behavior concerned cannot be taken for granted when questionnaire instruments are used. Self-assessment of perceptually more distinctive behaviors such as smoking or condom use may be a less difficult task.

Our results suggest that neither SAS nor TSQ is accurate enough to be used independently as a basis for personalized interventions in physical activity. Since it is probably easier for the respondent to perceive the aim of SAS than that of TSQ, we are inclined to prefer SAS in counseling contexts. It is advisable to invest additional time in a personal structured interview covering themes from general physical activity to the details of specified behaviors, especially routine ones.

In surveys, TSQ shows no obvious advantages over SAS. The respondents did not prefer TSQ to SAS. Further, the number of respondents not staged at all is bound to be higher with TSQ: with the three questions instead of one, it is prone to missing and contradictory responses. Also, the use of three questions may enhance the effect of social desirability and thus artificially accumulate respondents in stages of Maintenance.

The reliability of staging depends, of course, on the number of stages applied. In population-based interventions in physical activity, the distinctions between the two stages of Precontemplation, on the one hand, and between the two stages of Maintenance, on the other, hardly contribute to increased accuracy of measurement. Having the last two stages combined into one makes the question (SAS) or questions (TSQ) simpler and also reduces the problems with the time-frame component. Both Preparation and Action are by definition temporary stages, and a small proportion of respondents is expected to be found in these stages at any given time. The need to discriminate Action from Maintenance depends on the aims of the specific intervention. Combining Preparation with the stage of Ongoing Occasional Activity is in violation of the theory, as intention, a theoretically necessary component of the Stages of Change Model, discriminates these two stages.

**References**


Kerns, R.D., Rosenberg, R., Jamison, R.N., Caudill, M.A. and Haythornwaite, J. (1997) Readiness to adopt a self-


Received on April 24, 2002; accepted on July 9, 2002