Available online at www.sciencedirect.com

SCIENCE DIRECT®

Addictive Behaviors 31 (2006) 1101–1115

ADDICTIVE BEHAVIORS

Replication of subtypes for smoking cessation within the precontemplation stage of change

Milena D. Anatchkova *, Wayne F. Velicer, James O. Prochaska

Cancer Prevention Research Center, 2 Chafee Road, University of Rhode Island, Kingston, RI 02881-0808, United States

Abstract

The Transtheoretical Model constructs have been used in the development of effective tailored interventions for smoking cessation. Recent studies have suggested the existence of clusters within each stage of change, based on measures of the Pros, Cons and the Situational Temptations. The goal of this study is to replicate cluster subtypes within the Precontemplation stage of change in a secondary analysis of data from a sample of current smokers ($N=3967$). Four random samples ranging from 312 to 400 participants were selected from the Precontemplators subgroup. The cluster analyses were performed using Ward's Method on the standard scores from the three scales. Interpretability of the pattern, pseudo F test, and dendograms were used to determine the number of clusters. Four cluster subtypes were found and replicated across samples. The ten processes of change and two smoking behavior variables were used in the external validation of clusters. Statistically significant multivariate effects were found for the processes of change ($p<.05$) and the smoking behavior variables ($p<.001$) in all samples. The cluster patterns closely replicate earlier findings and provide evidence for the existence of clusters subtypes within the Precontemplation stage of change.

Keywords: Stage subtypes; Cluster analysis; TTM; Tailored interventions

* Corresponding author. QualityMetric Inc., 640 George Washington Highway, Suite 201, Lincoln, RI 02865-4207, United States. Tel.: +1 401 334 8800x244; fax: +1 401 334 8801.

E-mail addresses: manatchkova@qualitymetric.com (M.D. Anatchkova), velicer@uri.edu (W.F. Velicer).

URL: http://www.uri.edu/research/cprc (W.F. Velicer).

0306-4603/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.addbeh.2005.08.002

1. Introduction

Smoking is widely accepted as a serious hazard to health and well-being, but regardless of this knowledge large number of people still smoke. Many studies have suggested that among smokers there are different levels of motivation and intention to quit the habit, which should be used for tailoring of intervention programs. The idea is initially proposed within the framework of the Transtheoretical Model (TTM) of behavior change (Prochaska & DiClemente, 1983; Prochaska & Velicer, 1997; Velicer et al., 2000). TTM has been used for the successful development of population-based tailored interventions for smoking cessation (Hollis et al., 2005; Prochaska, DiClemente, Velicer, & Rossi, 1993; Prochaska, Velicer, Fava, Rossi, & Tsoh, 2001; Prochaska et al., 2001; Prochaska et al., 2004; Prochaska et al., 2005; Velicer & Prochaska, 1999; Velicer, Prochaska, Fava, Laforge, & Rossi, 1999). The model has been conceptualized as involving three key dimensions: the temporal, represented by the Stages of change; the dependant variable dimension, which includes the constructs of Decisional Balance, Situational Temptations, and behavioral measures, and the independent variables dimension, which includes the Processes of change (Velicer, Prochaska, Fava, Norman, & Redding, 1998).

The best known of the dimensions is the temporal organizing construct of the Stages of Change (precontemplation, contemplation, preparation, action and maintenance), which represents the intention and readiness to change a target behavior. Precontemplation is the stage in which people do not plan to take any action to quit smoking in the near future, defined as the next six months. Intervention programs that rely on volunteer samples traditionally ignore Precontemplators, since they are the people with the strongest resistance and the lowest motivation for change (Velicer et al., 2005; Glasgow, Lichtenstein, & Marcus, 2003).

Based on the Stages of Change construct, TTM based programs provide interventions that are tailored to participants' readiness to change, instead of the general "one fits all" approach often employed in population-based programs (Norman, Velicer, Fava, & Prochaska, 2000). In recent years the idea of identifying subgroups within the Precontemplation stage has been explored in a number of studies (Norman et al., 2000; Velicer, Hughes, Fava, Prochaska, & DiClemente, 1995; Dijkastra & De Vries, 2000; Schmidt & Gmel, 1999; Crittenden, Manfredi, Warnecke, Cho, & Parsons, 1998). The reason for this interest is that reliable identification of such subgroups increases the number of potential interventions by a factor of 4 or 5 (Norman et al., 2000), allowing a more narrowly targeted intervention approach. Such a targeted approach is of particular interest for interventions designed for people in the Precontemplation stage, since they have the strongest resistance to change and interventions with these smokers are more likely to fail.

Several approaches have been used to distinguish between potential subgroups within the Precontemplation stage. Crittenden et al. (1998) and Dijkastra, De Vries, and Roijackers (1999) distinguished between groups in Precontemplation based on the intention to quit in some longer period than the 6 months used in the staging algorithm. Schmidt and Gmel (1999) used physical health and well-being to distinguish between clusters within the early stages. While these studies illustrate the popularity of the idea of identifying subgroups within the early stages of change, they do not attempt to identify these variables within the TTM. When the goal is to identify subgroups within the stages of change, it seems that the best approach would be to investigate such groups within the TTM framework, since the staging algorithm is an integral part of the model.

Several studies have employed Transtheoretical Model-based variables to identify subgroups. In the first empirical investigation of subtypes within the stage of change, Velicer et al. (1995) identified three

distinct profiles within the Precontemplation stage in a convenience sample of smokers, using the Pros, Cons and Situational Temptations measures in a cluster analysis. The subgroup with a profile most closely corresponding to the stage was labeled Immotive and a profile corresponding to the Contemplation stage was labeled Progressing. Finally a subgroup demonstrating lack of interest and detachment in the cognitive and physiological aspects of smoking and situational temptations was called Disengaged.

Five years later Norman et al. (2000) published a replication of these findings within the first three stages, using a large representative sample of adult smokers. The study produced cluster profiles for Precontemplation similar to the ones reported in the first study. The major departure was in the presence of more than one Disengaged subtype within the stage. The two papers combined provided evidence for the existence of stage subtypes that can be used as a complementary typology system in the design of new tailored interventions. A very similar approach was used by Dijkastra and De Vries (2000) in a sample of European smokers. Three of the resulting clusters closely matched the profiles of the Progressing, Immotives and Disengaged clusters, even though the authors used different labels for the subgroups.

1.1. Cluster analysis

The three studies described above as well as the current study use cluster analysis to determine the number and nature of subgroups within a stage of change. Cluster analysis is a parsimony procedure that operates on the *subjects* in a data set in much the same way as factor analysis operates on the *variables*. Cluster analysis is used to discover groups or subtypes of homogeneous observations in a multivariate data set (Everitt, Landau, & Leese, 2001). This analysis can be described as an empirical method of classification, which leads to more objective and stable results than narrative taxonomy procedures. Such classifications provide a useful summary of the data, and within the field of medicine can be used to separate disease with different etiology, which require differential treatment (Everitt et al., 2001). Cluster analysis can be viewed as a method of exploratory data analysis (Behrens & Yu, 2003).

Agglomerative hierarchical cluster analysis merges units who are the most similar, starting with the fusion of two individuals, running from single individuals through a small number of clusters and ending with a single cluster containing all the individuals. In this manner cluster analysis proceeds from the individual level to the group level with data determining the patterns. The intermediate level is an appropriate level for tailored interventions. Clusters retain the full configural information, i.e., the pattern of high scores and low scores that is unique to each cluster.

Cluster analysis follows a series of decisions (Everitt et al., 2001; Milligan, 1996). The process involves (1) careful selection of relevant independent variables, (2) variable standardization, (3) choice of an appropriate proximity (distance) measure, (4) choice of clustering method, and (5) choice of a stopping rule for determining the number of clusters in the solution. The latter is the most problematic. Many different stopping rules have been suggested, but some subjectivity in the decision is almost inevitably involved. For this reason replication of cluster solutions is an important step for cluster analysis, providing certainty in the selected solution. The details of the steps taken in the current study are described in the Methods section.

1.2. Overview of the current study

The goal of the current study is to try to replicate the findings of cluster subtypes within the Precontemplation stage of change for smoking cessation reported by Norman et al. (2000). Because of the complexity and richness of the data, the results for the Contemplation and Preparation stages are reported in

separate papers (Anatchkova, Velicer, & Prochaska, 2005; Anatchkova, Velicer, & Prochaska, 2006). The study follows the steps and procedures described by Norman et al. (2000) but uses a bigger sample that allows for a larger number of replications. The study can provide additional evidence for the existence of theoretically interpretable, internally consistent and externally valid cluster subtypes within the Precontemplation stage as the last step of cross-replication required in cluster analytical studies.

2. Method

2.1. Participants

The present study is a secondary data analysis. The initial sample was recruited after a screening was completed on 19,236 adults in four offices of a New England managed care system via mail and telephone surveys. Out of these, a total of 4653 were identified as smokers with no serious illness, between the ages of 18 and 75 and competent in English. 85.3% of the eligible subjects ($N=3967$) were recruited at baseline. The stage distribution of the sample was Precontemplation (PC), 1480 (37.3%); Contemplation (*C*), 1775 (44.7%); and Preparation (PR), 712 (17.9%). The primary results for the two outcome studies based on this sample can be found elsewhere (Prochaska et al., 2001; Velicer et al., 1999). Only the data of participants in the Precontemplation stage of change were used for the current analysis. The average age of the Precontemplation subgroup was 38.2 years and 57.1% of the sample was female.

2.2. Measures

The Stages of Change for Smoking Cessation was assessed using the algorithm method, which consists of 5 yes/no questions and a screening question to determine if there is a smoking history. The first question assesses current smoking status. If currently smoking, subsequent questions focus on intention to quit, and if currently not smoking, questions focus on the length of the quit. There are five ordered stages, *Precontemplation (PC), Contemplation (C), Preparation (PR), Action (A),* and *Maintenance (M)*, that describe one's intention and readiness to quit smoking. Smokers in the Precontemplation stage are currently smoking and not thinking about quitting smoking in the next 6 months. The battery of measures administered at baseline also included the Pros of Smoking and Cons of Smoking from the Decisional Balance Inventory (Velicer, DiClemente, Prochaska, & Brandenberg, 1985) and the total score from the Situational Temptations Inventory (Velicer, DiClemente, Rossi, & Prochaska, 1990). These three variables were employed as the basis for the cluster analysis. Also administered at baseline were the ten Processes of Change (Prochaska, Velicer, DiClemente, & Fava, 1988). Two smoking behavior items measured the number of cigarettes smoked per day during the last week and the time elapsed before the first cigarette of the day (Fagerstrom, 1978). These were used in the external validation of the clusters. All measures have been shown to demonstrate adequate reliability and validity in previous studies.

2.3. Procedure

2.3.1. Sample selection

Participants in the Precontemplation stage with missing data on any of the key variables were excluded from the study ($N=44$). The remaining 1436 subjects were randomly divided into four samples ranging

between 300 and 400 subjects each and a cluster analysis was performed within each sample independently. This strategy allowed for an assessment of the extent to which the results replicated and, therefore, were stable across samples.

2.3.2. Variable selection

The variables employed in the cluster analysis were the same used in the earlier studies (Norman et al., 2000; Velicer et al., 1995), namely the Pros of Smoking and Cons of Smoking from the Decisional Balance Inventory and the total Temptation score from the Situational Temptations Scale. The choice of measures is based on the outcome measurement model proposed by Velicer, Rossi, Prochaska, and DiClemente (1996).

2.3.3. Standardization of variables

The three variables used in each cluster analysis were standardized to T-scores (Mean (M)=50; Standard Deviation (SD)=10). Standardization was performed across the entire sample of 1436 subjects and served both to equalize the contribution of each variable and to put the variables on a comparable metric.

2.3.4. Distance metric and clustering algorithm

All analyses were performed using the squared Euclidean distance metric and Ward's minimum variance algorithm (Ward, 1963). Euclidean distance metric was chosen because of its sensitivity to the cluster profile characteristics of shape, level, and scatter (Cronbach & Gleser, 1953; Edelbrock, 1979) and the relative independence of the three variables. Ward's algorithm is a hierarchical agglomerative procedure that has been shown to be one of the better clustering methods in several simulation studies (Edelbrock, 1979; Blashfield, 1976; Milligan, 1980; Milligan & Cooper, 1987; Milligan & Hirtle, 2003).

2.3.5. Determining the number of clusters

Several methods were used to determine the number of clusters as described in the original study. These included the inverse scree test (Lathrop & Williams, 1987, 1989) and the pseudo F test (Calinski & Harabasz, 1974), which were used to narrow the range of cluster solutions that would be interpreted. As a next step, visual inspection of the cluster profiles was performed with a focus on the shape (configuration of scores), level (the mean score of the variables), and scatter (the standard deviation) of the profiles.

2.3.6. External validation

While cluster analysis interpretation and replication establish the internal validity of the cluster, the relationship of cluster membership to variables not included in the development of the clusters provides external validation. Cluster membership served as the independent variable in a series of multivariate analyses of variances (MANOVAs). Significant multivariate results were followed by univariate analysis of variance (ANOVA) and post hoc Tukey tests. For this study, two different variable sets were employed as the dependant variables: (1) two variables assessing smoking behavior and (2) ten Processes of Change scales. The smoking behavior items measured the number of cigarettes smoked per day during the last week and the time elapsed before the first cigarette of the day (Fagerstrom, 1978). The Processes of change, representing the independent variable dimension of the TTM, are the activities that people employ to change their behavior. There are two distinct groups of processes:

behavioral and experiential (Velicer et al., 1996). The behavioral processes included in this study are Helping Relationship (HR), Self-Liberation (SL), Counter Conditioning (CC), Reinforcement Management (RM) and Stimulus Control (SC). The experiential processes included are Consciousness Raising (CR), Environmental Reevaluation (ER), Self-Reevaluation (SR), Social Liberation (SO) and Dramatic relief (DR). These variables are appropriate for establishing external validation of the cluster solution because, while not used in the formation of the clusters, they have theoretical relevance to the clustering variables.

3. Results

3.1. Cluster analysis

The procedures described above have been used to identify the number of clusters in each sample. Three to five clusters were found to describe the data best and were evaluated in detail. A four cluster solution was retained based on cluster profiles and interpretability. The solution replicated well across the four samples of participants in Precontemplation. The cluster profiles for the four samples are presented in Fig. 1. The cluster means and standard deviations are presented in Table 1. The cluster profiles generally replicated the findings of Norman et al. (2000) and the same cluster labels were retained for clarity.

3.1.1. Cluster 1

The first cluster profile was labeled "Immotives" and is characterized by a "V" shape with high scatter and average level. The profile is what would be expected for people in the Precontemplation stage. The means for the Pros and Temptation scores were about a standard deviation above the average and the Cons were a standard deviation below the average. This is a profile characteristic for people who do not consider changing their behavior. The cluster replicated very well across the four samples.

3.1.2. Cluster 2

The second cluster was labeled "Progressing". A flat shape, low scatter, and a medium level characterize this profile. Smokers in this cluster had scores slightly above average on all three scales. This cluster pattern is similar to the profile expected for people in the Contemplation stage of change, only about a half a standard deviation lower in level. The flat shape indicates that these people are still highly tempted to smoke and experience a cognitive conflict, demonstrated by the similar mean scores of the Pros and the Cons. Since they are similar to the next stage, they are viewed as making progress. The cluster replicated well across three of the samples, but demonstrated a higher scatter and lower values on the Pros and Temptations in the first sample (Fig. 1).

3.1.3. Cluster 3

The third cluster was labeled 'Disengaged 1'. A shallow "V" shape, moderate scatter and low level characterize this profile. The scores for Pros and Temptations are just below average, while the Cons are about a standard deviation below average. They are similar to the Immotive profile but are at or below average on all variables, indicating a low level of engagement with smoking for a regular smoker. The cluster replicated well across three of the samples, but demonstrated a very different pattern in Sample 2

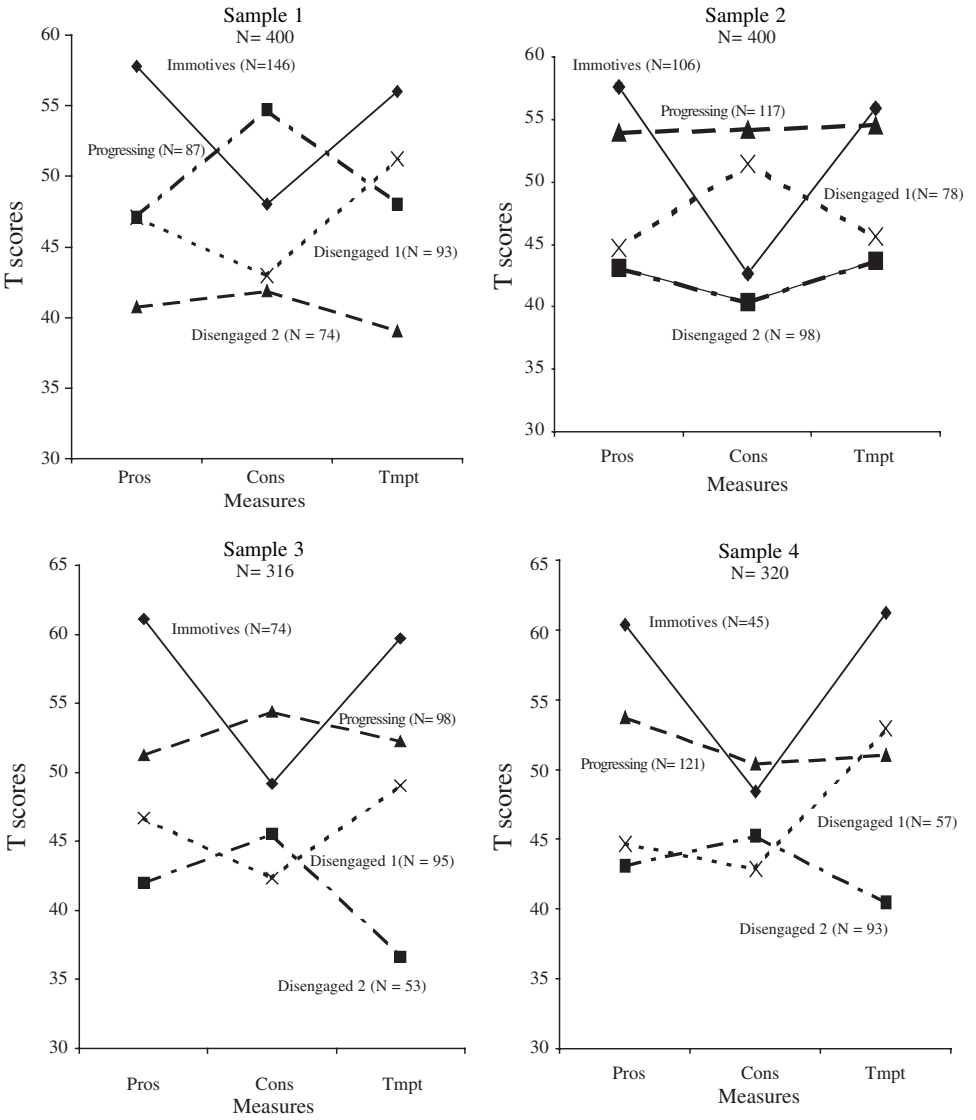

Fig. 1. Four cluster profiles from the four samples.

(Fig. 1), where the Cons were at the average and the Pros and Temptations were about a standard deviation below average, forming an inverted "V" shape.

3.1.4. Cluster 4

The fourth cluster was labeled 'Disengaged 2'. A flat, slightly inverted "V" shape, low scatter and low level characterize this profile. The scores on all scales were about a standard deviation below average. People in this group are smokers who report not being tempted to smoke and do not consider the Pros of smoking important. However, they also do not view the Cons as being important.

Table 1
Means and standard deviations for the four clusters in the four samples for the pros, cons, and temptation scales

	Sample 1 (N=400)		Sample 2 (N=400)		Sample 3 (N=316)		Sample 4 (N=320)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pros								
1. Immotive	57.78	5.14	57.60	6.17	61.10	5.49	60.35	6.23
2. Progressing	47.08	4.3	53.96	5.8	51.24	5.57	53.76	4.91
3. Disengaged 1	47.10	4.44	44.71	4.48	46.72	5.35	44.70	5.17
4. Disengaged 2	40.73	3.22	43.17	4.28	41.97	5.13	43.10	4.68
Cons								
1. Immotive	48.25	6.75	42.70	4.70	49.15	8.19	48.46	8.30
2. Progressing	54.72	4.31	54.13	4.8	54.41	6.21	50.43	6.31
3. Disengaged 1	42.94	4.97	51.44	4.60	42.31	5.44	42.84	4.51
4. Disengaged 2	41.89	5.43	40.42	3.90	45.53	6.21	45.24	6.86
Temptation								
1. Immotive	55.99	6.17	55.86	6.38	59.71	4.07	61.19	3.11
2. Progressing	48.04	6.55	54.52	5.58	52.23	5.34	51.02	5.33
3. Disengaged 1	51.24	5.16	45.58	5.58	49.04	6.51	52.89	3.98
4. Disengaged 2	39.03	6.04	43.74	8.28	36.66	5.88	40.49	6.69

3.2. External validation: behavioral variables

Significant multivariate effects were found on the smoking behavior variables in all samples using MANOVA (Sample 1, Wilk's $\Lambda = .897$, $p < .001$; Sample 2, Wilk's $\Lambda = .904$, $p < .001$; Sample 3, Wilk's $\Lambda = .913$, $p < .001$; Sample 4, Wilk's $\Lambda = .857$, $p < .001$). A summary of the results is presented in Table 2.

These significant multivariate results were followed by two separate ANOVA's. Results revealed that the clusters differed significantly across all four samples on the number of cigarettes smoked per day (Sample 1, $F(3,384) = 4.29$, $p < .05$, $\eta^2 = .032$; Sample 2, $F(3,383) = 7.23$, $p < .05$, $\eta^2 = .054$; Sample 3, $F(3,305) = 4.01$, $p < .05$, $\eta^2 = .039$; Sample 4, $F(3,303) = 6.24$, $p < .05$, $\eta^2 = .058$). The variable was characterized by small effect sizes. A post hoc Tukey test revealed that in all four samples participants in the Immotives cluster smoked more cigarettes than people in the Disengaged 2 cluster. The means for the other clusters positioned them in between these two extremes, but the differences were not significant.

Univariate ANOVA also indicated significant differences among clusters in the number of minutes waited before the first cigarette in the morning (Sample 1, $F(3,384) = 13.46$, $p < .05$, $\eta^2 = .095$; Sample 2, $F(3,383) = 8.85$, $p < .05$, $\eta^2 = .065$; Sample 3, $F(3,305) = 7.91$, $p < .05$, $\eta^2 = .072$; Sample 4, $F(3,303) = 14.86$, $p < .05$, $\eta^2 = .128$). The effect sizes for this variable were in the medium range. The follow up Tukey tests revealed that the Immotives had the shortest waiting time while participants in the Disengaged 2 cluster waited the longest before lighting their first cigarette of the day. These findings are consistent with the results reported by Norman et al. (2000).

3.3. External validation: processes of change

A Multivariate Analysis of variance indicated significant multivariate effects for the 10 processes of change in all four samples (Sample 1, Wilk's $\lambda = .740$, $p < .001$; Sample 2, Wilk's $\lambda = .673$, $p < .001$;

Table 2
Means, standard deviations and cluster differences for the smoking behavior items across four samples

Smoking behavior by cluster	Sample 1, *df*(3,384)			Sample 2, *df*(3,383)			Sample 3, *df*(3,305)			Sample 4, *df*(3,303)		
	M	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern
Cigarettes per day	F=4.29*, η^2=.032		1>2=4	*F*=7.23*, η^2=.054		1>2=4=3	*F*=4.01*, η^2=.039		1=2>4	*F*=6.24*, η^2=.058		1=2=3>4
1. Immotive	24.4	16.1		27.8	18.5		26.0	18.5		25.2	9.7	
2. Progressing	18.2	12.7		22.0	16.0		22.4	15.2		23.5	15.8	
3. Disengaged 1	23.2	16.9		16.7	11.8		21.0	13.3		21.8	12.8	
4. Disengaged 2	18.0	17.3		21.3	16.1		16.4	12.5		16.0	14.6	
Minutes before first cigarette	F=13.46*, η^2=.095		4=2>1=3	*F*=8.85*, η^2=.065		4=3>2=1	*F*=7.91*, η^2=.072		4>2>1; 4>3	*F*=14.86*, η^2=.128		4>3=2=1
1. Immotive	34.9	60.5		28.1	51.2		26.2	35.4		15.3	19.6	
2. Progressing	193.2	144.8		41.9	58.1		67.0	94.4		36.8	50.6	
3. Disengaged 1	45.2	44.8		83.4	128.2		56.1	80.6		53.9	81.1	
4. Disengaged 2	120.2	160.1		80.7	115.9		110.2	160.0		106.0	135.1	

*$p<.05$; 1 = Immotive, 2 = Progressing, 3 = Disengaged 1, and 4 = Disengaged 2.

Sample 3, Wilk's $\lambda=.693$, $p<.001$; Sample 4, Wilk's $\lambda=.797$, $p<.001$). These descriptive statistics, univariate *F* values and Tukey patterns for the four samples are presented in Table 3.

In the first sample follow up ANOVA indicated significant differences among clusters on seven of the ten processes: (1) Consciousness Raising, $F(3,378)=14.78$, $p<.05$, $\eta^2=.105$; (2) Self-Reevaluation, $F(3,378)=20.73$, $p<.05$, $\eta^2=.141$; (3) Environmental Reevaluation, $F(3,378)=5.05$, $p<.05$, $\eta^2=.039$; (4) Dramatic Relief, $F(3,378)=14.02$, $p<.05$, $\eta^2=.10$; (5) Social Liberation, $F(3,378)=5.38$, $p<.05$, $\eta^2=.041$; (6) Stimulus Control, $F(3,378)=2.50$, $p<.05$, $\eta^2=.019$; and (7) Reinforcement Management, $F(3,378)=5.07$, $p<.05$, $\eta^2=.039$. The effect sizes ranged from medium to high with the exception of the results for stimulus control, which had a small effect size. Across all process that reached significance follow up Tukey tests revealed that participants in the Progressing cluster are using the processes significantly more than participants in the Disengaged clusters. Some variation in the patterns was detected, but in general Immotives and Progressing participants were using the processes more actively than the ones in the two Disengaged clusters. No differences among clusters were discovered in the utilization of the processes of Helping Relationship, Counter Conditioning and Self-liberation.

In the second sample, significant differences among clusters were revealed for all but one of the processes: (1) Consciousness Raising, $F(3,381)=12.41$, $p<.05$, $\eta^2=.089$; (2) Self-Reevaluation, $F(3,381)=31.31$, $p<.05$, $\eta^2=.198$; (3) Environmental Reevaluation, $F(3,381)=13.97$, $p<.05$, $\eta^2=.099$; (4) Dramatic Relief, $F(3,381)=20.78$, $p<.05$, $\eta^2=.141$; (5) Social Liberation, $F(3,381)=8.82$, $p<.05$, $\eta^2=.065$; (6) Counter Conditioning, $F(3,381)=4.32$, $p<.05$, $\eta^2=.033$; (7) Stimulus Control, $F(3,381)=8.75$, $p<.05$, $\eta^2=.065$; (8) Self-Liberation, $F(3,381)=8.68$, $p<.05$, $\eta^2=.064$; and (9) Reinforcement Management, $F(3,381)=14.75$, $p<.05$, $\eta^2=.104$. As in the previous sample the effect sizes ranged from medium to large. The follow up Tukey tests revealed a pattern similar to the one described for the first sample with the single exception of Counter Conditioning, which was slightly

Table 3
Means, standard deviations and cluster differences for the processes of change across four samples

Processes of change by cluster	Sample 1, *df*(3,378)			Sample 2, *df*(3,381)			Sample 3, *df*(3,303)			Sample 4, *df*(3,306)		
	M	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern
Consciousness raising	F=14.78*, η^2=.105		2>1>4=3	*F*=12.41*, η^2=.089		2=3>4, 2>1	*F*=11.22*, η^2=.100		2>1>3=4	*F*=5.69*, η^2=.053		2>4=3
1. Immotive	47.5	8.8		44.6	8.3		48.0	10.2		46.1	10.1	
2. Progressing	50.5	7.8		49.7	10.3		51.7	9.4		48.9	9.1	
3. Disengaged 1	42.8	6.7		47.1	9.2		44.9	7.8		43.6	7.9	
4. Disengaged 2	44.0	8.7		42.4	7.2		43.3	9.2		45.0	8.5	
Self-reevaluation	F=20.73*, η^2=.141		2>1>3=4	*F*=31.31*, η^2=.198		2>3>4, 2>1	*F*=25.77*, η^2=.203		2>1>3=4	*F*=8.88*, η^2=.080		2>4=1=3
1. Immotive	45.9	9.3		42.0	7.9		46.1	10.4		43.6	8.9	
2. Progressing	49.1	8.7		49.6	9.1		50.8	9.5		47.7	9.2	
3. Disengaged 1	39.9	6.6		45.1	8.5		41.9	7.3		43.3	8.2	
4. Disengaged 2	41.7	7.07		39.5	5.6		39.2	6.3		41.8	8.2	
Environmental reevaluation	F=5.05*, η^2=.039		2>3	*F*=13.97*, η^2=.099		2=3>1=4	*F*=8.31*, η^2=.076		2>1>3, 2>4	*F*=4.14*, η^2=.039		2>4=3
1. Immotive	48.6	9.3		45.0	8.4		49.1	10.2		46.5	9.7	
2. Progressing	50.6	9.5		50.5	10.0		52.5	10.3		49.4	9.7	
3. Disengaged 1	46.9	8.3		47.0	9.4		46.2	7.8		44.6	7.7	
4. Disengaged 2	45.5	8.9		43.9	7.3		46.2	7.8		46.5	8.4	
Dramatic relief	F=14.02*, η^2=.100		2>1=3=4	*F*=20.78*, η^2=.141		2=3>1=4	*F*=10.15*, η^2=.091		2=1>3>4	*F*=4.99*, η^2=.047		2>4=1=3
1. Immotive	45.9	8.4		43.8	8.1		48.4	10.5		43.9	7.5	
2. Progressing	50.8	8.6		50.4	9.4		50.3	9.7		47.9	9.0	
3. Disengaged 1	43.1	7.4		48.1	8.8		45.5	8.0		43.4	7.5	
4. Disengaged 2	44.1	7.5					45.5	8.0		44.7	8.7	
Social liberation	F=5.38*, η^2=.041		2=1=3>4	*F*=8.82*, η^2=.065		2=1>4, 2>3	*F*=9.00*, η^2=.082		1=2>3=4	*F*=1.15, η^2=.011		2=4=1=3
1. Immotive	49.4	10.5		49.2	10.4		51.9	10.5		47.6	12.3	
2. Progressing	51.3	8.7		52.2	8.7		52.4	8.4		48.5	10.0	
3. Disengaged 1	49.2	11.9		48.4	9.7		46.2	9.6		45.5	9.5	
4. Disengaged 2	44.7	10.1		45.2	10.8		43.5	9.2		46.6	11.3	
Helping relationship	F=1.68, η^2=.013		2=3=4=1	*F*=2.39, η^2=.019		3=2=4=1	*F*=2.21, η^2=.021		2>4	*F*=2.76*, η^2=.026		2=4>3
1. Immotive	47.7	9.2		46.7	7.8		49.0	11.0		48.6	11.0	
2. Progressing	50.6	9.8		49.5	10.2		50.9	9.6		50.0	9.7	
3. Disengaged 1	48.5	10.1		50.1	10.4		48.6	9.7		45.6	8.8	
4. Disengaged 2	48.3	9.1		47.7	10.5		46.5	10.0		49.6	10.1	

Table 3 (*continued*)

Processes of change by cluster	Sample 1, *df*(3,378)			Sample 2, *df*(3,381)			Sample 3, *df*(3,303)			Sample 4, *df*(3,306)		
	M	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern	*M*	SD	Tukey HSD pattern
Counter conditioning	F=.467, η^2=.004		2=3=4=1	*F*=4.32*, η^2=.033		3>4=1	*F*=2.48, η^2=.024		2>1	*F*=4.77*, η^2=.045		4=2>1
1. Immotive	47.0	9.4		45.6	9.4		45.7	11.4		43.4	9.3	
2. Progressing	48.5	8.8		48.4	9.6		50.3	11.5		48.5	9.7	
3. Disengaged 1	47.6	10.1		50.2	10.4		48.5	9.2		46.3	9.1	
4. Disengaged 2	47.6	8.6		46.1	8.9		47.8	11.2		49.3	8.7	
Stimulus control	F=2.50*, η^2=.019		2>3	*F*=8.75*, η^2=.065		2=3>1=4	*F*=1.31, η^2=.013		2=3=1=4	*F*=2.72*, η^2=.026		2>3
1. Immotive	47.5	8.5		44.8	4.0		47.7	9.5		46.1	5.6	
2. Progressing	49.5	9.1		48.2	7.9		49.4	11.1		48.7	9.7	
3. Disengaged 1	45.9	9.8		47.3	7.0		47.9	8.1		45.5	5.2	
4. Disengaged 2	47.9	6.6		44.7	4.0		46.3	6.7		46.7	7.3	
Self-liberation	F=1.52, η^2=.012			*F*=8.68*, η^2=.064		2=3>1=4	*F*=6.63*, η^2=.062		2>1=3=4	*F*=1.98, η^2=.019		2=4=1=3
1. Immotive	44.1	9.8		42.0	9.0		43.9	10.3		43.3	10.5	
2. Progressing	46.5	9.9		47.5	9.9		48.9	9.3		46.1	9.2	
3. Disengaged 1	44.6	10.1		46.0	8.5		43.8	8.5		43.2	9.4	
4. Disengaged 2	43.4	9.2		42.5	9.2		43.5	9.2		45.9	8.9	
Reinforcement management	F=5.07*, η^2=.039		2>3=4	*F*=14.75*, η^2=.104		2>3>4, 2>1	*F*=5.37*, η^2=.051		1=2>3=4	*F*=7.05*, η^2=.065		2>4>3, 1>3
1. Immotive	49.0	9.3		46.1	7.2		49.5	10.3		48.4	9.8	
2. Progressing	52.3	10.2		52.8	11.0		49.3	10.1		49.8	9.7	
3. Disengaged 1	47.9	8.0		49.1	9.6		45.7	6.4		43.7	3.6	
4. Disengaged 2	46.9	8.4		45.5	7.2		45.0	6.3		46.6	8.2	

*$p<.05$; 1 = Immotive, 2 = Progressing, 3 = Disengaged 1, and 4 = Disengaged 2.

more widely used by participants in the Disengaged 1 cluster than the other clusters. In this sample only the results for Helping Relationships failed to reach significance.

In the third sample ANOVA demonstrated differences among clusters on the following seven processes: (1) Consciousness Raising, $F(3,303)=11.22$, $p<.05$, $\eta^2=.10$; (2) Self-Reevaluation, $F(3,303)=25.77$, $p<.05$, $\eta^2=.203$; (3) Environmental Reevaluation, $F(3,303)=8.31$, $p<.05$, $\eta^2=.076$; (4) Dramatic Relief, $F(3,303)=10.15$, $p<.05$, $\eta^2=.091$; (5) Social Liberation, $F(3,303)=9.00$, $p<.05$, $\eta^2=.082$; (6) Self-Liberation, $F(3,303)=6.63$, $p<.05$, $\eta^2=.062$; and (7) Reinforcement Management, $F(3,303)=5.37$, $p<.05$, $\eta^2=.051$. Once again the effect sizes were in the medium to large range. Participants in the Progressing cluster used the processes the most and participants in the two Disengaged clusters used the processes the least. No significant results were obtained for Stimulus Control, Helping Relationship and Counter Conditioning.

The results from the univariate ANOVA's in the fourth sample revealed significant differences in the use of eight of the processes: (1) Consciousness Raising, $F(3,306)=5.69$, $p<.05$, $\eta^2=.053$; (2) Self-Reevaluation, $F(3,306)=8.88$, $p<.05$, $\eta^2=.080$; (3) Environmental Reevaluation, $F(3,306)=4.14$,

$p<.05$, $\eta^2=.039$; (4) Dramatic Relief, $F(3,306)=4.99$, $p<.05$, $\eta^2=.047$; (5) Helping Relationship, $F(3,306)=2.76$, $p<.05$, $\eta^2=.026$; (6) Counter Conditioning, $F(3,306)=4.77$, $p<.05$, $\eta^2=.045$; (7) Stimulus Control, $F(3,306)=2.72$, $p<.05$, $\eta^2=.026$ and (8) Reinforcement Management, $F(3,306)=7.05$, $p<.05$, $\eta^2=.065$. As in the previous samples the effect sizes were in the medium range. Participants in the Progressing cluster used the processes the most and participants in the two Disengaged clusters used the processes the least. Counter conditioning was the only exception, in which participants in the Disengaged 2 cluster were using the processes the most. Social liberation and Self-Liberation failed to reach significance in this sample.

4. Discussion

The main finding of this study is the replication of cluster subtypes within the Precontemplation stage of change reported by Norman et al. (2000). In addition these results replicated three of the clusters in Precontemplation reported by Dijkastra and De Vries (2000). These results seem to support the authors' conclusion that some cultural differences may be accountable for the existence of specific clusters of unmotivated smokers in European samples that do not emerge in American samples. One reason for the existence of this difference could be the stronger social pressure to quit characteristic for the United States.

Cluster analyses replication with a different sample is an important step in the research process. The current study matched cluster solution patterns across four randomly drawn samples of a large representative study. The existence of cluster subtypes within any of the stages of change is important as it allows for more precise tailoring of interventions. The understanding of cluster subtypes within the Precontemplation stage is of particular interest, because it would allow for more precise and targeted interventions within the stage that traditionally present the most serious challenge for behavior change. Classification into different cluster subtypes within the Precontemplation stage suggests that the main reason for this resistance to change can have various sources. In the current study, four separate clusters were replicated.

The most stable cluster was the Immotive cluster which has the profile of classic Precontemplator with high Pros and Temptations to smoke and low Cons. The main intervention strategy for these participants should aim to increase the Cons of smoking and teach effective strategies to deal with Temptations. Participants in the Progressing cluster are a step ahead, since for them the Cons and Pros of smoking are at similar levels. This leads to a cognitive dissonance caused by the equal levels of the Pros and the Cons, complicated by the rather high level of Temptations. Intervention efforts need to focus on reducing the Pros and training in temptation resistance. The cluster profile was somewhat less stable than the Immotives, with more expressed scatter in one of the samples.

The two Disengaged clusters also replicated the results of the study by Norman et al. (2000). The cluster labeled Disengaged 2 had a very stable pattern across all samples, allowing the conclusion that it represents smokers for whom neither the Pros and Cons nor the Temptations are perceived as important. This pattern differed from the Disengaged 1 cluster, which demonstrated high levels of Temptations in three of the samples. The Disengaged clusters are rather challenging from an intervention standpoint since participants in these clusters are not very involved with their smoking. Interventions could focus on efforts to increase the Cons of smoking and clarification of the damages that occasional smoking can bring.

External validation results supported the validity of four clusters. When the processes of change were used, results indicated that people in the Progressing cluster usually used the processes the most and participants in the two Disengaged clusters used the processes the least. This observation is consistent with the results reported by Norman et al. (2000). Also as a replication of previous findings (Norman et al., 2000) the experiential processes of Consciousness Raising, Self-Reevaluation and Dramatic Relief had effect size greater than half a standard deviation (medium to large effect sizes (Cohen, 1988)) across all four samples. In the current study such high effect size was also detected for Reinforcement Management in three of the samples. Again consistent with previous reports that the magnitude of cluster differences for the remaining processes were weaker, varying from small to medium effect size.

External validation was also provided by the two smoking variables. People in the Immotive cluster smoked the largest number cigarettes per day. Participants in the Disengaged 2 cluster waited the longest before they lighted the first cigarette of the day. Exactly the same pattern of finding was reported by Norman et al. (2000). Such close replication provides strong evidence for the existence and stable presence of these clusters within the Precontemplation stage. Not only similar profiles across the selected for cluster analysis variables emerge, but the differences in process use and behavioral patterns across the two studies are remarkable.

It needs to be pointed out that the presence of these subtypes is considered a complementary rather than competing classification of smokers that can provide an additional level for tailoring of interventions. Compared to individual tailoring, interventions based on the cluster subtypes could be more cost-effective, while at the same time providing more specific information than programs tailored only on the stages of change variable. This more detailed level of tailoring may be especially important for Precontemplators, since they are the group that is most resistant to change. Finally, while this replication of the cluster subtypes within Precontemplation increases the confidence in their presence and utility for smoking interventions, prospective studies demonstrating the relationship between cluster changes in behavior at later time points would provide even stronger support.

4.1. Limitations

This study has the inherent limitation of the method used. Cluster analysis is an exploratory procedure and the final solution depends to some extent on the interpretation of the researcher. This fact underscores the importance of independent replications of previous findings, which can increase the confidence in the results.

Another limitation of the study is that it focuses on a single stage. The reason for this decision was the large number of replications (four) and the extensive external validity. Corresponding analyses for the Contemplation and Preparation stages of change will be presented in separate research reports (Anatchkova et al., 2005, 2006).

Acknowledgement

This study was supported by NCI Grants 71356, Grants CA 50087 and CA 27821. A preliminary version of this paper was presented at the annual meeting of the Society for Behavioral Medicine, Baltimore, MD, March 2004.

References

Anatchkova, M. D., Velicer, W. F., & Prochaska, J. O. (2005). Replication of subtypes for smoking cessation within the contemplation stage of change. *Addictive Behaviors*, *30*, 915–927.

Anatchkova, M. D., Velicer, W. F., & Prochaska, J. O. (2006). Replication of subtypes for smoking cessation with the preparation stage of change. *Addictive Behaviors*, *31*, 1101–1115.

Behrens, J. T., & Yu, C. H. (2003). Exploratory data analysis. In I. B. Weiner (Ed.), *Handbook of psychology* (pp. 33–64). New York: John Wiley and Sons.

Blashfield, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. *Psychological Bulletin*, *83*, 377–388.

Calinski, R. B., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*, 1–27.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, NJ: Erlbaum.

Crittenden, K. S., Manfredi, C., Warnecke, R. B., Cho, Y. I., & Parsons, J. A. (1998). Measuring readiness and motivation to quit smoking among women in public health clinics: Predictive validity. *Addictive Behaviors*, *23*, 191–199.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing the similarity between profiles. *Psychological Bulletin*, *50*, 456–473.

Dijkastra, A., & De Vries, H. (2000). Clusters of precontemplating smokers defined by the perception of the pros, cons and self-efficacy. *Addictive Behaviors*, *25*, 273–285.

Dijkastra, A., De Vries, H., & Roijackers, T. (1999). Targeting smokers with low readiness to change with tailored and non-tailored self-help materials. *Preventive Medicine*, *28*, 203–211.

Edelbrock, C. (1979). Mixture model tests of hierarchical clustering algorithms: The problem of classifying everybody. *Multivariate Behavioral Research*, *14*, 367–384.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis*. (4 ed.). New York: Oxford University Press.

Fagerstrom, K. (1978). Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors*, *3*, 235–241.

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? *Amercian Journal of Public Health*, *93*, 1261–1267.

Hollis, J. F., Polen, M. R., Whitlock, E. P., Lichtenstein, E., Mullooly, J. P., & Velicer, W. F. (2005). TEEN REACH: Outcomes from a randomized controlled trial of a tobacco reduction program among teens seen in primary medical care. *Pediatrics*, *115*, 981–999.

Lathrop, R. G., & Williams, J. E. (1987). The reliability of the inverse scree tests for cluster analysis. *Educational and Psychological Measurement*, *47*, 953–959.

Lathrop, R. G., & Williams, J. E. (1989). The shape of the inverse scree tests for cluster analysis. *Educational and Psychological Measurement*, *49*, 827–837.

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation of fifteen clustering algorithms. *Psychometrika*, *45*, 325–342.

Milligan, G. W. (1996). Clustering validation: Results and implications for applied analyses. In P. Arabie, L. G. Hubert, & G. DeSorte (Eds.), *Clustering and classification* (pp. 341–375).

Milligan, G. W., & Cooper, M. C. (1987). Methodology review: Clustering methods. *Applied Psychological Measurement*, *11*, 329–354.

Milligan, G. W., & Hirtle, S. C. (2003). Clustering and classification methods. In I. B. Weiner (Ed.), *Handbook of psychology* (pp. 165–186). New York: John Wiley and Sons.

Norman, G. J., Velicer, W. F., Fava, J. L., & Prochaska, J. O. (2000). Cluster subtypes within stage of change in a representative sample of smokers. *Addictive Behaviors*, *25*, 183–204.

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-help smoking: toward an integrative model of change. *Journal of Consulting Psychology*, *51*, 390–395.

Prochaska, J. O., DiClemente, C., Velicer, W. F., & Rossi, J. S. (1993). Standardized, individualized, interactive and personalized self-help programs for smoking cessation. *Health Psychology*, *12*, 399–405.

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, *12*, 38–48.

Prochaska, J. O., Velicer, W. F., DiClemente, C. C., & Fava, J. L. (1988). Measuring the processes of change: Applications to the cessation of smoking. *Journal of Consulting and Clinical Psychology*, *56*, 520–528.

Prochaska, J. O., Velicer, W. F., Fava, J. L., Rossi, J. S., & Tsoh, J. Y. (2001). Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation. *Addictive Behaviors*, *56*, 583–602.

Prochaska, J. O., Velicer, W. F., Fava, J. L., Ruggiero, L., Laforge, R. G., Rossi, J. S., et al. (2001). Counselor and stimulus control enhancements of a stage-matched expert system intervention for smokers in a managed care setting. *Preventive Medicine*, *32*, 23–32.

Prochaska, J. O., Velicer, W. F., Redding, C. A., Rossi, J. S., Goldstein, M., DePue, J., et al. (2005). Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer and receive regular mammograms. *Preventive Medicine*, *41*, 406–416.

Prochaska, J. O., Velicer, W. F., Rossi, J. S., Redding, C. A., Green, G. W., Rossi, S. R., et al. (2004). Multiple risk expert system interventions: Impact of simultaneous stage-matched expert system interventions for smoking, high fat diet and sun exposure in a population of parents. *Health Psychology*, *23*, 503–516.

Schmidt, H., & Gmel, G. (1999). Identification and characteristics of clusters of smokers within the early stages of change. *Swiss Journal of Psychology*, *58*, 111–122.

Velicer, W. F., DiClemente, C. C., Prochaska, J. O., & Brandenberg, N. (1985). A decisional balance measure for assessing and predicting smoking status. *Journal of Personality and Social Psychology*, *48*, 1279–1289.

Velicer, W. F., DiClemente, C., Rossi, J. S., & Prochaska, J. O. (1990). Relapse situations and self efficacy: An integrative model. *Addictive Behaviors*, *15*, 271–283.

Velicer, W. F., Hughes, S. L., Fava, J. L., Prochaska, J. O., & DiClemente, C. C. (1995). An empirical typology of subjects within stage of change. *Addictive Behaviors*, *20*, 299–320.

Velicer, W. F., Keller, S., Friedman, R. H., Fava, J. L., Gulliver, S. B., Ward, R. M., et al. (2005). Comparing participants and non-participants recruited for an effectiveness study of nicotine replacement therapy. *Annals of Behavioral Medicine*, *29*, 181–191.

Velicer, W. F., & Prochaska, J. O. (1999). An expert system intervention for smoking cessation. *Patient Education and Counseling*, *36*, 119–129.

Velicer, W. F., Prochaska, J. O., Fava, J. L., Laforge, R. G., & Rossi, J. S. (1999). Interactive versus non-interactive interventions and dose-response relationships for stage matched smoking cessation programs in a managed care setting. *Health Psychology*, *18*, 21–28.

Velicer, W. F., Prochaska, J. O., Fava, J. L., Norman, G. J., & Redding, C. A. (1998). Smoking and stress: Applications of the transtheoretical model of behavior change. *Homeostasis*, *38*, 216–233.

Velicer, W. F., Prochaska, J. O., Fava, J. L., Rossi, J. S., Redding, C. A., Laforge, R. G., et al. (2000). Using the transtheoretical model for population-based approaches to health promotion and disease prevention. *Homeostasis in Health and Disease*, *40*, 174–195.

Velicer, W. F., Rossi, J. S., Prochaska, J. O., & DiClemente, C. C. (1996). A criterion measurement model for health behavior change. *Addictive Behaviors*, *21*, 555–584.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244.