PSYCHOMETRIKA—VOL. 77, NO. 1, 85–105
JANUARY 2012
DOI: 10.1007/S11336-011-9235-3

# THE CLASSI-N METHOD FOR THE STUDY OF SEQUENTIAL PROCESSES

EVA VANDE GAER, EVA CEULEMANS, IVEN VAN MECHELEN, AND PETER KUPPENS

KATHOLIEKE UNIVERSITEIT LEUVEN

In many psychological research domains stimulus-response profiles are explained by conjecturing a sequential process in which some variables mediate between stimuli and responses. Charting sequential processes is often a complex task because (1) many possible mediating variables may exist, and (2) interindividual differences may occur in the relationship between these mediating variables and the response. Recently, Ceulemans and Van Mechelen (Psychometrika 73(1):107–124, 2008) addressed these challenges by developing the CLASSI model. A major drawback of CLASSI is that it requires information about the same set of stimuli for all participants (i.e., crossed data), whereas recently a number of data gathering techniques have been proposed in which the set of stimuli differs across participants, yielding nested data. Therefore we present the CLASSI-N model, which extends the CLASSI model to nested data. A simulated annealing algorithm is proposed. The results of a simulation study are discussed as well as an application to data concerning depression.

Key words: sequential processes, CLASSI, individual differences, binary data, clusterwise regression, clustering.

## 1. Introduction

Many of the research questions facing psychological science are concerned with understanding how people respond to particular circumstances or stimuli (so-called S-R profiles). Such questions range from what determines consumer choice behavior in marketing research (Bigné, Mattila, & Andreu, 2008), what drives stress reactions following important life events in health psychology (Robinson, Garber, & Hilsman, 1995), examination of responses to certain tasks or test items in educational or assessment settings (Knuth, Stephens, McNeil, & Alibali, 2006), to studies of people's emotional reactions in everyday situations (Mischel & Shoda, 1998). The basic issue that researchers are trying to address in these domains often involves understanding the processes that drive such S-R profiles (why do people respond the way they do?), and how these processes may differ across individuals (e.g., across different types of consumer or demographic groups, across different types of patient groups, across students with different backgrounds, etc.).

To arrive at such an understanding, researchers often put forward a theoretical model in which a number of mediating variables (M) intervene between the stimuli and the responses. Such theories imply that the response comes about through a sequential S-M-R process, with the stimulus eliciting the mediating variables, which, in turn are associated with a specific response. For example, theories of emotion elicitation (Scherer, 1999) explain the experience of an emotion in a particular situation by postulating that cognitive evaluations of the situation, called appraisals, act as mediating variables (M) between the situations (S) and emotion (R) such that (1) people appraise their circumstances on a number of criteria, and (2) specific appraisal outcomes may elicit the experience of a particular emotion. When someone experiences anger, for instance, this could be because he or she appraised an event as frustrating and blamed someone

Requests for reprints should be sent to Eva Vande Gaer, Methodology of Educational Sciences Research Group and Research Group of Quantitative Psychology and Individual Differences, Katholieke Universiteit Leuven, Andreas Vesaliusstraat 2, P.O. Box 3762, 3000 Leuven, Belgium. E-mail: eva.vandegaer@ppw.kuleuven.be

else for it, and because these appraisals are associated with anger for that individual. As a second example, one may consider emotion regulation processes (Gross, 1998) aimed at increasing, maintaining, or decreasing affective responses (Davidson, Jackson, & Kalin, 2000; Gross, 2001). As such, emotion regulation implies the following sequential process: (1) a particular emotion eliciting event (S) activates regulation strategies (M), and (2) such strategies may increase, maintain or decrease the initial affective response (R). For instance, when a person becomes angry in a particular situation, he may try to regulate this feeling by hiding it from the outside world; however, such inhibition often leads to an even more intense internal feeling of anger. Moreover, many theories leave room for individual differences in the S-M and M-R links of the sequential process (see, e.g., Kuppens, Van Mechelen, & Rijmen, 2008 and Van Mechelen & Hennes, 2009 for individual differences in emotion elicitation, and Koole, 2009 for individual differences in emotion regulation).

Given data on the mediating variables and responses that people report for a number of stimuli, obtaining a precise picture of the underlying sequential process can be a complex and tedious task. Often, there are many different variables that possibly mediate between the stimuli and the response, and it is not always clear how they interrelate and differentially play a role in driving the sequential process. To complicate things further, both the S-M and M-R processes may be subject to individual differences. To deal with this complex task, it may be convenient to have at one's disposal a data analysis tool that reduces the mediating variables to a few mediating variable types and, at the same time, captures the individual differences in the link between stimuli and mediating variables and in the link between mediating variables and the response.

Recently, Ceulemans & Van Mechelen (2008) addressed these challenges by developing a data analysis technique called CLASSI. CLASSI reduces the number of stimuli and mediating variables by inducing a stimulus typology matrix and a mediating variable typology matrix from the data. Furthermore, CLASSI induces an S-M person typology and an M-R person typology from the data, with the S-M person types being characterized in terms of *if* <stimulus type>–*then* <mediating variable type> rules and the M-R person types in terms of *if* <mediating variable type>–*then* <response> rules. The S-M and M-R person types and their associated *if-then* rules provide a summary description of the individual differences in the S-M and M-R links of the sequential process under study.

A major drawback of CLASSI, however, is that it requires crossed data in that information about the same set of stimuli should be available for all participants. Yet, recently, a number of data collection methods, like experience sampling (Barrett & Barrett, 2001; Csikszentmihalyi & Larsen, 1987), were developed that gather data in real-life, in that people report on their responses to real-life stimuli. The resulting data are no longer crossed but nested since the set of stimuli differs across persons (i.e., stimuli are nested within persons). Therefore, in this paper we present a CLASSI method for nested data, called CLASSI-N. Because the stimuli under study differ across subjects, summarizing individual differences in the S-M link is not straightforward; therefore, CLASSI-N focusses on individual differences in the M-R link.

The remainder of the paper will be organized as follows. In Section 2 the CLASSI-N method is introduced. Section 3 covers the aim of and an algorithm for CLASSI-N analysis. In Section 4 the performance of the algorithm is evaluated by means of a simulation study. In Section 5, the CLASSI-N method is illustrated by applying it to data about emotion regulation in depression. In Section 6, we relate the CLASSI-N method to both the original CLASSI method and to other techniques.

TABLE 1.
Hypothetical data matrices $\mathbf{X}^M$ and $\mathbf{x}^R$.

| Person | Situation | $\mathbf{X}^M$: Mediating variables | | | | $\mathbf{x}^R$: Response |
|---|---|---|---|---|---|---|
| | | Threat to self-esteem | Other-blame | Injustice | Severe neg. consequences | Anger |
| 1 | Conflict with sister | 0 | 1 | 1 | 0 | 1 |
| | Sister moved far | 1 | 0 | 0 | 1 | 1 |
| | Child sick | 0 | 0 | 0 | 0 | 0 |
| | Child bad grades | 1 | 1 | 1 | 1 | 1 |
| | Conflict with boss | 1 | 0 | 0 | 1 | 1 |
| 2 | Conflict with partner | 1 | 0 | 0 | 1 | 1 |
| | Fail first exam | 1 | 0 | 0 | 1 | 1 |
| | Fail second exam | 1 | 1 | 1 | 1 | 1 |
| | Punished by parents | 0 | 0 | 0 | 0 | 0 |
| 3 | Bad news about health | 0 | 0 | 0 | 0 | 0 |
| | Receiving bills | 1 | 0 | 0 | 1 | 1 |
| | Conflict with child | 0 | 1 | 1 | 0 | 0 |
| | Argument with partner | 0 | 0 | 0 | 0 | 0 |
| 4 | Conflict with boss | 0 | 1 | 1 | 0 | 0 |
| | Fail job interview | 1 | 0 | 0 | 1 | 1 |
| | Missed train | 0 | 0 | 0 | 0 | 0 |
| 5 | Fell while walking | 0 | 0 | 0 | 0 | 0 |
| | Conflict with partner | 1 | 0 | 0 | 1 | 1 |
| | Got parking ticket | 0 | 1 | 1 | 0 | 0 |

## 2. The CLASSI-N Model

### 2.1. Data

CLASSI-N requires that for $K$ persons and for $I_k$ $(k = 1, \ldots, K)$ person-specific stimuli, information is available about the absence or presence of $J$ mediating variables and about the response under investigation. The data of person $k$ are represented in a binary $I_k \times J$ mediating variable data matrix $\mathbf{X}_k^M$ and a binary $I_k \times 1$ response data vector $\mathbf{x}_k^R$. Concatenating the data of all persons yields a binary $I \times J$ mediating variable data matrix $\mathbf{X}^M$ and a binary $I \times 1$ response data vector $\mathbf{x}^R$, with $I = \sum_{k=1}^{K} I_k$. In the remainder of this section, we will use the hypothetical data matrix $\mathbf{X}^M$ and data vector $\mathbf{x}^R$ in Table 1 as a guiding example. Table 1 shows for 19 person-situation combinations, whether or not the situation activates each of four appraisals in the person and whether or not the person feels angry in the situation.

### 2.2. Model

As stated in the introduction, studying sequential S-M-R processes can be a complex task because many mediating variables may be involved. Moreover, the *if* <mediating variable>–*then* <response> rules may be subject to individual differences. In this section, we discuss how CLASSI-N deals with both complicating factors by reducing the mediating variables to a few mediating variable types and by summarizing the individual differences in the *if* <mediating variable>–*then* <response> rules.

*2.2.1. Submodel for the Reconstruction of* $\mathbf{X}^M$*: Reduction of the Mediating Variables* To reduce the mediating variables to a few mediating variable types, CLASSI-N decomposes the $I \times J$ mediating variable data matrix $\mathbf{X}^M$ into a binary $J \times Q$ mediating variable typology matrix $\mathbf{M}$ and a binary $I \times Q$ mediator profile matrix $\mathbf{S}$. The mediating variable typology, which is represented in $\mathbf{M}$, is a partition implying that the different types are mutually exclusive and

TABLE 2.
(2, 2) CLASSI-N decomposition of $\mathbf{X}^M$ and $\mathbf{x}^R$ in Table 1.

Mediator profile matrix **S**

| Person | Situation | Mediating variable type $MT_1$ | Mediating variable type $MT_2$ |
|---|---|---|---|
| Person 1 | Conflict with sister | 0 | 1 |
| | Sister moved far | 1 | 0 |
| | Child sick | 0 | 0 |
| | Child bad grades | 1 | 1 |
| | Conflict with boss | 1 | 0 |
| Person 2 | Conflict with partner | 1 | 0 |
| | Fail first exam | 1 | 0 |
| | Fail second exam | 1 | 1 |
| | Punished by parents | 0 | 0 |
| Person 3 | Bad news about health | 0 | 0 |
| | Receiving bill | 1 | 0 |
| | Conflict with child | 0 | 1 |
| | Argument with partner | 0 | 0 |
| Person 4 | Conflict with boss | 0 | 1 |
| | Fail job interview | 1 | 0 |
| | Missed train | 0 | 0 |
| Person 5 | Fell while walking | 0 | 0 |
| | Conflict with partner | 1 | 0 |
| | Got parking ticket | 0 | 1 |

Mediating variable typology matrix **M**

| Mediating variable | Mediating variable type $MT_1$ | Mediating variable type $MT_2$ |
|---|---|---|
| Threat to self-esteem | 1 | 0 |
| Other-blame | 0 | 1 |
| Injustice | 0 | 1 |
| Severe neg. conseq. | 1 | 0 |

Person typology matrix **P**

| Person | Person type $PT_1$ | Person type $PT_2$ |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 1 | 0 |
| 3 | 0 | 1 |
| 4 | 0 | 1 |
| 5 | 0 | 1 |

Linking matrix **L**

| Mediating variable type | Person type $PT_1$ | Person type $PT_2$ |
|---|---|---|
| $MT_1$ | 1 | 1 |
| $MT_2$ | 1 | 0 |

nonempty. Therefore, the rows of **M** are restricted to sum to 1 and the columns to at least 1, with each entry indicating whether (1-entry) or not (0-entry) the corresponding mediating variable belongs to the type in question. For example, from Table 2, which shows a CLASSI-N solution with two mediating variable types and two person types for the data in Table 1, it can be derived that the appraisals of 'threat to self-esteem' and 'severe negative consequences' constitute the first mediating variable type $MT_1$, which can be labeled 'harm', and the appraisals of 'other-blame' and 'injustice' constitute the second mediating variable type $MT_2$, which can be labeled 'external attribution'.

The mediator profile matrix **S** specifies for each of the person-stimulus combinations which of the mediating variable types it activates, where a 1-entry (respectively, a 0-entry) indicates activation (respectively, no activation). For example, from the mediator profile matrix **S** in Table 2, it can be concluded that when Person 1's child came home with bad grades, this situation activated the harm appraisals in mediating variable type 1 ($MT_1$) as well as the external attribution appraisals in mediating variable type 2 ($MT_2$), whereas the situation in which the same child was sick, activated none of the appraisals. Note that the CLASSI-N method does not take into account that different persons may encounter very similar stimuli (for instance, both Persons 1 and 4 report on a situation 'conflict with boss'), in that CLASSI-N allows such similar stimuli to elicit different mediating variable types in these persons.

Given matrices **M** and **S**, the mediating variable data matrix $\mathbf{X}^M$ can be reconstructed by means of the following rule: A stimulus $i_k$ will activate mediating variable $j$ in person $k$ (i.e. $\hat{x}^M_{i_k j} = 1$) if and only if stimulus $i_k$ activates the mediating variable type $MT_q$ to which mediating variable $j$ belongs. For example, it can be deduced from Table 2 that Person 4 experienced a threat to his self-esteem when he failed in a job interview because the mediator profile matrix **S** indicates that this situation elicits mediating variable type $MT_1$ in Person 4, to which the appraisal of threat to self-esteem belongs. This rule can be formalized as follows:

$$x^M_{i_k j} \approx \hat{x}^M_{i_k j} = \bigoplus_{q=1}^{Q} s_{i_k q} m_{jq}, \tag{1}$$

where $m_{jq}$ indicates the entries of the mediating variable typology matrix **M** and $s_{i_k q}$ indicates the entries of the mediator profile matrix **S**. Note that $\oplus$ denotes the Boolean sum (i.e., $1 \oplus 1 = 1$).

*2.2.2. Submodel for the Reconstruction of* $\mathbf{x}^R$*: Individual Differences in the if <Mediating Variable Type>–then <Response> Rules* To capture individual differences in the *if* <mediating variable type>–*then* <response> rules, the CLASSI-N method induces a binary $K \times T$ person typology matrix **P** and a binary $Q \times T$ linking matrix **L**. The person typology matrix **P** represents the partition of the $K$ persons into $T$ person types. For example, from Table 2 it can be derived that Persons 1 and 2 constitute the first person type $PT_1$ and that Persons 3, 4, and 5 constitute the second person type $PT_2$.

The linking matrix **L** specifies the *if* <mediating variable type>–*then* <response> rules for each person type separately, where a 1-entry indicates that the corresponding mediating variable type elicits the response under investigation from the person type at hand and a zero-entry indicates that this is not the case. For example, from the linking matrix **L** in Table 2 it can be derived that both harm appraisals ($MT_1$) and external attribution appraisals ($MT_2$) are sufficient to elicit anger from persons belonging to $PT_1$. On **L** the restriction is imposed that the *if* <mediating variable type>–*then* <response> rules of the person types have to differ, implying that no column of **L** may equal another column of **L**.

Given **P** and **L**, the response data vector $\mathbf{x}^R$ can be reconstructed as follows: A stimulus $i_k$ will activate the response under investigation in person $k$ (i.e., $\hat{x}^R_{i_k} = 1$) if stimulus $i_k$ activates at least one mediating variable type in person $k$ that elicits the response under investigation from the person type to which $k$ belongs. For example, from Table 2 it can be read that Person 3 became angry when he received a bill, since this situation activated ‘harm’ appraisals ($MT_1$) in Person 3, which are sufficient to elicit anger from person type 2 ($PT_2$) to which Person 3 belongs. Similarly, it can be concluded from Table 2 that Person 3 did not feel angry when he had a conflict with his child, since this situation activated ‘external attribution’ appraisals ($MT_2$) which are not sufficient to elicit anger from $PT_2$ to which Person 3 belongs. This decomposition rule can be formalized as follows:

$$x^R_{i_k} \approx \hat{x}^R_{i_k} = \bigoplus_{q=1}^{Q} \bigoplus_{t=1}^{T} s_{i_k q} p_{kt} l_{qt}, \tag{2}$$

where $s_{i_k q}$ indicates the entries of the mediator profile matrix **S**, $p_{kt}$ indicates the entries of the person typology matrix **P**, and $l_{qt}$ indicates the entries of the linking matrix **L**.

*2.2.3. Graphical Representation* A CLASSI-N solution can be given a comprehensive graphical representation. For example, Figure 1 shows a graphical representation of the CLASSI-N solution in Table 2. This figure can be obtained by first displaying the mediating variable types by means of boxes in the middle of the figure. Next, person-stimulus combinations that activate the same mediating variable types are grouped together in boxes on the left side of the figure and arrows are drawn between the person-stimulus combination boxes and the mediating variable type boxes to indicate which mediating variable types are activated by each of the person-stimulus combinations. For instance, from Figure 1 it can be read that the situations ‘fail first exam’ (Person 2) and ‘fail job interview’ (Person 4) both activate ‘harm’ appraisals but not

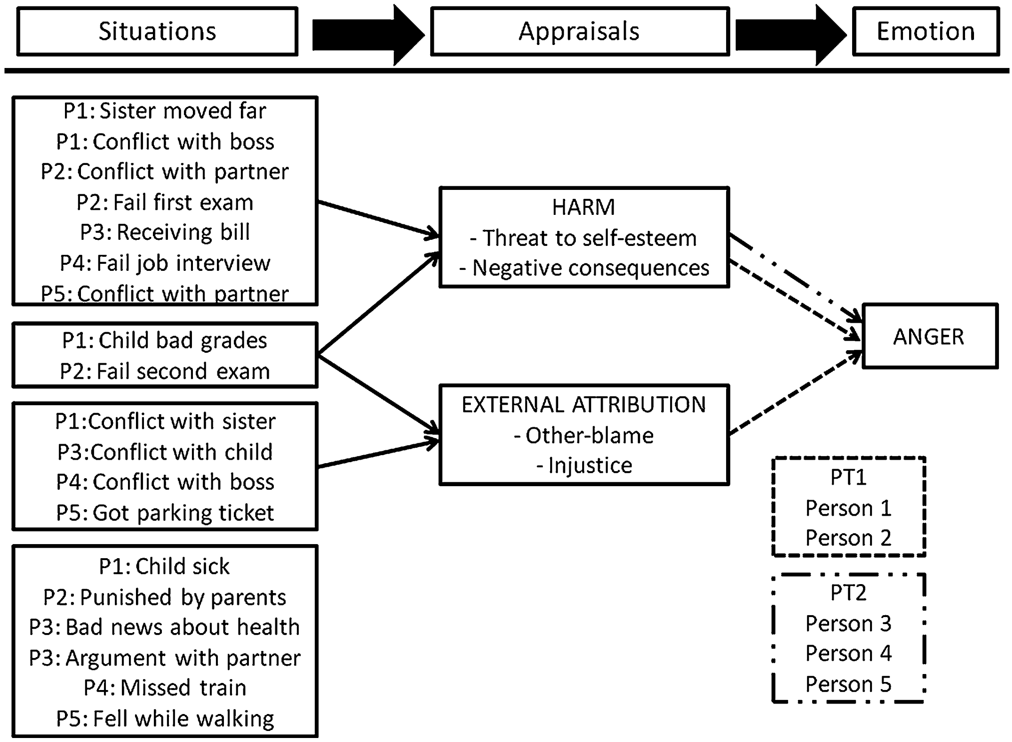


FIGURE 1.
Graphical representation of the $(Q, T) = (2, 2)$ CLASSI-N decomposition of $\mathbf{X}^M$ and $\mathbf{x}^R$ in Table 2.

'external attribution' appraisals. Next, in the lower right corner of the graphical representation, the person types are represented by a stack of boxes. Finally, the *if* <mediating variable type>–*then* <response> rules are indicated by interconnecting the mediating variable type boxes with the response under investigation, using different line styles for each person type. For instance, from Figure 1 it can be derived that for Person 1, who belongs to $PT1$, both mediating variable types are sufficient to elicit anger.

*2.2.4. Uniqueness* With respect to the uniqueness of a CLASSI-N decomposition of a reconstructed data matrix $\hat{\mathbf{X}}^M$ and reconstructed vector $\hat{\mathbf{x}}^R$, the following uniqueness theorem can be proved (see Appendix): If a $(Q, T)$ CLASSI-N decomposition of an $I \times J$ binary reconstructed data matrix $\hat{\mathbf{X}}^M$ and a $I \times 1$ binary reconstructed data vector $\hat{\mathbf{x}}^R$ exists, such that for all persons $k$ $(k = 1, \dots, K)$ and all mediating variable types $MT_q$ $(q = 1, \dots, Q)$ at least one stimulus $i_k$ exists that activates this mediating variable type only (i.e., $s_{i_k q'} = 1$ iff $q = q'$), this decomposition is unique upon a permutation of the mediating variable types and the person types.

More specifically, this theorem implies that when for a particular person information is missing on the separate effect of one or more mediating variables types—i.e., a mediating variable type never occurs or always occurs in combination with other mediating variable types—this person can sometimes be assigned to multiple person types without altering the reconstructed data. For example, if two person types only differ with respect to the effect of a particular mediating variable type, a person for whom this latter type never occurs may be assigned to either of these person types without changing $\hat{\mathbf{x}}^R$. Moreover, in case a mediating variable type never occurs or is always activated together with other mediating variable types for all persons that belong to a particular person type, it is also possible that two or more different linking matrices $\mathbf{L}$ lead to the same reconstructed data vector $\hat{\mathbf{x}}^R$, implying that the linking matrix is not unique. However, we never encountered a non-unique linking matrix in practice, neither in the simulation study nor in the analysis of several real data sets. We conclude that in case CLASSI-N solutions are non-unique, this almost always pertains to specific parts of the person typology matrix **P** without affecting **S**, **M** and **L**. Therefore, it may be more fair to refer to this phenomenon as partial non-uniqueness. In the application in Section 5, this partial non-uniqueness will be further illustrated, showing that it is not problematic.

## 3. Data Analysis

### 3.1. Aim

Given a binary $I \times J$ mediating variable matrix $\mathbf{X}^M$ and a binary $I \times 1$ response vector $\mathbf{x}^R$ and given complexity $(Q, T)$, the aim of a CLASSI-N analysis is to look for the reconstructed binary $I \times J$ mediating variable matrix $\hat{\mathbf{X}}^M$ and the reconstructed binary $I \times 1$ response vector $\hat{\mathbf{x}}^R$ that have a minimal value on the following least squares loss function:

$$L = \sum_{k=1}^{K} \sum_{i_k=1}^{I_k} \sum_{j=1}^{J} \left(\hat{x}_{i_k j}^M - x_{i_k j}^M\right)^2 + \sum_{k=1}^{K} \sum_{i_k=1}^{I_k} \left(\hat{x}_{i_k}^R - x_{i_k}^R\right)^2 \qquad (3)$$

and that can be represented by a CLASSI-N solution of the specified complexity. However, the complexity $(Q, T)$ of the CLASSI-N solution that best describes a given data set without being overly complex is often unknown. As a way out, one will usually fit CLASSI-N solutions of different complexities to these data. Note that solutions with $T > 2^Q$ can be discarded, because in such solutions some columns of **L** will be identical, implying that a solution with one less person type exists that fits the data equally well. Having obtained CLASSI-N solutions of different complexities, a final solution may be selected on the basis of formal model selection heuristics (which weigh the fit of the different solutions against the complexity), the interpretability of the solutions, and their stability.

### 3.2. Algorithm

In this subsection, we propose a simulated annealing (SA) algorithm for fitting a $(Q, T)$ CLASSI-N solution to a given data matrix $\mathbf{X}^M$ and data vector $\mathbf{x}^R$ and thus for estimating a mediator profile matrix **S**, a mediating variable typology matrix **M**, a person typology matrix **P**, and a linking matrix **L** of the prespecified complexity. Simulated annealing (for a general introduction, see Aarts, Korst, & Michiels, 2005) is a local search technique that is often used for complex combinatorial optimization problems (Brusco, 2001; Trejos & Castillo, 2000; Wilderjans, Ceulemans, & Van Mechelen, 2008). Making use of the pseudo code in Algorithm 1 and the associated notation in Table 3, we first recapitulate the general principle of the SA algorithm and subsequently discuss the specific implementation for fitting CLASSI-N solutions.

Given an initial solution to the problem at hand, the current solution $S_{\text{current}}$ with an associated loss function value $L_{\text{current}}$, SA generates a sequence of trial solutions $S_{\text{trial}}$ with associated loss function values $L_{\text{trial}}$. These trial solutions are constructed by changing one or a few randomly selected parameter values of $S_{\text{current}}$. If the trial solution fits the data better than the current solution ($L_{\text{trial}} < L_{\text{current}}$), $S_{\text{trial}}$ is always accepted and $S_{\text{current}}$ is replaced by $S_{\text{trial}}$. However, since only accepting better fitting solutions would make the algorithm susceptible to

**Algorithm 1** The SA algorithm for CLASSI-N analysis.

$\alpha = 0.90$;
$L_{\text{best}} := IJ + I$;
initialize $T_{\text{current}}$;
initialize $S_{\text{current}}$ and associated $L_{\text{current}}$;
**repeat**
  $i_{\text{gen}} := 0$;
  $i_{\text{acc}} := 0$;
  **while** $(i_{\text{gen}} < (IQ + JQ + KT + QT))$ and $(i_{\text{acc}} < 0.1 * (IQ + JQ + KT + QT))$ **do**
    $i_{\text{gen}} := i_{\text{gen}} + 1$;
    generate $S_{\text{trial}}$ by changing random parameter value;
    **if M** was changed **then**
      update **S**;
    **end if**
    update $L_{\text{trial}}$;
    draw $h$ from $U(0, 1)$
    **if** $(L_{\text{trial}} < L_{\text{current}})$ or $(h < \exp(\frac{L_{\text{current}} - L_{\text{trial}}}{T_{\text{current}}}))$ **then**
      **if** $L_{\text{trial}} < L_{\text{best}}$ **then**
        $S_{\text{best}} := S_{\text{trial}}$;
        $L_{\text{best}} := L_{\text{trial}}$
      **end if**
      $S_{\text{current}} := S_{\text{trial}}$;
      $L_{\text{current}} := L_{\text{trial}}$;
      $i_{\text{acc}} := i_{\text{acc}} + 1$
    **end if**
  **end while**
  $T_{\text{current}} := \alpha * T_{\text{current}}$;
**until** $(T_{\text{current}} \leq 0.000001)$ or $(i_{\text{acc}} = 0)$;
return $S_{\text{best}}$;

local minima, sometimes also worse fitting trial solutions ($L_{\text{trial}} \geq L_{\text{current}}$) are accepted. More specifically, these worse fitting trial solutions are accepted with probability:

$$p_{\text{acc}} = \exp\left(\frac{L_{\text{current}} - L_{\text{trial}}}{T_{\text{current}}}\right), \tag{4}$$

where $T_{\text{current}}$ indicates the temperature of the algorithmic process (the term temperature stems from the algorithm's analogy to metallurgical cooling processes), with $T_{\text{current}} > 0$. $T_{\text{current}}$ is initially set high, leading to the acceptance of many worse fitting trial solutions. During the algorithm, each time a prespecified cooling criterion is met, $T_{\text{current}}$ is decreased. This cooling is usually done by multiplying $T_{\text{current}}$ by a cooling factor $\alpha$ $(0 < \alpha < 1)$. The lower $T_{\text{current}}$ becomes, the lower the probability that a worse fitting trial solution will be accepted. In the end, $p_{\text{acc}}$ approaches 0. As a consequence, no more worse trial solutions are accepted. Thus, the algorithm converges on a specific solution and the best obtained solution is returned.

To estimate CLASSI-N solutions, the SA algorithm is implemented as follows: First, $S_{\text{current}}$ is initialized randomly. More specifically, the mediating variable and person typology matrices (**M** and **P**) are generated by drawing type memberships from a multinomial distribution with the probabilities of each type set to 1 divided by the number of types, subject to the restriction that each type contains at least one element. The entries of the linking matrix **L** are independent

TABLE 3.
Notation for the SA algorithm for CLASSI-N analysis.

| Label | Indicates |
|---|---|
| $S_{\text{current}}, S_{\text{trial}}, S_{\text{best}}$ | The current solution, the trial solution, and the best encountered CLASSI-N solution |
| $L_{\text{current}}, L_{\text{trial}}, L_{\text{best}}$ | The loss function value of the current, trial, and best encountered CLASSI-N solution |
| $T_{\text{current}}$ | The current temperature |
| $\alpha$ | The cooling factor by which $T_{\text{current}}$ is multiplied to reduce the temperature, $0 < \alpha < 1$ |
| $i_{\text{gen}}$ | The number of trial solutions that have already been generated at the current temperature |
| $i_{\text{acc}}$ | The number of trial solutions that have already been accepted at the current temperature |

realizations of a Bernouilli variable with parameter value 0.5, subject to the restriction that no column of **L** equals another column. Next, the mediator profile matrix **S** is estimated conditionally upon the current **M**, **P**, and **L** matrices by means of greedy Boolean regression (Leenen & Van Mechelen, 2001). Note that due to a separability property of the loss function value (Chaturvedi & Carroll, 1994), the mediator profile of each person-stimulus combination can be estimated separately.

Second, trial solutions are generated by altering the type membership of one randomly chosen mediating variable in **M** or person in **P** or the value of one randomly chosen entry of the linking matrix **L**, with all mediating variable and person type memberships and all entries of the linking matrix **L** having an equal chance of being altered. When altering the type membership of a mediating variable or person, the restriction is imposed that all types must contain at least one element. When altering the *if* <mediating variable>–*then* <response> rules, it is required that no column of **L** equals another column. Because of the relatively large size of the mediator profile matrix **S**, estimating its entries by randomly altering one entry at a time would make estimating CLASSI-N solutions very intensive if not infeasible. Therefore, we opted to update **S** by means of greedy Boolean regression each time the mediating variable matrix **M** is altered. Note that updating **S** each time **P**, **M**, or **L** are altered, would also be too computationally intensive since the number of persons and thus **P** will often be large.

Third, the initial $T_{\text{current}}$ was chosen so as to result in an average acceptance probability of 0.80 (Aarts et al., 2005). Fourth, on the basis of some pilot studies, the cooling factor $\alpha$ was set to 0.90 (typical values of $\alpha$ lie between 0.8 and 0.99, Aarts et al., 2005; Trejos & Castillo, 2000). Fifth, $T_{\text{current}}$ is cooled down when either $IQ + JQ + KT + QT$ trial solutions have been generated or $0.1 * (IQ + JQ + KT + QT)$ trial solutions have been accepted (see e.g. Brusco, 2001; Kirkpatrick, Gelatt, & Vecchi, 1983), which implies that the number of trial solutions depends on the size of the data set and the complexity $(Q, T)$ of the obtained CLASSI-N solution. Moreover, at lower temperatures, when the SA algorithm is generally exploring more interesting subsets of solutions (Kirkpatrick et al., 1983), more trial solutions are generated and evaluated. Sixth, the algorithm stops when $T_{\text{current}} \leq 0.000001$ or when no $S_{\text{trial}}$ has been accepted at a specific value of $T_{\text{current}}$ (Kirkpatrick et al., 1983; Trejos & Castillo, 2000).

Finally, in order to avoid local minima, we opt to run the CLASSI-N algorithm 25 times, where each run starts from a different random initialization of $S_{\text{current}}$. From the 25 resulting CLASSI-N solutions, only the solution with the lowest loss function value is retained. On the obtained solution a postprocessing procedure can be applied to identify the persons for whom not enough information is available to uniquely assign them to a person type. Specifically, the postprocessing procedure indicates to which person types these persons can be assigned without altering the reconstructed data vector $\hat{\mathbf{x}}^R$. Moreover, this procedure also flags all entries of the linking matrix $\mathbf{L}$ that can be changed in value without affecting $\hat{\mathbf{x}}^R$.

## 4. Simulation Study

In this section, we will discuss the results of a simulation study that was conducted in order to evaluate the performance of the CLASSI-N algorithm. Specifically, to evaluate how well the algorithm succeeds in minimizing the loss function (goodness-of-fit), it was examined how often the CLASSI-N algorithm returned a local minimum. Furthermore, it was assessed how well the algorithm succeeds in recovering the underlying truth (goodness-of-recovery) both on the level of the entries of the reconstructed data sets and on the level of the underlying typology, mediator profile and linking matrices.

First, the design and procedure of the simulation study will be described. Next, the results concerning local minima will be discussed, followed by the goodness-of-recovery results.

### *4.1. Design and Procedure*

The evaluation of the algorithm involves three different pairs of an $I \times J$ binary matrix and a $I \times 1$ binary vector: true matrices and vectors, data matrices and vectors, and reconstructed data matrices and vectors. The true matrices $\mathbf{T}^M$ and true vectors $\mathbf{t}^R$, which can be perfectly represented by a $(Q, T)$ CLASSI-N solution, were generated as follows: The typology matrices $\mathbf{M}$ and $\mathbf{P}$ were obtained by randomly assigning each element to a type where all types had equal probability of being assigned to, subject to the restriction that all types must contain at least one element. The entries of the linking matrix $\mathbf{L}$ and the mediator profile matrix $\mathbf{S}$ were independent realizations of a Bernouilli variable with a probability parameter of 0.5, subject to the restriction that $\mathbf{L}$ contains no identical columns. Subsequently, combining the $\mathbf{P}$, $\mathbf{M}$, $\mathbf{S}$ and $\mathbf{L}$ matrices by means of the CLASSI-N reconstruction rules (Equations (1) and (2)) yielded true matrices $\mathbf{T}^M$ and true vectors $\mathbf{t}^R$. Next, data matrices $\mathbf{X}^M$ and data vectors $\mathbf{x}^R$ were obtained by altering the value of $\varepsilon \times 100$ randomly chosen entries of $\mathbf{T}^M$ and $\mathbf{t}^R$, where $\varepsilon$ indicates the error level (see below). Finally, reconstructed data matrices and vectors $\hat{\mathbf{X}}^M$ and $\hat{\mathbf{x}}^R$ were acquired by analyzing the data matrices $\mathbf{X}^M$ and vectors $\mathbf{x}^R$ with the CLASSI-N algorithm in complexity $(Q, T)$.

The simulation study had a trifactorial design. More specifically, the following three factors were manipulated.

(a) The *Size*, $(I, J)$; i.e. the number of person-stimulus combinations $I = \sum_{k=1}^{K} I_k$ was varied at three levels: $K = 25$ and $I_k \sim bin(100, 0.50)$, $K = 50$ and $I_k \sim bin(150, 0.50)$, and $K = 100$ and $I_k \sim bin(200, 0.50)$, while the number of mediating variables $J$ was varied at two levels: $J = 12$ and $J = 24$. Thus, fully crossing the three levels for $I$ and the two levels for $J$, this factor consisted of six different levels.
(b) The *True complexity*, $(Q, T)$, of the CLASSI-N solution for $\mathbf{T}^M$ and $\mathbf{t}^R$ was varied at nine levels: (2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4), (4, 2), (4, 3), (4, 4).
(c) The *Error level*, $\varepsilon$, which is the proportion of entries $x^M_{i_k j}$ (resp. $x^R_{i_k}$) differing from $t^M_{i_k j}$ (resp. $t^R_{i_k}$), varied at three levels: 0.00, 0.10, 0.20.

Five data sets were generated for each combination of Size, Complexity and Error level, yielding 6 (Size) $\times$ 9 (Complexity) $\times$ 3 (Error level) $\times$ 5 (replicates) = 810 simulated data sets. To decrease the probability of ending up in a local minimum, we ran the CLASSI-N algorithm 25 times in the true complexity where each run started from a different random initialization of the current solution. From the 25 resulting CLASSI-N solutions, the solution with the lowest loss function value (Equation (3)) was retained. The average CPU time per run of the algorithm was approximately 308.71 seconds. However, for the smallest data sets CPU time per run was only a few seconds while for the largest, most complex datasets CPU times for one run could exceed one hour. In total, the CPU time for the whole simulation study amounted to approximately $6.24 \times 10^6$ seconds.

### 4.2. *Local Minima*

In this subsection, we examine whether or not the CLASSI-N algorithm tends to end in a local minimum. This question is not a simple one to answer since in all cases in which the simulated data matrices $\mathbf{X}^M$ and vectors $\mathbf{x}^R$ are obtained by perturbing $\mathbf{T}^M$ and $\mathbf{t}^R$ with nonzero random error, the global minimum for the CLASSI-N analysis of $\mathbf{X}^M$ and $\mathbf{x}^R$ is unknown. As a way out, we compare the loss function value $L$ (Equation (3)) of the obtained CLASSI-N solution to the badness-of-data (*BOD*) value

$$BOD = \sum_{k=1}^{K} \sum_{i_k=1}^{I_K} \sum_{j=1}^{J} \left(x_{i_k j}^M - t_{i_k j}^M\right)^2 + \sum_{k=1}^{K} \sum_{i_k=1}^{I_K} \left(x_{i_k}^R - t_{i_k}^R\right)^2. \tag{5}$$

This *BOD*-value can be considered an upper bound for the loss function value $L$ of the global minimum of the CLASSI-N analysis with complexity $(Q, T)$. Therefore, if $L > BOD$, we are certain the algorithm yielded a local minimum. However, $L \leq BOD$ does not imply that the algorithm found the global minimum, since there may exist other $\hat{\mathbf{X}}^M$ and $\hat{\mathbf{X}}^R$ pairs that are closer to $\mathbf{X}^M$ and $\mathbf{x}^R$ than the obtained solution.

Comparing the loss function value $L$ of the best obtained CLASSI-N solution across 25 runs of the algorithm with the *BOD* value, revealed that $L$ exceeded the *BOD* value for 63 of the 810 data sets (8%), that $L$ was equal to the *BOD* value in 278 of the 810 data sets (34%), and that $L$ was smaller than the *BOD* value for the other 469 data sets (58%). Although the CLASSI-N algorithm ended for sure in a local minimum for 8% of the data sets, the obtained solutions for these 63 data sets still fitted the data reasonably well:

$$\frac{L - BOD}{IJ + I} \tag{6}$$

equaled on average 0.002, implying that the obtained solutions yielded only 0.2% more erroneously reconstructed data entries than the true solutions. These results suggest that the CLASSI-N algorithm succeeds rather well in minimizing the loss function.

To investigate further the issue of local minima, we examined how many of the 25 runs per simulated data set ended in a solution with a loss function value equal to the loss function value of the best obtained solution for that data set. On average, this was the case for 8.00 of the 25 runs ($SD = 10.22$), with the average standard error of the mean (*SEM*) per cell of the design amounting to 1.08. An analysis of variance with Size, Complexity, and Error level as independent variables and the number of runs as dependent variable yielded an intraclass correlation $\hat{\rho}_I$ (Haggard, 1958; Kirk, 1995) of 0.74 for the main effect of Error level. Specifically, the more error the data contained, the lower the number of runs that ended in a solution with a loss function value equal to that of the retained solution. Note that in this and the following analyses of variance only effects accounting for at least 10% of the variance of the dependent variable are considered (i.e., $\hat{\rho}_I \geq 0.10$). We conclude that 25 runs is usually sufficient to obtain a good solution when not much error is present in the data. However, when one expects the amount of error in the data to be large, one could consider increasing the number of runs.

TABLE 4.
Mean Badness-of-Recovery values at different levels of Size × Error.

| Size | Error level | | | |
|---|---|---|---|---|
| | 0.00 | 0.10 | 0.20 | Overall |
| $K = 25$, $J = 12$, $E(I_k) = 50$ | 0.0000 | 0.0242 | 0.0846 | 0.0363 |
| $K = 50$, $J = 12$, $E(I_k) = 75$ | 0.0000 | 0.0251 | 0.0849 | 0.0367 |
| $K = 100$, $J = 12$, $E(I_k) = 100$ | 0.0001 | 0.0249 | 0.0877 | 0.0376 |
| $K = 25$, $J = 24$, $E(I_k) = 50$ | 0.0001 | 0.0046 | 0.0309 | 0.0119 |
| $K = 50$, $J = 24$, $E(I_k) = 75$ | 0.0001 | 0.0043 | 0.0314 | 0.0120 |
| $K = 100$, $J = 24$, $E(I_k) = 100$ | 0.0005 | 0.0052 | 0.0316 | 0.0124 |
| Overall | 0.0002 | 0.0147 | 0.0585 | 0.0245 |

### 4.3. *Goodness-of-Recovery*

The goodness-of-recovery of the 810 obtained CLASSI-N solutions will be evaluated both on the level of the entries of the true matrices $\mathbf{T}^M$ and vectors $\mathbf{t}^R$ as well as on the level of the underlying typologies, mediator profiles and linking matrices.

*4.3.1. Recovery of* $\boldsymbol{T}^M$ *and* $\boldsymbol{t}^R$ To evaluate the recovery of the entries of the true matrices and true vectors, we computed a badness-of-recovery (*BOR*) value based on the proportion of discrepancies between the reconstructed data matrices and vectors, $\hat{\mathbf{X}}^M$ and $\hat{\mathbf{X}}^R$, and the corresponding true matrices and vectors, $\mathbf{T}^M$ and $\mathbf{t}^R$:

$$BOR = \frac{\sum_{k=1}^{K}\sum_{i_k=1}^{I_K}\sum_{j=1}^{J}(\hat{x}_{i_k j}^M - t_{i_k j}^M)^2 + \sum_{k=1}^{K}\sum_{i_k=1}^{I_K}(\hat{x}_{i_k}^R - t_{i_k}^R)^2}{IJ + I}. \tag{7}$$

First of all, the *BOR* values were quite stable within the cells of the design, as on average, the standard error per cell amounted to 0.00. For 264 of the 810 data sets, the *BOR* value exactly equaled 0, indicating that the true matrices were reconstructed perfectly. For these data sets, we checked whether the obtained solution was unique. Note that the other data sets cannot be used for studying uniqueness, because differences in model parameters can either be due to non-uniqueness, or to the fact that the reconstructed matrices differ from the true matrices. For only one of these 264 data sets could one person be switched from one person type to another without altering the reconstructed data vector $\mathbf{x}^R$. This was due to the fact that for the person in question one of the mediating variable types always occurred together with a second mediating variable type, implying that the effect of the first mediating variable type could not be evaluated properly. The mean *BOR* value of the other 546 data sets equaled 0.04 ($SD = 0.03$), which means that the reconstructed data matrices yielded by the algorithm differed on average 4% from the underlying true matrices. Taking into account that given the goodness-of-fit results, at most 278 data sets could be reconstructed perfectly, these results imply that the CLASSI-N algorithm succeeds quite well in recovering the underlying truth.

An analysis of variance with *BOR* as the dependent variable and Size, Complexity, and Error level as independent variables yielded an intraclass correlation of 0.61 for the main effect of Error level: The more error, the worse the recovery (see Table 4). Also, an interaction effect between Size and Error level was found ($\hat{\rho}_I = 0.16$), indicating that the effect of Error level becomes less pronounced when the number of mediating variables increases.

*4.3.2. Recovery of Typologies, Mediator Profiles and Linking Matrices* To evaluate the recovery of the underlying true typology matrices and true mediator profile matrices, the corrected Rand index (Hubert & Arabie, 1985) was used. This index equals 1 if the mediator profiles and

TABLE 5.
Mean corrected Rand Index for the mediator profile matrices **S** at different levels of Complexity × Error.

| | | Error level | | | |
|---|---|---|---|---|---|
| | Complexity | 0.00 | 0.10 | 0.20 | Overall |
| Mediator profile matrices **S** | (2, 2) | 1.000 | 0.974 | 0.813 | 0.929 |
| | (2, 3) | 1.000 | 0.961 | 0.815 | 0.925 |
| | (2, 4) | 1.000 | 0.975 | 0.824 | 0.933 |
| | (3, 2) | 1.000 | 0.873 | 0.634 | 0.836 |
| | (3, 3) | 1.000 | 0.891 | 0.621 | 0.838 |
| | (3, 4) | 0.994 | 0.862 | 0.622 | 0.826 |
| | (4, 2) | 1.000 | 0.749 | 0.447 | 0.732 |
| | (4, 3) | 1.000 | 0.757 | 0.436 | 0.731 |
| | (4, 4) | 0.998 | 0.762 | 0.433 | 0.731 |
| | Overall | 0.999 | 0.867 | 0.627 | 0.831 |

typology matrices are recovered perfectly and 0 when the true and reconstructed typologies and mediator profiles do not correspond more than expected by chance. Because the corrected Rand index can only be used for comparing partitions, the mediator profile matrices were reformulated into partition matrices that assign the person-stimulus combinations to $2^Q$ partition classes, where each partition class contains all the person-stimulus combinations that activate a particular subset of mediating variable types.

For the mediating variable typologies, the corrected Rand indices always equaled 1. This implies that for all data sets the mediating variable typology matrix **M** was reconstructed perfectly.

With a mean corrected Rand index of 0.83 ($SD = 0.21$, average $SEM = 0.01$), the mediator profile matrix **S** was in general reconstructed reasonably well, implying a moderate to good recovery of the 'true' mediator profiles (Steinley, 2004). An analysis of variance with the corrected Rand index for the mediator profile matrix **S** as dependent variable and Size, Complexity, and Error level as independent variables yielded a main effect of Error level ($\hat{\rho}_I = 0.59$), indicating that higher Error levels lead to lower Rand indices (see Table 5). Additionally, an interaction effect of Complexity and Error level was found ($\hat{\rho}_I = 0.12$), implying that the effect of Error level becomes more pronounced when the number of mediating variable types increases.

With respect to the recovery of the person typology matrix **P**, the mean corrected Rand index equaled 0.95 ($SD = 0.13$, average $SEM = 0.03$), indicating a very high correspondence between the true and reconstructed person type matrices. An analysis of variance revealed a main effect of Error level ($\hat{\rho}_I = 0.21$): The more error, the lower the corrected Rand index. Specifically, the mean corrected Rand indices for 0%, 10%, and 20% error amounted to 1.00, 0.98, and 0.87, respectively.

To evaluate the recovery of the linking matrix **L**, we computed the proportion of discrepancies between the true and reconstructed linking matrices:

$$\sum_{q=1}^{Q}\sum_{t=1}^{T}\frac{(l_{qt}^{T}-l_{qt}^{R})^2}{QT}. \tag{8}$$

As the person types and mediating variable types of CLASSI-N solution can be permuted without altering the loss function value (3), we permuted the types so as to minimize (8). For only 38 out of the 810 CLASSI-N data sets, the obtained proportion of discrepancies exceeded 0, indicating that the linking matrix was recovered very well. An analysis of variance with (8) as dependent variable and Size, Rank, and Error as independent variables, revealed that none of the effects explained more than 10% of the variance ($\hat{\rho}_I < 0.10$).

## 5. Illustrative Application

In this section, we illustrate the CLASSI-N method by applying it to data from a study on emotion regulation. Emotion regulation refers to any process or strategy that is aimed at increasing, maintaining, or decreasing affective responses (Davidson et al., 2000; Gross, 2001). As such, emotion regulation can be considered a sequential S-M-R process in that a particular emotion eliciting event (S) is said to elicit one or several regulative processes (M) which in turn can change or maintain the initial emotional response (R). These two links of this emotion regulation process are subject to individual differences. First, people may differ in the extent that they use particular emotion regulation strategies (Gross, 2008; Gross & John, 2003). Different emotion regulation strategies have different ease of use and some may be more preferred than others by certain individuals (Koole, 2009). Second, individuals also may differ in their ability to use particular regulation strategies effectively (Hemenover, Augustine, Shulman, Tran, & Barlett, 2008). Different emotion regulation strategies involve different mechanisms that may be facilitated or counteracted by specific individual characteristics (Koole, 2009; O'Brien & DeLongis, 1996). Certain emotion regulation strategies may thus be more fitting and suitable for some persons and less for others, resulting in better or, respectively, worse outcomes (Hemenover et al., 2008).

In this application, the focus is on individual differences in the emotion of depression. With respect to depression, emotion regulation research has mainly focused on the role of turning one's attention to one's problems and thoughts, usually referred to as rumination, as this emotion regulation strategy seems to play a crucial role in the development and maintenance of depressed mood (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). This emotion strategy is often contrasted with turning one's attention away from one's feelings and thoughts (avoidance), which is thought to have differential effects, at least depending on the individual (Koole, 2009).

The data were collected as follows: 31 persons carried a palmtop computer for seven days. During each day, this palmtop computer beeped at 10 randomly selected moments, prompting the participants to answer several questions about their cognitions and actions during the period since the previous beep and about their feelings at that moment. Specifically, the participants rated by means of a six-level Likert-type scale (ranging from 0 = not at all to 5 = very much so) to what extent they felt depressed at the moment of the beep and to what extent, since the previous beep, they had engaged in each of the following activities: constantly thinking about their feelings, reflecting on their feelings, talking about their feelings, avoid thinking about their feelings, changing their way of thinking about the cause of their feelings (= reappraising), avoid expressing their feelings, focusing on their feelings, and distracting from their feelings by means of activities. These variables were dichotomized by recoding 0 and 1 to zero and 2 to 5 to one. Moreover, the participants also indicated with a yes/no score whether they had met and talked with other people, did something that they wanted to do, and did something that they had to do. As such, this study resulted in a data matrix consisting of 1538 observation periods $\times$ 12 mediating variables and an associated depression data vector $\mathbf{x}^R$ with the 1538 observation periods (31 persons $\times$ 10 observations per day $\times$ 7 days—missing observations) being nested in the participants. Note that in this application observation periods are considered stimuli as in each observation period an event may have occurred that induced feelings of depression.

In order to obtain a summary description of the individual differences in the associations between the regulatory actions and cognitions and the feelings of depression, CLASSI-N models of complexity $(1, 1)$ to $(6, 6)$ were fitted to the mediating variable data matrix $\mathbf{X}^M$ and response

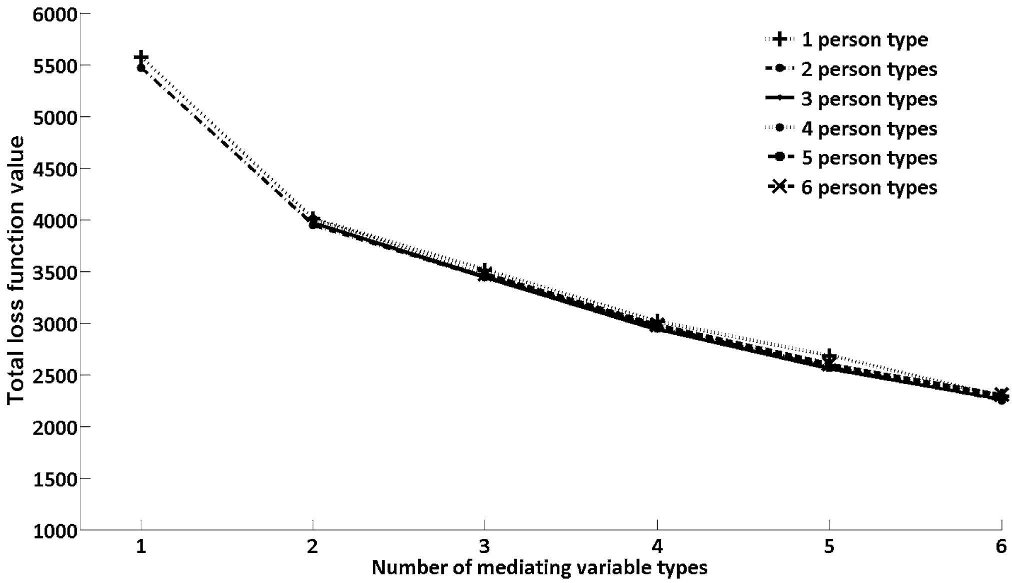


FIGURE 2.
Loss function values for the depression data as a function of the number of mediating variable types $Q$ and the number of person types $T$.

data vector $\mathbf{x}^R$ using 25 randomly started runs. To establish the number of mediating variable types $Q$, the loss function values (3) of the obtained solutions were compared by drawing a separate scree line for each value of $T$. As can be seen in Figure 2, the scree lines show no obvious elbow except for $Q = 2$. Since the solutions with 2 mediating variable types only separated the action related variables from the cognitive variables, they were not very informative. Therefore, we decided to retain four mediating variable types as the resulting mediating variable types were well interpretable.

To subsequently decide on the number of person types $T$, only the loss function value for the response vector $\hat{\mathbf{X}}^R$:

$$L^R = \sum_{k=1}^{K} \sum_{i_k=1}^{I_k} \left(\hat{x}_{i_k}^R - x_{i_k}^R\right)^2, \tag{9}$$

was taken into account since the number of person types has no influence on the reconstruction of the mediating variable data matrix $\mathbf{X}^M$. In Figure 3, which shows a scree plot for the number of person types, it can be seen that the decrease in $L^R$ levels off at $T = 2$ and $T = 3$.

Because a CLASSI-N solution with three person types allows for a more detailed investigation of the interindividual differences in the link between the regulation strategies and the feeling of depression than the solution with two person types, we selected the (4, 3) CLASSI-N solution for further investigation. This solution is graphically represented in Figure 4.

With respect to the reduction of the regulatory cognitions and actions, it can be read from Figure 4 that 'thinking and reflecting on feelings' and 'reappraising and focusing on feelings and problems' are combined together into one mediating variable type ($MT_1$), which we will label 'rumination'. 'Talking about feelings' forms a type on its own ($MT_2$). The third mediating variable type ($MT_3$) consists of all cognitions and actions that pertain to avoiding the problem and the related feelings (avoid thinking about feelings, avoid expressing feelings, distract from problem) and thus will be labeled 'avoidance'. Note that this partitioning of the mediating variables into a rumination type and an avoidance type links up with the 'rumination' versus 'avoidance'

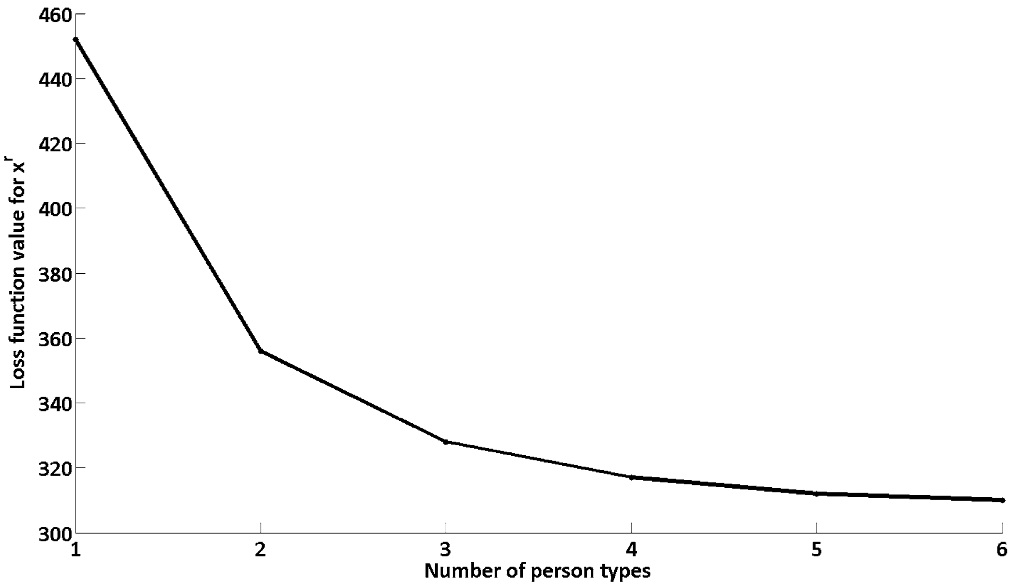


FIGURE 3.
Number of person types $T$ by response loss function values for the depression data, given $Q = 4$.

division found in the literature (Koole, 2009; Nolen-Hoeksema et al., 2008). Whereas these first three mediating variable types contain emotion related cognitions and actions, the fourth mediating variable type ($MT_4$) consists of actions that have no direct relation to the problem (talk to people, I did something I wanted to do, and I did something I had to do). We will label this mediating variable type ‘activity not related to the problem’.

Figure 4 shows how many of the 1538 observations periods activated each of the 15 possible combinations of the four mediating variable types. Note that the 97 observations that activated none of the mediating variable types are not included in the figure. In general, one can conclude that some combinations occurred very frequently, while others were rarely observed. Only discussing the combinations that were elicited for more than 10% of the observations, we first note that for 363 observations (24%) only activities not related to the problem ($MT_4$) were reported. This suggests that these observations took place in emotionally neutral periods. Second, in 216 observations (14%) rumination ($MT_1$), avoidance ($MT_3$) as well as activities not related to the problem were reported ($MT_4$). Finally, in 227 observations (15%) all the cognitions and actions under study occurred. It is important to note that except for activities not related to the problem, the mediating variable types occurred only seldom without one or more of the other mediating variables types. These results imply that it is very likely that the obtained CLASSI-N solution will only be partially unique, as for some persons not enough information will be available about the sufficiency of the cognitions and actions for feeling depressed to discard some of the person types.

With respect to individual differences in the *if* <regulatory action>–*then* depression rules, the retained CLASSI-N solution distinguishes between three person types. For the ten persons belonging to the first person type ($PT_1$) both rumination and avoidance are sufficient to report a feeling of depression later on. Secondly, the ten persons belonging to the second person type only feel depressed if they ruminated on the event. Finally, for the ten persons belonging to the third person type, none of the obtained types of regulatory strategy is related to depression. Note that the latter ten persons report depression at a limited number of beeps only. As expected, the retained CLASSI-N solution is only partially unique, since one person can be switched from $PT_1$

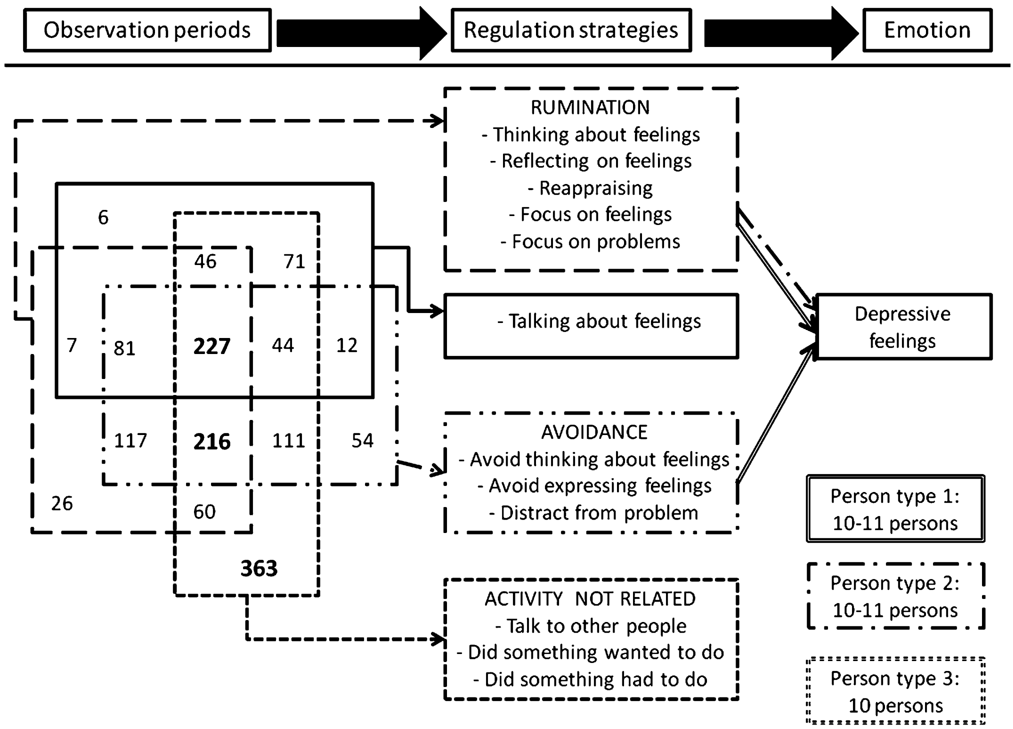


FIGURE 4.
Graphical representation of the (4, 3) CLASSI-N decomposition of the depression data. Note that one person can be switched from person type 1 to person type 2 without altering the fit of the CLASSI-N solution.

to $PT_2$ without altering the loss function value. This can be explained by noting that this person never tried to avoid the problem without ruminating also. Therefore we cannot evaluate what happens if the person only avoids the problem, and thus we cannot distinguish between the first and second person type.

In conclusion, these results confirm that individual differences exist in the way regulatory cognitions and actions are linked to depression (Koole, 2009). Whereas some people are not easily depressed or are very effective in addressing their feelings ($PT_3$), others are less effective in that they feel depressed when they ruminate about the problem as well as when they are avoiding it ($PT_1$). The second group ($PT_2$) takes an intermediate position in that for them avoiding the problem is not related to depression afterwards. These results are in line with previous findings that rumination generally leads to depression (Nolen-Hoeksema et al., 2008), whereas the effectiveness of avoidant strategies differs across persons (Koole, 2009). We externally validated the interpretation of these person types by performing an analysis of variance with person type as independent variable and the scores on a depression scale, the CES-D scale (Radloff, 2006; Shafer, 2006), as dependent variable. Note that the person that could not be uniquely assigned to one person type was omitted from this analysis. The results of this ANOVA indicate that the three groups differ significantly ($F(2, 27) = 5.17$, $p < 0.05$). As expected, a Tukey test revealed that $PT_1$ is significantly more depressed than $PT_3$ with $PT_2$ taking an intermediate position.

## 6. Conclusion and Discussion

Sequential processes are manifold in psychological theories—see for instance the ABC-model in rational-emotive therapy (Dryden & David, 2008); the Cognitive Affective Personality

System theory (Mischel & Shoda, 1998); appraisal theory (Roseman & Smith, 2001), etc. Therefore, the development of data analysis tools that allow to obtain a concise picture of such sequential processes and individual differences therein, is an important topic of research. In this paper, we proposed the CLASSI-N method, which extends the CLASSI technique to nested data, that is, data where the stimuli under study may differ across persons. Like CLASSI, CLASSI-N aims at charting sequential processes and individual differences therein, by reducing the mediating variables to a few mediating variable types and summarizing the individual differences in the *if* <mediating variable>–*then* <response> rules by inducing a person typology from the data. As many recently developed data gathering techniques result in nested data, the range of problems to which the CLASSI framework can be applied has strongly increased through the development of CLASSI-N.

In this section, we first compare CLASSI-N to the original CLASSI model. Next the CLASSI-N model is related to several other existing methods.

### *6.1. Relation to the CLASSI Method*

The key difference between CLASSI and CLASSI-N is that CLASSI requires three-way three-mode data and CLASSI-N nested two-way data. In other words, whereas CLASSI assumes that people's responses to the same set of stimuli have been studied, CLASSI-N allows the set of stimuli to differ across persons. As such, CLASSI-N can handle CLASSI data sets, but not the other way around; therefore, one could say that CLASSI-N subsumes CLASSI as a special case.

It is important to note, however, that because crossed data yield person-specific information for the same set of stimuli, such data allow the study of additional aspects of sequential processes and individual differences therein. Specifically, one may investigate whether some of the stimuli are functionally equivalent in that they activate the same mediating variables and responses across persons. Moreover, it can be examined whether the relationship between the stimuli and the mediating variables is subject to individual differences. To this end, apart from inducing a mediating variable typology and a M-R person typology that summarizes individual differences in the relationship between the mediating variables and the response, CLASSI also includes a stimulus typology matrix $\mathbf{S}^\star$ (which reduces the stimuli to $P$ stimulus types), an S-M person typology matrix $\mathbf{P}^{\text{S-M}}$, and an associated linking array $\underline{\mathbf{L}}^{\text{S-M}}$ that holds the *if* <stimulus type>–*then* <mediating variable type> rules of the $R$ S-M person types. Specifically, in the CLASSI model the $x^M_{ijk}$ entries of the mediating variable data array, which indicate whether or not stimulus $i$ elicits mediating variable $j$ in person $k$, are reconstructed by means of the following rule:

$$x^M_{ijk} \approx \hat{x}^M_{ijk} = \bigoplus_{p=1}^{P} \bigoplus_{q=1}^{Q} \bigoplus_{r=1}^{R} s^\star_{ip} m_{jq} p^{\text{S-M}}_{kr} l^{\text{S-M}}_{pqr}. \tag{10}$$

This rule reveals that CLASSI is a special case of CLASSI-N in that the restriction is imposed that the mediator profile matrix $\mathbf{S}$ can be further decomposed as follows:

$$s_{i_k q} = \bigoplus_{p=1}^{P} \bigoplus_{r=1}^{R} s^\star_{ip} p^{\text{S-M}}_{kr} l^{\text{S-M}}_{pqr}. \tag{11}$$

Finally, another less important difference between CLASSI and CLASSI-N is that CLASSI-N can handle only one response variable, whereas in CLASSI multiple responses can be studied simultaneously.

### 6.2. *Relations to Other Methods*

The CLASSI-N method can be related to several other existing methods. We will focus first on the CLASSI-N submodels for reconstructing $X^M$ and $x^R$. Subsequently, we will consider the global CLASSI-N model.

The CLASSI-N submodel for reconstructing the mediating variable data matrix $X^M$ (reconstruction rule 1), boils down to a special case of a K-means clustering model (Hartigan, 1975), in which the variables (rather than the objects) are clustered and in which the centroids are binary instead of real-valued (Li, 2005). The CLASSI-N submodel for the reconstruction of the response data vector $x^R$ (reconstruction rule 2), can be considered a Boolean clusterwise regression model (DeSarbo, Oliver, & Rangaswamy, 1989). In clusterwise regression, one examines the relationship between a set of predictor variables and a criterion where it is assumed that the relationship can differ across observations. Clusterwise regression deals with these differences in the relationship between predictors and criterion by assigning the observations to a few clusters and by specifying a separate regression model for each cluster. Analogously, the CLASSI-N method reduces the persons under study and thus the stimuli that are nested in each person, into a few types, which are characterized by a different Boolean regression model.

The global CLASSI-N model is related to the covariates regression approach (de Jong & Kiers, 1992). This approach aims to regress a set of criterion variables on several predictor variables, where the dimensionality of the predictor variables is reduced by means of a principal component analysis. In a similar vein, the global CLASSI-N model relates the responses to the different mediating variables, which are reduced to a few mutual exclusive types, by means of a Boolean regression. Thus, the CLASSI-N method differs from covariates regression in that Boolean algebra is used instead of linear algebra, and in that the reduction of the mediating variables is not dimensional but categorical.

Summarizing, the CLASSI-N method combines the key idea behind clusterwise regression—namely observations are clustered and for each cluster of observations a different regression model is estimated—with the key principle of the covariates regression approach—simultaneously reducing the number of predictor variables and relating these predictors to a criterion.

## Acknowledgement

The research reported in this paper was partially supported by the Research Council of K.U. Leuven (GOA/05/04 and GOA/10/02).

## Appendix

**THEOREM 1.** *If a* $(Q, T)$ *CLASSI-N decomposition of an* $I \times J$ *binary reconstructed data matrix* $\hat{\mathbf{X}}^M$ *and an* $I \times 1$ *binary reconstructed data vector* $\hat{\mathbf{x}}^R$ *exists such that for all persons* $k$ $(k = 1, \ldots, K)$ *and for all mediating variable types* $q$ $(q = 1, \ldots, Q)$, *at least one stimulus* $i_k$ *exists that activates this mediating variable type only, this decomposition is unique upon a permutation of the mediating variable and person types.*

*Proof:* In this proof we will assume, without loss of generality, that the first $Q$ rows of $\mathbf{M}$ refer to one element of each of the $Q$ mediating variable types. In particular, for $1 \leq j \leq Q$, $m_{jq} = 1$ iff $q = j$. This implies that the $Q \times Q$ submatrix $\mathbf{M}^{\Delta}$ constituted by the first $Q$ rows of $\mathbf{M}$, is an identity matrix. Similarly, we assume that, for each person $k$ $(k = 1, \ldots, K)$, the $q$th $(q = 1, \ldots, Q)$ row of $\mathbf{S}_k$ refers to a stimulus which activates the $q$th mediating variable type only, $\mathbf{S}_k$ being the submatrix of $\mathbf{S}$ that holds the mediator profiles of person $k$. Specifically, for $1 \le i_k \le Q$, $s_{i_k q} = 1$ iff $q = i_k$. In that case the $Q \times Q$ submatrix $\mathbf{S}_k^*$, based on the first $Q$ rows of $\mathbf{S}_k$, is an identity matrix.

To prove the uniqueness of $\mathbf{S}$, we consider only the first $Q$ mediating variables. According to Equation (1), the corresponding columns of $\hat{\mathbf{X}}^M$, collected in the $I \times Q$ matrix $\hat{\mathbf{X}}^{\Delta M}$, equal:

$$\hat{\mathbf{X}}^{\Delta M} = \mathbf{S} \otimes \mathbf{M}'^{\Delta} = \mathbf{S} \tag{A.1}$$

where $\otimes$ denotes the Boolean matrix product. Therefore, no entry of $\mathbf{S}$ can be modified without modifying $\hat{\mathbf{X}}^{\Delta M}$ and, thus, $\mathbf{S}$ is unique.

With respect to $\mathbf{M}$, for all persons $k$ $(k = 1, \dots, K)$ the first $Q$ rows of $\hat{\mathbf{X}}_k^M$, constituting the matrix $\hat{\mathbf{X}}_k^{*M}$, equal

$$\hat{\mathbf{X}}_k^{*M} = \mathbf{S}_k^* \otimes \mathbf{M}' = \mathbf{M}'. \tag{A.2}$$

Hence, no entry of $\mathbf{M}$ can be changed without changing $\hat{\mathbf{X}}_k^{*M}$ and, thus, $\mathbf{M}$ is unique.

To prove the uniqueness of $\mathbf{L}$ and $\mathbf{P}$, for all persons $k$ $(k = 1, \dots, K)$, we find that the subvector $\hat{\mathbf{x}}_k^{*R}$, which contains the first $Q$ elements of $\hat{\mathbf{x}}_k^R$, equals

$$\hat{\mathbf{x}}_k^{*R} = \mathbf{S}_k^* \otimes \mathbf{L}_{k \in t} = \mathbf{L}_{k \in t}, \tag{A.3}$$

with $\mathbf{L}_{k \in t}$ denoting the column of $\mathbf{L}$ that pertains to the person type $PT_t$ to which person $k$ belongs. This implies that no entry of $\mathbf{L}$ can be changed without changing $\hat{\mathbf{x}}^{*R}$ and, thus $\mathbf{L}$ is unique. Finally, if $\mathbf{P}$ is non-unique, a person $k$ exists that can be switched from one person type $PT_{t_1}$ to another person type $PT_{t_2}$, implying that

$$\hat{\mathbf{x}}_k^{*R} = \mathbf{L}_{k \in t_1} = \mathbf{L}_{k \in t_2}. \tag{A.4}$$

This result violates the CLASSI-N restriction that no column of the linking matrix $\mathbf{L}$ equals another column. □

## References


Aarts, E., Korst, J., & Michiels, W. (2005). Simulated annealing. In E.K. Burke & G. Kendall (Eds.), *Search methodologies* (pp. 187–210). New York: Springer.

Barrett, L.F., & Barrett, D.J. (2001). An introduction to computerized experience sampling in psychology. *Social Science Computer Review*, *19*(2), 175–185.

Bigné, J.E., Mattila, A.S., & Andreu, L. (2008). The impact of experiential consumption cognitions and emotions on behavioral intentions. *Journal of Services Marketing*, *22*(4), 303–315.

Brusco, M.J. (2001). A simulated annealing heuristic for unidimensional and multidimensional (city-block) scaling of symmetric proximity matrices. *Journal of Classification*, *18*(1), 3–33.

Ceulemans, E., & Van Mechelen, I. (2008). CLASSI: a classification model for the study of sequential processes and individual differences therein. *Psychometrika*, *73*(1), 107–124.

Chaturvedi, A., & Carroll, J.D. (1994). An alternating combinatorial optimization approach to fitting the INDCLUS and generalized INDCLUS models. *Journal of Classification*, *11*(2), 155–170.

Csikszentmihalyi, M., & Larsen, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, *175*(9), 526–536.

Davidson, R.J., Jackson, D.C., & Kalin, N.H. (2000). Emotion plasticity, context, and regulation: perspectives from affective neuroscience. *Psychological Bulletin*, *126*(6), 890–909.

de Jong, S., & Kiers, H.A.L. (1992). Principal covariates regression. Part I. Theory. *Chemometrics and Intelligent Laboratory Systems*, *14*(1–3), 155–164.

DeSarbo, W.S., Oliver, R.L., & Rangaswamy, A. (1989). A simulated annealing methodology for clusterwise linear regression. *Psychometrika*, *54*(4), 707–736.

Dryden, W., & David, D. (2008). Rational emotive behaviour therapy: current status. *Journal of Cognitive Psychotherapy: An International Quarterly*, *22*(3), 195–209.

Gross, J.J. (1998). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*, 224–237.

Gross, J.J. (2001). Emotion regulation in adulthood: timing is everything. *Current Directions in Psychological Science*, *10*(6), 214–219.
Gross, J.J. (2008). Emotion regulation. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (pp. 497–512). New York: Guilford.
Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: implications for affect relationships and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362.
Haggard, E.A. (1958). *Intraclass correlation and the analysis of variance*. Troy: Dryden.
Hartigan, J.A. (1975). *Clustering algorithms*. New York: Wiley.
Hemenover, S.H., Augustine, A.A., Shulman, T., Tran, T., & Barlett, C.P. (2008). Individual differences in negative affect repair. *Emotion*, *8*(4), 468–478.
Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, *2*(1), 193–218.
Kirk, R.E. (1995). *Experimental design: procedures for the behavioral sciences*. Belmont: Brooks and Cole.
Kirkpatrick, S., Gelatt, C.D., & Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*, *220*(4598), 671–680.
Knuth, E.J., Stephens, A.C., McNeil, N.M., & Alibali, M.W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*(4), 297–312.
Koole, S.L. (2009). The psychology of emotion regulation: an integrative review. *Cognition and Emotion*, *23*(1), 4–41.
Kuppens, P., Van Mechelen, I., & Rijmen, F. (2008). Towards disentangling sources of individual differences in appraisal and anger. *Journal of Personality*, *76*(4), 969–1000.
Leenen, I., & Van Mechelen, I. (2001). An evaluation of two algorithms for hierarchical classes analysis. *Journal of Classification*, *18*(1), 57–80.
Li, T. (2005). A general model for clustering binary data. In *KDD'05: Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining* (pp. 188–197). New York: ACM.
Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology*, *49*(1), 229–258.
Nolen-Hoeksema, S., Wisco, B.E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, *3*(5), 400–424.
O'Brien, T.B., & DeLongis, A. (1996). The interactional context of problem-, emotion-, and relationship-focused coping: the role of the big five personality factors. *Journal of Personality*, *64*(4), 775–813.
Radloff, L.S. (2006). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.
Robinson, N.S., Garber, J., & Hilsman, R. (1995). Cognitions and stress: direct and moderating effects on depressive versus externalizing symptoms during the junior high school transition. *Journal of Abnormal Psychology*, *104*(3), 453–463.
Roseman, I.J., & Smith, C.A. (2001). Appraisal theory: overview, assumptions, varieties, controversies. In K.R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: theory, methods, research* (pp. 3–19). London: Oxford University Press.
Scherer, K.R. (1999). Appraisal theory. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). Chichester: Wiley.
Shafer, A.B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*, *62*, 123–146.
Steinley, D. (2004). Properties of the Hubert-Arabie adjusted Rand index. *Psychological Methods*, *9*(3), 386–396.
Trejos, J., & Castillo, W. (2000). Simulated annealing optimization for two-mode partitioning. In W. Gaul & R. Decker (Eds.), *Classification and information at the turn of the millenium* (pp. 135–142). Heidelberg: Springer.
Van Mechelen, I., & Hennes, K. (2009). The appraisal basis of anger occurence and intensity revisited. *Cognition and Emotion*, *23*(7), 1373–1388.
Wilderjans, T.F., Ceulemans, E., & Van Mechelen, I. (2008). The CHIC model: a global model for coupled binary data. *Psychometrika*, *73*(4), 729–751.

*Manuscript Received: 11 MAR 2010*
*Final Version Received: 5 NOV 2010*
*Published Online Date: 17 NOV 2011*