

2009

Validity of Stage Assessment in the Adoption and Maintenance of Physical Activity and Fruit and Vegetable Consumption

Sonia Lippke

Jochen P. Ziegelmann

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/cprc_facpubs

Authors

Sonia Lippke, Jochen P. Ziegelmann, Ralf Schwarzar, and Wayne F. Velicer



Published in final edited form as:

Health Psychol. 2009 March ; 28(2): 183–193. doi:10.1037/a0012983.

Validity of Stage Assessment in the Adoption and Maintenance of Physical Activity and Fruit and Vegetable Consumption

Sonia Lippke,

Department of Psychology, Freie Universität Berlin

Jochen P. Ziegelmann,

Department of Psychology, Freie Universität Berlin

Ralf Schwarzer, and

Department of Psychology, Freie Universität Berlin

Wayne F. Velicer

Cancer Prevention Research Center, University of Rhode Island

Abstract

Objective—Stage assessments are examined to develop and test refined measurements that can be used for classifying individuals.

Design—Stages were assessed in 1,850 persons in terms of their physical activity and dietary behaviors.

Main Outcome Measures—Stages for both behaviors were compared to behavior and other test variables. Misclassification, sensitivity, specificity, receiver-operation-curves, and discontinuity patterns were computed. Discontinuity patterns were tested with trends across stages and planned contrasts between adjacent stages.

Results—In comparison to previous studies, sensitivity (70% to 80%) and specificity (80% to 87%) were high. When using lower level criteria (such as less intensive activity), sensitivity was lower, whereas specificity was higher. When behavioral maintenance was assessed, results suggested that the temporal cut-off point between action and maintenance was equally optimal at different cut-off points. Applying contrast analyses, nonlinear trends across the stages and a match of 87% of predictions of stage differences resulted.

Conclusion—Stage assumptions are supported in general, and refined stage assessment in particular. Levels of psychological variables (e.g., easiness, habit) may discriminate stages as well as or even better than temporal stage definitions.

Keywords

stages; misclassification; specificity; sensitivity; health behavior

Behavior change toward a particular goal (such as smoking cessation) can be understood either as a continuous process or as a series of distinct stages that represent an ordered set of categories into which people can be classified. These categories reflect cognitive or behavioral characteristics, such as the motivation to perform a behavior or the habit of doing so for an

extended period of time. Following Weinstein, Rothman, and Sutton (1998), the common defining properties of stage models are: (a) Individuals can be classified into different stages and a valid assessment exists. (b) The stages are ordered, that is, Stage 3 is closer to the criterion behavior than Stages 1 and 2, and Stage 1 is farthest from the criterion behavior. A person in Stage 1 has to move first to Stage 2 before proceeding to Stage 3 and finally adopting the criterion behavior. (c) Individuals in the same stage are more similar than those in different stages, that is, they face the same barriers, but these barriers are different from those in other stages. The most popular stage theory of health behavior change is the Transtheoretical Model (TTM; Prochaska & DiClemente, 1983) that proposes five stages of change: precontemplation, contemplation, preparation, action, and maintenance. The TTM was originally developed for addictive behaviors in general, and for smoking cessation in particular (e.g., Herzog & Blagg, 2007).

Although stage models have received much attention as well as empirical support (e.g., Hall & Rossi, 2008), there has also been criticism (Brug et al., 2005; Herzog & Blagg, 2007; Sutton, 2000; Weinstein et al., 1998). A meta-analysis found that stage-matched treatments worked significantly better than unmatched ones. However, effects were rather small ($r = .09$; Noar, Benac, & Harris, 2007). An extensive review of the multitude of papers that have evaluated aspects of the TTM can be found elsewhere (e.g., Brug et al., 2005; Hall & Rossi, 2008; Nigg, 2005; Noar et al., 2007; Prochaska et al., 1994; Sutton, 2000; Weinstein et al., 1998).

Most of the criticism has focused on operationalizations by the use of *time frames*: The stage definitions are viewed as being problematic because it seems unclear whether the chosen time frame is actually the proper one. However, because stage is one of the best predictors of outcome with a large effect size (Velicer, Redding, Sun, & Prochaska, 2007), an improvement in measurement can only serve to increase the importance of the stage construct. Time frames for defining stages may be more or less appropriate for different behaviors. An example of a staging algorithm with a very different stage definition is mammography screening (Rakowski et al., 1998), where the target behavior is infrequent, for example, annually. For nutrition and physical activity, empirical evidence in favor of the chosen time frames is lacking (cf. Sutton, 2000). There are three implications: (a) Assessment qualities should be examined systematically (cf. Nigg, 2005); (b) time frames might be studied in more depth; (c) substitutes for time frames (i.e., psychological variables such as habituation) should be investigated.

Theoretically Indicative Variables

Besides variables such as behavior and intention, others are also important when investigating stages. Behavior maintenance may be reflected by variables such as easiness of performance or habituation. Also, the passage of time may be considered when categorizing people. Typically, maintenance is indicated by performing the behavior for “half a year or longer” after the initiation (Marcus, Rossi, Selby, Niaura, & Abrams, 1992). However, because of the desiderata of empirical tests, other time frames, such as 1 or 2 years, should also be tested to see how well they classify individuals into the two active stages (i.e., action, maintenance). Furthermore, due to the fact that no studies could be found testing duration of behavior pattern performance in any stage, the reported time frames should be investigated in addition. The only assumptions deduced from the literature are that the time span since behavior adoption should be significantly different between the two active stages (action, maintenance). Thus, stage assessment quality might be improved by omitting time frames completely in favor of psychological distinctions between the stage groups. This has been done in the present study to avoid the time-frame critique and to create refined stage assessments.

Examining the Psychometric Quality of Stage Assessments

Stage is a complex construct as it includes different variables such as behavior, intention, and duration (time since last behavior change) components. In this article, we examine each of these components. One method of evaluating measurement quality is to examine the percentages of misclassification with respect to behavior. This approach combines the three “nonactive” stages (precontemplation, contemplation, preparation) and the two “active” stages (action, maintenance), and it compares them to the performance of the goal behavior. The comparison focuses on two aspects (Heneghan & Badenoch, 2006): *sensitivity* (agreement between classification as being active and performing the goal behavior¹) and *specificity* (agreement between classification as being nonactive, and the nonperformance of the goal behavior²). Sensitivity is defined as the proportion of persons who are accurately classified as performing the behavior over the total number who meet the criteria for the behavior (correct plus individuals in nonactive stages who meet the behavior criteria). Specificity is defined as the proportion of individuals accurately classified as not performing the behavior over the total number who are not performing the behavior (correct plus individuals in active stages who do not meet the behavior criteria).

Most studies on stages have not reported specificity and sensitivity. The few investigations explicitly addressing assessment qualities have shown dependency on behavior intensity and age: Lee, Nigg, Courneya, and DiClemente (2001) reported higher sensitivity and specificity for strenuous than for moderate physical activity. Plotnikoff et al. (2007) found in individuals with diabetes that sensitivity was 76 to 90% and specificity was 56 to 78%, with linear trends for age (the older the sample, the lower the specificity and the higher the sensitivity). In a review of over 30 studies, Nigg (2005) aggregated specificity and sensitivity for different activity intensities. Overall, the sensitivity for strenuous physical activity was 86%, whereas for moderate activity it was 71%, and for mild activity it was 54%. The specificity was even lower: For strenuous physical activity it was 71%, for moderate activity 63%, and for mild activity 47%. However, no previous studies have tested statistically which algorithm was more adequate. In addition, no such specificity and sensitivity estimates were reported for other components of stage, such as performance duration, which therefore was addressed in this study.

Statistical Test of Stage Assumptions: Testing for Nonlinearity

Sutton (2000) suggested investigating whether change was continuous by testing if discontinuity patterns would result in nonlinear trends. A linear trend would produce a series of ordered significant differences between stages (precontemplation < contemplation < preparation < action < maintenance), with approximately equal differences between each adjacent pair of means. A nonlinear trend is indicated not only by significant increases in means of test variables, but also by unequal magnitudes in the difference between pairs of adjacent means. A linear trend, in contrast, would imply a continuum (Sutton, 2000; Weinstein et al., 1998). Thus, testing the explicit assumptions of linearity between stages is another test of the assumptions of a stage model (Sutton, 2000). The results of the few studies that have tested linearity/discontinuity patterns in the TTM have been inconsistent (e.g., support found by Armitage, Povey, & Arden, 2003; no support, Armitage & Arden, 2002). However, the majority of the TTM studies have not explicitly tested for nonlinear trends even if they have found discontinuity patterns (e.g., studies on smoking cessation, Herzog & Blagg, 2007; Prochaska et al., 1994; Velicer, Prochaska, Rossi, & DiClemente, 1996). Indeed, for smoking cessation, many of the relationships between the TTM variables across stages either increase

¹cf. www.fu-berlin.de/gesund/validation/spec_sens.pdf

²See footnote 1.

monotonically (Processes of Change subscales, Pros of Quitting), or decrease (Cons of Quitting, Self-Efficacy, Temptation scales), but in a curvilinear fashion rather than in a linear one (Johnson, Fava, Velicer, Monroe, & Emmons, 2002; Velicer, Norman, Fava, & Prochaska, 1999). The current study tested whether the relationship between the stages departs from linearity, and this was done separately for two behaviors, nutrition and physical activity.

Statistical Test of Stage Assumptions: Planned Contrasts

The tests for linearity assume that the grouping variable represents equal units on a continuous variable. The stage variable assumes that an implicit ordering exists between the five stages (i.e., precontemplation—contemplation—preparation—action—maintenance), but does not assume that they are equal units apart on a continuum. For example, Armitage and Arden (2002) measured intention in all five stage groups and found varying differences between stages. This means that if differences between stages are regarded as units, these units are unequal if validated with intention, and nonlinear patterns transpire. Thus, merely finding a nonlinear pattern does not provide adequate evidence for rejecting the pseudostage hypothesis. Nonlinearity could be a result of unequal units between the stages. Therefore, a nonlinear term should not lead to a rejection of the linearity assumption because the unequal units represent a plausible alternative explanation. In other words, discontinuity could also result if the difference between one pair of adjacent stages represents a larger or smaller difference between the stages than elsewhere. A test of linearity that suggests a quadratic relationship between stages and a test variable does not necessarily prove the existence of stages. A plausible alternative explanation would be that there are non-equal units. Consequently, when quadratic patterns emerge, the stage model needs to be tested by a subsequent series of planned contrasts. With planned contrasts, theoretically derived assumptions about differences or similarities between adjacent stages can be tested. However, a cubic relationship pattern or above would lead directly to a rejection of the pseudostage hypothesis. For example, such an assumption is that behavior is similar in the nonactive stages (precontemplation, contemplation) and significantly different from the active stages (action, maintenance). The assumption regarding intention would be different: Here an increase across the early, nonactive stages (precontemplation to contemplation to preparation), and no further differences in the two active stages (action, maintenance) would be assumed. Another relevant construct is action plans. Here, no differences are expected within the “nonintentional” stages (precontemplation, contemplation) and within the active stages (action, maintenance). However, after setting a goal (in preparation), plans are crucial, and therefore a substantial increase in contrast to the other nonactive stages should be evident (precontemplation, contemplation). In addition, those who actually succeed in changing their behavior (action) should exhibit higher plan levels than those who remain inactive (preparation). To summarize, this study examined whether the stages of the TTM are qualitatively different by exploring discontinuity patterns of theoretically indicative variables. This was done for two behaviors separately: nutrition and physical activity. Moreover, a refined stage assessment was used.

Research Questions

1. How valid are the refined stage assessments when tested with behavior, intention, and duration of behavior pattern performance? Are there differences in terms of different behavior intensities and between physical activity and dietary behaviors? Validity is understood here in terms of misclassification, specificity, sensitivity, and confirmed hypotheses on stage differences (cf. Table 1, left-hand side, upper part).
2. Can the stage assumptions be supported by discontinuity patterns, that is, nonlinear terms across stages in the test variables, accompanied by planned contrasts?

3. When discontinuities exist, do they confirm the theoretical predictions on the differences between adjacent stages, derived from previous findings and the considerations above (Table 1, left-hand side, lower part)?

Method

Procedure

Online studies give researchers the potential to reach large samples of persons with diverse socioeconomic status and age and from different geographic regions (Rademacher & Lippke, 2007). German-speaking Internet users provided cross-sectional data. The online study was conducted using the software dynQuest (Rademacher & Lippke, 2007) examining dietary behaviors (consuming fruit and vegetables) and physical activity. After the study was introduced, participants provided informed consent and followed a link to a self-administered questionnaire.

Participants

Two thousand two hundred seven potential study participants responded to the initial web page. Of these, 78 persons (3.5%) dropped out after the first page of the questionnaire. Further 279 individuals did not answer one or both of the two questions of the staging algorithm, which was a precondition for being included in the analyses. Thus, the final sample consisted of 1,850 participants (83.8% of those initially recruited). Except for sex and time since behavior change, no other differences between the two groups were found (more men and people performing their current behavior for longer times were nonresponders, $ps > .05$).

The participants in the final sample were 15 to 80 years of age ($M = 37.68$, $SD = 12.56$); 63.7% of them were women. About 50% of the respondents were living with a partner, 67.8% of the sample had completed senior high school. The body mass index (BMI) ranged from 14.26 to 59.43 ($M = 25.21$, $SD = 5.16$). After classifying, 55 (3.0%) of the individuals were underweight (BMI < 18), 928 (50.2%) had normal weight (BMI = 18 to 24), 430 (23.2%) were overweight (BMI = 25 to 30), 27 (1.5%) were obese (BMI > 30), 178 (9.6%) did not provide BMI-related data.

Measures

Stage, behavior, intention, plans, duration of behavior pattern performance, easiness, and habit were assessed for both nutrition and physical activity. Goal criteria for physical activity were that individuals should accumulate 30 minutes or more of at least moderate-intensity physical activity on most days of the week, or 150 minutes per week. In the nutrition domain, the goal criterion was set at five portions of fruit and vegetables per day.

The assessment of stage was a refinement of the algorithm previously developed for exercise and diet (Greene & Rossi, 1998; Marcus et al., 1992) without using a specific time frame (in accordance with Sutton, 2000). For physical activity, participants were asked, "Please think about your typical weeks: Did you engage in physical activity at least 5 days per week for 30 minutes or more (or 2.5 hr during the week), in such a way that you were moderately exhausted?" Regarding nutrition, the question was: "Please think about what you have typically consumed during the last weeks: Did you eat five portions of fruit and vegetables per day?" For both behaviors, the instruction followed: "Please choose the statement that describes you best." Participants responded based on a rating scale with the verbal anchors "No, and I do not intend to start" (precontemplation stage), "No, but I am considering it" (contemplation stage), "No, but I seriously intend to start" (preparation stage), "Yes, but only for a brief period of time" (action stage), and "Yes, and for a long period of time" (maintenance stage). The algorithm was designed like a rating scale, with precontemplation on the very left and

maintenance on the very right, for two reasons: (a) This format is more similar to the assessment of other social–cognitive variables, such as intention and plans, and, by this, measurement correspondence is achieved; and (b) the format is less space consuming.

Directly after the stage assessment, questions regarding behavior maintenance followed (cf. Lippke & Ziegelmann, 2006). First, duration of behavior pattern performance was measured with the question “How long have you been as physically active on a regular basis as you are currently?” and, respectively, “How long have you been eating this way on a regular basis as you are currently?” The time span could be weeks, months, or years. Second, easiness was assessed with “How difficult was it to be so physically active?” and “How difficult was it to eat this way?” The five-point scales ranged from 1 (*totally easy*) to 5 (*totally difficult*). Third, the question about habit was worded “How much has it become habitual to be so physically active?” and “How much has it become habitual to eat this way?” The five-point scales ranged from 1 (*not at all*) to 5 (*absolutely*).

Physical activity behavior was measured with a modified version of the Godin Leisure-Time Exercise Questionnaire (GLTEQ; Godin & Shephard, 1985; Plotnikoff et al., 2007). This self-report measure has been validated with physiological and anthropometric measures (i.e., VO_2 max and body fat; Godin & Shephard, 1985; Jacobs, Ainsworth, Hartman, & Leon, 1993). Participants were asked to report the average number of sessions per week and average duration per week of strenuous (rapid heartbeats, sweating), moderate (not exhausting, light perspiration) and mild (minimal effort, no perspiration) physical activity in the past month. Only activities outside of work duties (not business or at home) were addressed. Responses (product of frequency and duration) for each of these three activity categories were computed. Two sum scores were then computed: strenuous and moderate activities on the one hand, and strenuous, moderate, and mild activities on the other. Regarding nutrition behavior, participants were asked, “How many portions of fruit and vegetables do you eat per day?” The instruction was, “Please think about a typical weekday within the last month. (Please note that potatoes do not count.)” Four categories were provided: “Salad and uncooked vegetables,” “fruit,” “fruit and vegetable juice (number of glasses),” and “cooked and steamed vegetables” (Schwarzer et al., 2007).

Physical activity intention was assessed with three items, as suggested by Nigg (2005), matching the three behavior intensities “I intend to perform the following activities at least 5 days per week for 30 minutes...”: (1) “... strenuous (rapid heartbeats, sweating) physical activities”; (2) “... moderate (not exhausting, light perspiration) physical activities”; and (3) “... mild (minimal effort, no perspiration) physical activity.” Two different scales were aggregated corresponding to the behavior measurement: On the one hand, strenuous and moderate activities (correlation of the two items $r = .17$), and on the other hand strenuous, moderate and mild activities (Cronbach’s alpha of the three items = .34). Thus, items with discriminant validity were combined to obtain an index that reflects a broad construct. Nutrition intention was assessed as in Schwarzer et al. (2007) with a single item: “I intend to eat at least five portions of fruit and vegetables per day.”

Plans were assessed with three items on the when, where, and how of activity and nutrition, taken from Lippke, Ziegelmann, and Schwarzer (2004). The three items assessing physical activity plans were worded: “I have already planned ...”: (1) “... where I will be physically active”; (2) “... how I will be physically active”; and (3) “... when and how often I will be physically active” (Cronbach’s alpha = .91). The three items assessing nutrition plans were worded “I have already planned ...”: (1) “... at which meals during the day I will eat fruit and vegetables”; (2) “... which kinds of fruit and vegetables I will eat”; and (3) “... how I will prepare fruit and vegetables” (Cronbach’s alpha = .83; Schwarzer et al., 2007).

Finally, pros and cons were assessed according to Lippke et al. (2004) and Schwarzer et al. (2007). The stems “If I engage in physical activity at least five days per week for 30 minutes...” and “If I eat five portions of fruit and vegetables every day...”, respectively, were followed by items on the pros and cons. Pros were measured with three items like “... then I would be doing something good for my health” (Cronbach’s $\alpha_{\text{physical activity}} = .57$; $\alpha_{\text{nutrition}} = .72$). Cons were assessed by three items like “... then this costs me a lot” ($\alpha_{\text{physical activity}} = .62$; $\alpha_{\text{nutrition}} = .56$).

Answers for intention, plans, pros and cons were assessed using four-point scales, ranging from 1 (*totally disagree*) to 4 (*totally agree*).

Data Analysis

Misclassification, specificity, and sensitivity were computed by grouping participants according to stage as compared to their behaviors, intentions, and time of behavior pattern performance, respectively.³ Whether sensitivity and specificity rates were significantly different between the particular criteria (lower vs. higher behavior criterion; 6-month vs. 1-year vs. 2-year cut-off points) was tested by a Z test for two proportions. Moreover, receiver operator curves (ROC) and post hoc test for areas under the curve (AUC) were computed.

Stage differences and discontinuity patterns (cf. Statistical Test of Stage Assumptions: Testing for Nonlinearity, Statistical Test of Stage Assumptions: Planned Contrasts, above) were examined by means of polynomial-based contrast analyses using analyses of variance (ANOVAs). Contrast analyses were employed to test the differences between the stages. Polynomial contrast analyses were used to test for nonlinear trends, that is, quadratic, cubic, and fourth-order terms. The trends were tested with adjustment (weighted terms) for unequal sample sizes. Missing data were handled by pairwise deletion. All analyses on stage differences, discontinuity patterns, and trends were performed with SPSS 13.0.

Results

Regarding physical activity, of the 1,850 study participants, 3.5% were in precontemplation, 20.3% in contemplation, 28.5% in preparation, 15.9% in action, and 31.8% in maintenance (for numbers see Table 2). Regarding nutrition, 11.2% study participants were in precontemplation, 29.7% in contemplation, 24.5% in preparation, 10.3% in action, and 24.2% in the maintenance stage. Thus, for physical activity, the majority of the sample was in preparation and maintenance. Regarding dietary behavior, most participants were in contemplation, preparation, and maintenance.

Sensitivity and Specificity Analysis: Physical Activity

The first set of analyses focused on sensitivity and specificity for the external criteria (behavior, intention, time). For behavior, individuals were classified as being nonactive if they were in precontemplation, contemplation, or preparation, as opposed to being active if they were in action or maintenance. This was compared to their scores on the GLTEQ scale, and a 2×2 table was calculated. Two physical activity scores from the GLTEQ were employed: strenuous and moderate activity, as well as strenuous, moderate and mild activity.

On average, 81% of the individuals in nonactive stages (i.e., precontemplation, contemplation, preparation) were correctly classified as being nonactive, 74% were correctly classified as being active by the strenuous and moderate behavior measure from the GLTEQ. Sensitivity and specificity were both 80% (cf. Table 2). For the strenuous, moderate, and mild behavior, 58% were correctly classified as being nonactive, 83% were correctly classified as being active.

³See footnote 1.

Sensitivity was 65% and specificity 81%. Comparing the sensitivity and specificity rates of the two criteria revealed that sensitivity was significantly higher when no mild behavior was included. Specificity rates were not significantly different.

Physical activity intention focused on comparing participants classified as not intending (precontemplation, contemplation) with the three intending stages (preparation, action, maintenance; Table 2). Specificity was significantly higher when no mild behavior was included. Sensitivity rates were not significantly different.

Duration of behavior pattern performance focused on comparing action and maintenance. Three cut-off points for the time dimension were chosen: 6 months, 1 year, and 2 years (see correctly classified individuals in Table 2). Sensitivity was significantly higher for the 2-year cut-off point in comparison to the 6-month and 1-year cut-off points. Specificity was significantly different for all three cut-off points.

The AUC estimates for physical activity behavior, intention, duration of behavior pattern performance, and maintenance indicators were computed for the different criteria.⁴ The AUCs, their standard errors and post hoc tests of significant differences are listed in Table 2 (except for maintenance indicators, which are reported in detail below). Both behavior and intention AUCs were significantly higher when only strenuous and moderate behaviors were regarded, than when mild behavior was additionally included, $t_{\text{behavior}}(1849) = 5.80$; $t_{\text{intention}}(1853) = 3.43$, $ps < .01$, indicating that they had a higher diagnostic accuracy. The AUCs for time were not significantly different. For duration of behavior pattern performance (no time specified) the AUC was 0.94 ($SE = .01$), for easiness the AUC was 0.73 ($SE = .02$) and for habit the AUC was 0.85 ($SE = .01$). All three were significantly different from each other, $t_{\text{duration-easiness}}(1568) = 10.14$, $p < .01$; $t_{\text{easiness-habit}}(1568) = 4.93$, $p < .01$; $t_{\text{duration-habit}}(1568) = -5.29$, $p < .01$.

Sensitivity and Specificity Analysis: Consumption of Fruit and Vegetables

The same comparison between stage and the behavioral assessment of fruit and vegetables was calculated. Two different behavior criteria were considered: five-a-day and three-a-day. On average, 83% of the persons in nonactive stages (i.e., precontemplation, contemplation, preparation) were correctly classified as being below the criterion, 73% of those in active stages were correctly classified as meeting the five-a-day criterion. Sensitivity was 70% and specificity 87% (cf. Table 2). For the three-a-day criterion, 39% were correctly classified as being below the criterion, 95% were correctly classified as meeting the criterion. Sensitivity was 45% and specificity 94%; both were significantly different (cf. Table 2).

Nutrition intention focused on the five-a-day nutrition criterion, comparing nonintenders (precontemplation, contemplation) with intenders in the three later stages (preparation, action, maintenance; Table 2). Duration of behavior pattern performance for nutrition focused on comparing action and maintenance. Three cut-off points for the time dimension were considered: 6 months, 1 year, and 2 years (see Table 2). As for physical activity, the 2-year cut-off point revealed significantly higher sensitivity than the two shorter time periods. Specificity was significantly different between all three cut-off points.

The AUC estimates for nutrition behavior, intention, duration of behavior pattern, performance, and maintenance indicators for the different criteria were computed (AUCs, their standard errors, post hoc test, see Table 2⁴). Behavior AUCs were significantly higher when only the five-a-day behavior was regarded than when the three-a-day behavior was included, $t_{\text{behavior}}(1849) = 7.49$; $p < .01$, indicating that this criterion had a higher diagnostic accuracy.

⁴The AUCs can be accessed at www.fu-berlin.de/gesund/validation/AUC.pdf

The AUCs for time were not significantly different. For duration of behavior pattern performance (no time specified) the AUC was 0.89 ($SE = .02$), for easiness the AUC was 0.76 ($SE = .02$), and for habit the AUC was 0.84 ($SE = .02$). All three were significantly different from each other, $t_{\text{duration-easiness}}(1629) = 5.10, p < .01$; $t_{\text{easiness-habit}}(1629) = -3.06, p < .01$; $t_{\text{duration-habit}}(1629) = 2.15, p < .05$.

Testing for Nonlinearity

Mean differences of stage groups were tested with ANOVA and polynomial contrast analyses for testing the differences between adjacent stages (Tables 3 and 4). Discontinuity patterns were examined by means of polynomial-based contrast analyses using ANOVAs. If only linear trends are observed, this can be interpreted as supporting a pseudostage model, and if quadratic or cubic or higher trends are present, this might suggest discontinuity (Armitage & Arden, 2002). Therefore, polynomial contrast analyses were used to test for nonlinear trends: quadratic, cubic, and fourth-order terms. The trends were tested with adjustment (weighted terms) for unequal sample sizes. Trends and planned contrasts were computed in line with Winer, Brown, and Michels (1991).

For physical activity, 10 tests were performed. All 10 tests required a term higher than linear to provide an adequate fit (see Table 3). For nutrition, 8 tests were performed, all requiring a term higher than linear to provide an adequate fit (cf. Table 4).

Discontinuity Assessment: Behavior

Polynomial contrasts of adjacent stage groups tested the predictions (as reported in Table 1, left-hand side). Analyzing means of physical activity behavior, including only strenuous and moderate physical activity (no mild), individuals in precontemplation, contemplation, and preparation reported on average less than 150 min/week, whereas individuals in action and maintenance reported, on average, more than 150 min/week. This was not true for the behavior measure including also mild activities. This finding corroborates the previous results based on the dichotomized variable (i.e., being below the criterion, or fulfilling the criterion; see previous part on misclassification, specificity, sensitivity). In terms of fruit and vegetable consumption (see Table 4), individuals in precontemplation, contemplation, and preparation reported on average less than five portions per day (which had been the criterion within the staging algorithm, cf. Method section) as opposed to individuals in action and maintenance, who consumed, on average, five portions or more. This finding also corroborates the result based on the dichotomized variable (see previous part on misclassification, specificity, sensitivity).

By pair comparisons (polynomial contrasts) for all behavior measures, it was found that individuals in precontemplation and contemplation as well as those in contemplation and preparation reported on average the same levels (cf. Table 3). This is in accordance with the hypothesis. However, persons in maintenance reported significantly higher levels of behavior than those in action (which opposes our prediction, cf. Table 1).

Discontinuity Assessment: Intentions and Plans

Regarding dietary as well as physical activity intentions, individuals in precontemplation scored, on average, significantly lower than those in contemplation, and in contemplation lower than in preparation. For both physical activity intention measures (with or without mild activity), no significant increase from preparation to action and to maintenance was found. However, this was not the case for nutrition. Participants in action reported significantly higher intentions than those in preparation (which opposes the prediction). All other hypotheses were confirmed; including the same intention levels of individuals in action and in maintenance (cf. Tables 1 and 3). Persons in precontemplation reported, on average, the same levels of activity plans as those in contemplation as well as those in action and maintenance (as predicted).

However, dietary plans differed significantly between the precontemplation and contemplation (opposed to our prediction) and scored equally high in the action and maintenance (in line with the hypothesis). In both behaviors, individuals in contemplation reported less plans than those in preparation, and the highest levels of plans were found in the action and maintenance (as predicted). The other characteristics of behavior, that is, easiness, habit, and duration of behavior pattern performance, were significantly higher in maintenance than in action (as predicted, Table 1).

Short-Term and Long-Term Execution of Behaviors

The characteristics of short-term and long-term execution of behaviors were assessed not only in the two active stages (action, maintenance), but in all stages, separately for physical activity and nutrition. Individuals in precontemplation reported to perform physical activity (or rather their nonactive lifestyle) for the last 21 days up to the last 47 years (cf. Table 3). This was not statistically different ($p > .05$) from contemplation. Contemplators reported to perform their physical activity for the last 7 days up to the last 49 years. Individuals in precontemplation and contemplation were different regarding their perceived easiness and habit of this behavior level (cf. Table 3). Individuals in contemplation and preparation reported equal levels of easiness and habit ($p > .05$), but significantly different time spans for their prior behavior ($p < .05$; preparation range 7 days to 51 years). No differences were found between individuals in preparation and action regarding their perceived easiness ($p > .05$). However, those in action perceived more difficulties and were performing their new behavior for a shorter duration (range 7 days to 20 years). As reported above, persons in the action and maintenance differed significantly ($p < .05$) from each other in all three indicators of behavioral maintenance (maintenance range 21 days to 51 years).

Regarding nutrition, that is, daily consumption of fruit and vegetables, participants in precontemplation reported to perform their dietary lifestyle (or, rather, not eating enough fruit and vegetables) for the last 7 days up to the last 51 years. This and their perceived easiness were statistically different ($p < .05$) from the contemplators (contemplators range 7 days to 51 years). Precontemplators perceived, on average, fewer difficulties ($p < .05$). Precontemplators, contemplators, and preparers reported the same mean level of habits (cf. Table 4, $p > .05$). However, preparers described more difficulties ($p > .05$) and a shorter time span of their past behavior ($p < .05$; preparers range 7 days to 51 years). Preparers and actors differed in all three indicators of behavioral maintenance: Those in action experienced more difficulties, higher habit levels, and shorter time spans of behavior performance (action range 7 days to 49.59 years). Also, actors and maintainers were significantly different in the other three variables: Maintainers experienced more easiness, higher habituation, and longer time spans of behavior performance (maintenance range 14 days to 50 years).

Discussion

This study extends previous research on the examination of the validity of stage assessment by (a) analyzing measurement qualities, mainly focusing on sensitivity and specificity (cf. review by Nigg, 2005), and (b) testing nonlinear terms across means and hypotheses across means of adjacent stages. Two refined stage measures were designed for physical activity and dietary behaviors. The first set of analyses was performed testing misclassification, sensitivity, and specificity, and ROC with different cut-off points. Regarding the ROC analyses, which combine sensitivity and specificity into one outcome, the higher criterion was significantly better than the lower behavior criterion, for both physical activity behavior and intention, and nutrition behavior.

In comparison to previous studies on physical activity (Lee et al., 2001; Nigg, 2005; Plotnikoff et al., 2007), sensitivity was comparably high, and specificity was superior. In addition,

assessment accuracy was examined with regard to lower behavior levels (including mild physical activity as well as consuming only three portions). The results were in accordance with those reviewed by Nigg in terms of lower sensitivity when the behavior criterion was reduced to a lower level. Due to the fact that the lower behavior criterion (mild) incorporated the higher one; specificity was slightly higher with the lower behavior criterion. However, the improvement of specificity was much less (1% in physical activity, 7% in nutrition) than the decline in sensitivity (15% in physical activity, 25% in nutrition).

In the present study, it was also examined whether specificity and sensitivity were different with regard to different intention criteria. An analogue pattern to Nigg's (2005) review emerged: Both specificity and sensitivity decreased, and the diagnostic accuracy was significantly better for the higher behavior criterion. This is intuitive because the behavior criterion stated in the stage assessment was set to physical activity accumulated only with strenuous and moderate activities (no mild activity included) and eating five portions of fruit and vegetables per day (three would not be enough). Thus, the criteria for stage assessments as well as behavior and intention measures should be selected accordingly. Otherwise, the validity of the stage assessment would suffer. It may be that persons who are incorrectly classified as being inactive are those who are doing the behavior, but who perceive themselves as still preparing for doing it, or who are typically performing the goal behavior, but not in the last month (e.g., because of an illness or vacation).

In addition, this was the first study that investigated a time variable (performance duration) of dietary behaviors and physical activity. This was done by testing the validity of stage assessment based on different time cut-off points. These analyses revealed that sensitivity increased and specificity decreased when a longer time span was chosen. When sensitivity and specificity were analyzed jointly (ROC analyses) for both behaviors, the differences between the half-year cut-off point, the 1-year cut-off point, and the 2-year cut-off point were almost negligible (sensitivity and specificity taken together). When comparing the diagnostic accuracy of the duration (no time specified) with other psychological variables, such as easiness and habit, duration functioned well. Thus, not only the 6-month criterion, but also the 1-year and the 2-year, or even later cut-off points may be useful. However, limitations are linked to the present test procedure: For testing the validity of the stage algorithm, a dichotomization was required, although the TTM highlights the importance of each of the five stages.

Another strategy to determine the quality of the stage measures was to test predicted pair comparisons. With that, no dichotomization had to be done, and mean patterns were investigated. Of the 30 predictions, 26 were confirmed (cf. Table 1). In the physical activity domain, all but one hypothesis were supported, and in the nutrition domain, all but three mean differences emerged as predicted. Thus, with a match of 87% of hypotheses and findings, results were promising regarding the refined stage assessment. These comparisons were based on theory and prior research (e.g., Armitage & Arden, 2002; Armitage et al., 2003; Greene & Rossi, 1998; Marcus et al., 1992; Prochaska et al., 1994; Velicer et al., 2007).

A further research question addressed the indicators of behavior maintenance. These analyses were innovative because no other study could be found that assessed easiness, habit, and durations of behavior pattern performance across stages. The results showed that duration of behavior pattern performance were not only different between the two active stages (action, maintenance), but also between almost all other stages. In contrast to other studies, we did not assess the time until the individuals intended to change their behavior (as it is usually done in TTM stage questions, cf. Marcus et al., 1992; Plotnikoff et al., 2007).

The present data reveal that individuals in precontemplation and contemplation seem to be also a kind of "maintainers": They reportedly maintain their inactivity routines (which can be

physical inactivity or not consuming enough fruit and vegetables daily) over a long time period as well. More interesting, because easiness and habit are psychological variables, the other indicators of behavioral maintenance underscored that precontemplators may be “risk behavior maintainers”: They experience less difficulties and more habitual engagement with their current risk behaviors.

These findings may show the usefulness of the refined stage assessment that omitted any particular time span: Not specifying an arbitrary time cut-off for short-term behaviors as opposed to long-term behaviors can be useful. In this study, individuals classified themselves equally correct if cross-validated with different time cut-off points and psychological variables, such as easiness and habit. The variation of time cut-offs to differentiate between action and maintenance showed no clear tendency, that is, it seems to be unimportant to select a specific time span. It may also be valuable to employ psychologically meaningful variables for the purpose of classification (Sutton, 2000). Other social–cognitive variables, such as perceived self-efficacy, have also been found to discriminate between stages (e.g., Luszczynska & Sutton, 2006). Although significantly different from duration of behavior pattern performance (no time specified), habit and easiness transpired as being additionally useful for differentiating action and maintenance. The exclusion of a time dimension may be the reason why variables such as “plans” help differentiate groups beyond the stage measures. However, in the past, the stage construct has been used for multiple purposes and has been adapted across a wide variety of behaviors. Meta-analyses such as the one by Hall and Rossi (2008) demonstrate the robustness of the construct. The findings of the current study can be seen as a potential development and how to overcome measurement weaknesses found by authors such as Resnicow, McCarty, and Baranowski (2003).

This study is limited to the research question of comparing stage assessment with behaviors and cognitions. It was not the purpose to assess stage transitions because this would have required a longitudinal research design. The results, especially on the time criteria, would not necessarily show the same results in the context of an intervention as in this cross-sectional study. However, many studies lack follow-ups to assess maintenance that would go beyond 6 or 12 months. Thus, this study may stimulate research on long-term maintenance. Furthermore, the participants in this online sample might have been at lower risk than the general population. Moreover, reliance on self-reported behaviors is a limitation. Although the measures of behavior employed in this study are well-validated, it would still be desirable to examine additional objective behavioral data for a better validation of the refined stage assessment procedure.

The results show the potential usefulness of the modified stage measures. In the future, these versions can be used in studies investigating physical activity and fruit and vegetable consumption, and they could be adapted to other behavior domains. The advantage of using psychological variables in addition to or instead of time periods is evident in psychological fields, such as assessment/diagnostic and interventions. Further studies should also examine whether the stages can be confirmed experimentally, for example, by using stage-matched interventions as opposed to stage-mismatched interventions (cf. Weinstein et al., 1998). Previous critique of the stage approach can be partly due to the failure to create innovative treatments that match the stages (Noar et al., 2007; Resnicow et al., 2003). This failure may be due to a lack of validity of the staging algorithms as well. If an experimental design fails to demonstrate the desired effects, this does not invalidate the stage approach per se. It may simply mean that the developers have failed to generate a fitting intervention package or did not use a valid staging algorithm.

References

- Armitage CJ, Arden MA. Exploring discontinuity pattern in the Transtheoretical Model: An application of the theory of planned behaviour. *British Journal of Health Psychology* 2002;7:89–103. [PubMed: 14596720]
- Armitage JC, Povey R, Arden MA. Evidence for discontinuity patterns across the stages of change: A role for attitudinal ambivalence. *Psychology & Health* 2003;18:373–386.
- Brug J, Conner M, Harré N, Kremers S, McKellar S, Whitelaw S. The Transtheoretical Model and stages of change: A critique. Observations by five commentators on the paper by Adams, J., and White, M. (2004). Why don't stage-based activity promotion interventions work? *Health Education Research* 2005;20:244–258. [PubMed: 15618324]
- Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences* 1985;10:141–146.
- Greene GW, Rossi SR. Stages of change for reducing dietary fat intake over 18 months. *Journal of the American Dietetic Association* 1998;98:529–534. [PubMed: 9597025]
- Hall KL, Rossi JS. Meta-analytic examination of the strong and weak principles across 48 health behaviors. *Preventive Medicine* 2008;46:266–274. [PubMed: 18242667]
- Heneghan, C.; Badenoch, D. Evidence-based medicine toolkit. 2nd ed.. Malden, MA: Blackwell; 2006.
- Herzog TA, Blagg CO. Are most precontemplators contemplating smoking cessation? Assessing the validity of the stages of change. *Health Psychology* 2007;26:222–231. [PubMed: 17385974]
- Jacobs DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Medical Science in Sports and Exercise* 1993;25:81–91.
- Johnson JL, Fava JL, Velicer WF, Monroe AD, Emmons KM. Testing stage effects in an ethnically diverse sample. *Addictive Behaviors* 2002;27:605–617. [PubMed: 12188595]
- Lee RE, Nigg C, Courneya K, DiClemente CC. Validating motivational readiness for exercise behavior with adolescents. *Research Quarterly in Exercise & Sport* 2001;72:401–410.
- Lippke S, Ziegelmann JP. Understanding and modeling health behavior change: The multi-stage model of health behavior change. *Journal of Health Psychology* 2006;11:37–50. [PubMed: 16314379]
- Lippke S, Ziegelmann JP, Schwarzer R. Behavioral intentions and action plans promote physical exercise: A longitudinal study with orthopedic rehabilitation patients. *Journal of Sport & Exercise Psychology* 2004;26:470–483.
- Luszczynska A, Sutton S. Physical activity after cardiac rehabilitation: Evidence that different types of self-efficacy are important in maintainers and relapsers. *Rehabilitation Psychology* 2006;51:314–321.
- Marcus BH, Rossi JS, Selby VC, Niaura RS, Abrams DB. The stages and processes of exercise adoption and maintenance in a worksite sample. *Health Psychology* 1992;11:386–395. [PubMed: 1286658]
- Nigg CR. There is more to stages of exercise than just exercise. *American College of Sports Medicine* 2005;33:32–35.
- Noar SM, Benac C, Harris M. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin* 2007;133:673–693. [PubMed: 17592961]
- Plotnikoff RC, Lippke S, Reinbold-Matthews M, Courneya KS, Karunamuni N, Sigal RJ, et al. Assessing the validity of a stage measure on physical activity in a population-based sample of individuals with type 1 or type 2 diabetes. *Measurement in Physical Education and Exercise Science* 2007;11:73–91.
- Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology* 1983;51:390–395. [PubMed: 6863699]
- Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for twelve problem behaviors. *Health Psychology* 1994;13:39–46. [PubMed: 8168470]
- Rademacher JDM, Lippke S. Dynamic online surveys and experiments with the free open source software dynQuest. *Behavior Research Methods* 2007;39:415–426. [PubMed: 17958153]
- Rakowski W, Ehrich B, Goldstein MG, Rimer BK, Pearlman DN, Clark MA, et al. Increasing mammography among women aged 40–74 by use of a stage-matched, tailored intervention. *Preventive Medicine* 1998;27:748–756. [PubMed: 9808807]

- Resnicow K, McCarty F, Baranowski T. Are precontemplators less likely to change their dietary behavior? A prospective analysis. *Health Education Research* 2003;18:693–705. [PubMed: 14654502]
- Schwarzer R, Schüz B, Ziegelmann JP, Lippke S, Luszczynska A, Scholz U. Adoption and maintenance of four health behaviors: Theory-guided longitudinal studies on dental flossing, seat belt use, dietary behavior, and physical activity. *Annals of Behavioral Medicine* 2007;33:156–166. [PubMed: 17447868]
- Sutton S. Interpreting cross-sectional data on stages of change. *Psychology & Health* 2000;15:163–171.
- Velicer WF, Norman GJ, Fava JL, Prochaska JO. Testing 40 predictions from the Transtheoretical Model. *Addictive Behaviors* 1999;24:455–469. [PubMed: 10466842]
- Velicer WF, Prochaska JO, Rossi J, DiClemente CC. A criterion measurement model for addictive behaviors. *Addictive Behaviors* 1996;21:555–584. [PubMed: 8876758]
- Velicer WF, Redding CA, Sun X, Prochaska JO. Demographic variables, smoking variables, and TTM-tailored outcomes across five studies of adults. *Health Psychology* 2007;26:278–287. [PubMed: 17500614]
- Weinstein ND, Rothman AJ, Sutton SR. Stage theories of health behavior: Conceptual and methodological issues. *Health Psychology* 1998;17:290–299. [PubMed: 9619480]
- Winer, BJ.; Brown, DR.; Michels, KM. *Statistical principles in experimental design*. New York: McGraw-Hill; 1991.

Table 1

Overview of Predictions and Confirmed Hypotheses

	Hypothesis (Prediction)	Findings: Hypothesis confirmation	
		Both behaviors	Physical activity
Behavior	PC = C	✓	✓
	C = P	✓	✓
	P < A	✓	✓
	A = M	—	—
Intention	PC < C	✓	✓
	C < P	✓	✓
	P = A	✓	—
	A = M	✓	✓
Plans	PC = C	✓	—
	C < P	✓	✓
	P < A	✓	✓
	A = M	✓	✓
Easiness (R)	A < M	✓	✓
Habit	A < M	✓	✓
Duration of behavior pattern performance	A < M	✓	✓
Match of predictions and confirmed hypotheses		14/15	12/15

Note. R means reverse-scored: The higher the value, the more difficulties the individual perceives. Physical activity = accumulating 150 min/week; nutrition = 5-a-day fruit and vegetable consumption; PC = precontemplation; C = contemplation; P = preparation; A = action; M = maintenance.

Table 2

Assessment Qualities of the Physical Activity and the Nutrition Stage Measure

	Correctly Classified ...	Correctly Classified ...	Sensitivity	Specificity	AUC (SE)
<i>Physical activity</i>	... as nonactive	... as active			
Strenuous and moderate behavior measure	81% PC = 80%; C = 84%; P = 80%	74% A = 62%; M = 85%	80% _a	80%	0.85 (.01) ^d _a
Strenuous, moderate, and mild behavior measure	58% PC = 55%; C = 60%; P = 57%	83% A = 75%; M = 90%	65% _a	81%	0.77 (.01) ^d _a
Physical activity intention	... as not intending	... as intending			
Strenuous and moderate behavior criteria	69% PC = 84%; C = 54%	75% p = 75%; A = 74%; M = 77%	85%	43% _b	0.72 (.01) ^b
Strenuous, moderate and mild behavior criteria	72% PC = 84%; C = 60%	58% p = 58%; A = 58%; M = 58%	83%	32% _b	0.66 (.01) ^b
Time of behavior pattern performance	...as initiating	...as maintaining			
6-month cut-off point	79% (= A)	92% (= M)	90% _a	83% _{b,c}	0.85 (.02)
1-year cut-off point	84% (= A)	88% (= M)	92% _b	78% _{b,d}	0.86 (.02)
2-year cut-off point	92% (= A)	80% (= M)	95% _{a,b}	70% _{c,d}	0.86 (.01)
<i>Nutrition (consumption of fruit and vegetables daily)</i>	...as nonactive	...as active			
Behavior: 5-a-day	83% PC = 87%; C = 87%; p = 75%	73% A = 63%; M = 83%	70% _a	87% _b	0.80(.01) ^c
Behavior: 3-a-day	39% PC = 45%; C = 46%; p = 26%	95% A = 91%; M = 98%	45% _a	94% _b	0.67(.01) ^c
5-a-day Intention	...as not intending 73% PC = 94%; C = 62%;	...as intending 82% P = 74%; A = 85%; M = 87%	83%	71%	0.82(.01)
Duration of behavior pattern performance	...as initiating	...as maintaining			
6-month cut-off point	69% (= A)	94% (= M)	88% _a	83% _{b,c}	0.81 (.02)
1-year cut-off point	77% (= A)	89% (= M)	90% _b	75% _{b,d}	0.83 (.02)
2-year cut-off point	85% (= A)	82% (= M)	93% _{a,b}	67% _{c,d}	0.83 (.02)

Note. Subscripts indicate significant differences between proportions and the AUC. AUC = area under the curve; SE = standard error (computed with receiver operator curves); PC = precontemplation; C = contemplation; P = preparation; A = action; M = maintenance.

^a $t(1849) = 5.80, p < .01.$

^b $t(1849) = 3.43, p < .01.$

^c $t(1849) = 7.49, p < .01.$

Table 3

Means and Standard Deviations (in Parentheses) of the Stage Groups, Results From Planned Pair Comparisons as Well as Linear, Quadratic, and Cubic Terms for Physical Activity

Physical activity stage	PC ^a		C ^b		P ^c		A ^d		M ^e		Pair comparisons	Linear term F =	Quadratic term F =	Cubic term F =	Fourth-order term F =	
	M (SD)		M (SD)		M (SD)		M (SD)		M (SD)							
Behavior																
Strenuous and moderate activities	95.19 (140.37)		76.27 (101.69)		100.45 (127.47)		199.03 (160.37)		305.14 (205.07)		PC = C < P < A < M	617.39**	67.58**	1.55	2.12	
Strenuous, moderate, and mild activities	205.08 (268.78)		156.93 (189.12)		180.12 (204.90)		261.99 (213.88)		385.76 (250.93)		PC = C = P < A < M	292.52**	49.80**	0.44	0.01	
Intention																
Strenuous and moderate activities	1.53 (0.61)		2.18 (0.76)		2.71 (0.77)		2.72 (0.83)		2.80 (0.83)		PC < C < P = A = M	182.70**	74.66**	2.21	8.32**	
Strenuous, moderate, and mild activities	1.76 (0.63)		2.32 (0.70)		2.65 (0.72)		2.66 (0.77)		2.69 (0.82)		PC < C < P = A = M	86.30**	49.93**	3.11	2.16	
Plans	2.31 (1.15)		2.41 (1.00)		2.99 (0.79)		3.34 (0.73)		3.37 (0.81)		PC = C < P < A = M	334.21**	21.28**	20.51**	3.03	
Maintenance																
Easiness (R)	2.63 (1.27)		3.06 (1.14)		3.01 (1.02)		2.89 (0.83)		2.13 (0.87)		PC < C = P = A > M	207.31**	98.35**	1.15	5.28*	
Habit	2.97 (1.38)		2.49 (1.20)		2.45 (1.01)		2.86 (0.78)		4.02 (0.77)		PC > C = P < A < M	665.82**	260.38**	2.98	0.59	
Duration of behavior pattern performance (years)	7.43 (9.20)		6.13 (8.10)		3.60 (6.10)		0.58 (1.54)		8.42 (10.10)		PC = C > P > A < M	19.34**	137.19**	57.82**	7.85*	
Social-cognitive variables																
Pros	3.23 (0.73)		3.43 (0.50)		3.65 (0.42)		3.61 (0.42)		3.63 (0.45)		PC < C < P = A = M	51.32**	36.15**	0.54	7.92*	
Cons	2.49 (0.71)		2.35 (0.70)		2.24 (0.65)		2.08 (0.68)		1.96 (0.64)		PC = C = P > A = M	110.40**	0.08	0.01	0.51	

Note. R= reverse scored: The higher the value, the more difficulties the individual perceives. > is larger at $p < .05$, < is smaller at $p < .05$, = is equal at $p > .05$. PC = precontemplation; C = contemplation; P = preparation; A = action; M = maintenance.

^a $n \leq 64$.

^b $n \leq 376$.

^c $n \leq 528$.

^d $n \leq 294$.
^e $n \leq 588$.
* $p < .05$.
** $p < .01$.

Table 4

Means and Standard Deviations (in Parentheses) of the Different Stage Groups, Results From Planned Pair Comparisons as Well as Linear, Quadratic, and Cubic Terms for Nutrition

Nutrition stage	PC ^a		C ^b		P ^c		A ^d		M ^e		Linear term F =	Quadratic term F =	Cubic term F =	Fourth-order term F =
	M(SD)		M(SD)		M(SD)		M(SD)		M(SD)					
Behavior: Portions of Fruit and vegetables per day	2.95 (1.77)	3.10 (3.14)	3.80 (2.65)	5.14 (2.49)	6.79 (10.12)	PC = C = P < A < M	125.46**	10.14**	0.03	0.03				
Intention	1.27 (0.53)	2.12 (0.83)	2.88 (0.87)	3.31 (0.75)	3.37 (0.94)	PC < C < P < A = M	1138.00**	124.00**	4.01*	0.36				
Plans	1.48 (0.65)	1.75 (0.73)	2.08 (0.76)	2.55 (0.75)	2.70 (0.97)	PC < C < P < A = M	534.59**	0.13	4.62*	2.44				
Maintenance														
Easiness (R)	1.53 (0.74)	1.85 (0.90)	2.11 (0.95)	2.56 (0.88)	1.66 (0.86)	PC < C < P < A > M	3.46	131.57**	46.27**	20.37**				
Habit	3.88 (1.06)	3.68 (0.99)	3.53 (0.98)	3.06 (0.80)	4.20 (0.85)	PC = C = P < A < M	21.62**	124.70**	56.70**	30.16**				
Duration of behavior pattern performance (yrs)	14.19 (12.85)	10.94 (11.42)	8.16 (10.46)	1.86 (6.26)	8.61 (9.13)	PC > C > P > A < M	60.34**	47.90**	29.70**	17.41**				
Social-cognitive variables														
Pros	2.65 (0.78)	3.15 (0.60)	3.39 (0.55)	3.54 (0.47)	3.54 (0.49)	PC < C < P < A = M	334.20**	79.04**	1.99	0.87				
Cons	2.32 (0.69)	2.22 (0.67)	2.06 (0.66)	1.95 (0.62)	1.60 (0.59)	PC = C < P < A > M	265.29**	8.76	0.78	2.28				

Note. For pair comparisons: > is larger at $p < .05$, < is smaller at $p < .05$, = is equally at $p > .05$. R = reverse scored; The higher the value, the more difficulties the individual perceives. PC = precontemplation; C = contemplation; P = preparation; A = action; M = maintenance.

^a $n \leq 208$.

^b $n \leq 550$.

^c $n \leq 454$.

^d $n \leq 191$.

^e $n \leq 447$.

* $p < .05$.

** $p < .01$.