Adequate assessment of adherence to medical treatment is critical for both research purposes and clinical practice. This study examined the factor structure and longitudinal invariance of the Medication Adherence Report Scale (MARS-A10) in a sample of asthmatic patients. We examined longitudinal data from 294 inner-city, adult participants with moderate to severe asthma. Because of ambiguous evidence regarding the dimensionality of the MARS-A10, the data was analysed with exploratory structural equation modelling. We first proceeded by determining the dimensionality of the scale at baseline and examined whether the structure, loadings, intercepts and errors were invariant over the four assessments points. Results indicated that a two-factor structure (factor 1: non-adherence based on experiential changes; factor 2: non-adherence based on intentional medication avoidance) had the best fit to the data ($\chi^2(25) = 37.69, p = 0.05$). Longitudinal analyses revealed that the nine items assessing intentional non-adherence were invariant over time. The evidence from the factor analysis suggests that intentional non-adherence is a multidimensional construct. Additionally, longitudinal data provided strong evidence that the items examining intentional non-adherence are invariant over time, indicating that changes in non-adherence scores can be validly attributed to changes in behaviour.

Keywords: adherence; asthma; longitudinal invariance; self-report

Introduction

Patients afflicted with chronic disease must adhere to prescribed medical treatments in order to reduce morbidity and mortality (World Health Organization, 2003). For asthmatic patients, adherence to inhaled corticosteroids (ICS) is the basis for the
management of moderate to severe asthma (National Asthma Education and Prevention Program: Expert Panel Report 3, 2007). Adequate assessment of adherence is fundamental from both a practical and a theoretical perspective. From a practical perspective, adequate assessment of adherence allows practitioners to determine whether patients’ asthma-related problems are due to the intrinsic disease severity, the inadequacy of treatment or a failure to follow prescribed recommendations (Cochrane, Horne, & Chanez, 1999). From a theoretical perspective, a valid measure of adherence allows researchers to properly identify the impact of different types of factors that may influence adherence behaviours directly and/or indirectly (Dunbar-Jacob et al., 2000; Horne, 1998). This theory building process is critical because it helps the researchers’ efforts to identify targets of interventions for improving adherence to medical treatment (Haynes et al., 2005).

There are several challenges involved in the assessment of adherence to medical treatment on a regular basis. For example, the use of electronic monitoring in large studies is logistically difficult and cost prohibitive. Other objective measures, such as biochemical and pharmacy data have similar limitations (e.g. cost and availability) that render them impractical as well (Dunbar-Jacob, Sereika, Burke, & Ockene, 2001). Although self-reports also have limitations, they enjoy the significant advantage of being easy to implement and are less intrusive. Additionally, unlike electronic monitoring and biochemical information, self-report measures allow researchers to examine psychological and behavioural processes, such as the type of non-adherence (i.e. intentional or unintentional) (Clifford, Barber, & Horne, 2008; Wroe, 2002). Thus, self-report data can provide important information for the development of tailored interventions (Smith et al., 2007).

The Medication Adherence Report Scale for Asthma (MARS-A10© (©Prof. Robert Horne, 1999)) developed by Horne and collaborators is a brief self-measure that has demonstrated good psychometric properties (Horne & Hankins, 2008; Horne & Weinman, 2002). Studies have shown that self-reports of non-adherence behaviours as assessed by full and short versions of the MARS-Asthma are significantly associated with objective measures of adherence, such as electronic monitoring (Cohen et al., 2008; Ohm & Aaronson, 2006) and pill counting (Menckeberg et al., 2008). Because of this, the MARS-A10 is a valuable option for large-scale studies examining adherence to medical treatment among asthmatics. Additionally, these characteristics make the MARS-A10 a valuable tool for clinicians as they rely on self-reports to determine the extent to which their patients are following the prescribed medical treatment.

Although the evidence from cross-sectional studies indicates that the MARS-A10 is a valid and reliable measure, little is known about whether its measurement properties are consistent when examined in a longitudinal framework (i.e. whether the MARS-A10 is invariant over time). Longitudinal measurement invariance is a crucial assumption that needs to be tested when a researcher wants to ensure that an instrument is assessing the same construct over time. Testing this assumption is fundamental if one is to determine whether mean differences in adherence behaviour over time and correlations between waves are due to real changes in the construct or due to extraneous factors. To demonstrate longitudinal measurement invariance it is necessary to establish that the factor structure, the meaning of items (i.e. loadings) and the items means (i.e. intercepts) and their uniqueness (i.e. residual variances) are the same over time.
The main goal of this study was to provide evidence for the longitudinal invariance of the MARS-A10 over four waves of data. Although several studies have treated various versions of the MARS-Asthma (e.g., 10 and 5 items) as a measure of a unidimensional construct (e.g. Ponieman et al., 2009; Smith et al., 2007), we decided to use exploratory structural equation modelling (ESEM) instead of confirmatory factor analysis (CFA) because the evidence regarding the dimensionality of the MARS-A10 is ambiguous (Horne & Hankins, 2008) and because the items assess two different types of non-adherence (i.e., intentional and unintentional). Thus, instead of using a confirmatory technique in an exploratory fashion (Browne, 2001), we decided to use a technique that allows the examination of the factor structure of a measure and its invariance over time without pre-specifying a factor structure (see more details in section ‘Method’).

Method

Participants
We analysed data from a four-wave longitudinal survey (measurement points: baseline, and 1 month, 3 month and 12 month). Participants were adult asthmatics who received care at one of the two hospital-based primary care clinics located in New York City or in New Brunswick, NJ. Patient eligibility for the study required: (1) a physician’s diagnosis of asthma (mild persistent, moderate or severe) according to the National Institutes of Health guidelines (National Asthma Education and Prevention Program: Expert Panel Report 3, 2007), (2) that patients be 18 years of age or older and (3) patient’s fluency in English and/or Spanish. Individuals with a smoking history of $\geq 10$ pack-years, a diagnosis of chronic obstructive lung disease or other chronic respiratory illnesses were excluded from the study. Twenty-one per cent of the participants completed the survey in Spanish (details of the recruitment process and sampling frame can be found in Mora et al., 2009).

For this study, we analysed data from 294 (out of 326) patients who had been prescribed ICS and had responded to the MARS-A10 items at any wave. The mean age of this sample was 48 years (range 20 to 87) and 82% were female. A majority of the sample (60.7%) had an annual household income of $< 15,000$ US dollars. The ethnic breakdown was 55% Latinos, 27% African-Americans and 13% European-Americans. The remaining 5% consisted of Native-Americans, Asian-Americans and others. Asthma morbidity was considerable, with 11% reporting a history of intubation, 52.4% reporting emergency department visits and 68% having taken oral steroids more than once in the past year.

The study protocol was approved by the Institutional Review Boards of both institutions (Protocols: GCO#03-0725 and 0220055507 respectively), and written informed consent was obtained from all study participants.

Data collection and procedures
All interviews were conducted in the clinics using a room dedicated to the study. Bilingual research staff conducted an interviewer-administered survey in English or Spanish. The survey collected information, such as educational level, income, asthma history, illness beliefs and use of health care. Interviewers were trained by the principal investigator and senior research staff. Training consisted of: (1) a workshop
in which theoretical and practical aspects related to the study were discussed, (2) a session in which interviewers familiarised themselves with the data collection instruments and procedures and (3) mock interviews that were observed and critiqued by the senior research staff.

**Measure**

*Medication Adherence Report Scale for Asthma*

The MARS-A10 is a 10-item questionnaire that includes both generic (‘I use it regularly every day’) and asthma-specific (‘I only use it when I feel breathless’) questions about medication use (Cochrane et al., 1999). Nine items assess intentional (‘I avoid using it if I can’), and one item assesses unintentional (‘I forget to use it’) non-adherence (Horne & Hankins, 2008). Questions are framed as negative statements to minimise social desirability bias and to indicate that non-adherence is a normal behaviour (Rand & Wise, 1994). The MARS-A10 items emphasise acts that obstruct the regular use of ICS medication rather than compliance with practitioners’ instructions. Responses are recorded on a 5-point Likert scale (1 = always to 5 = never). The following statement preceded the MARS items: ‘Many people find a way of using their asthma controller medicines, I mean: [NAME OF CONTROLLER USED BY PATIENT] which suits them. This may differ from the instructions on the label or from what their doctor has said. We would like to ask you a few questions about how you use your medicines: How often do you do the following?’ (Prof. Robert Horne, 1999).

**Statistical analysis and procedures**

The analyses for this study were conducted with a statistical technique recently described by Asparouhov and Muthén (2009): ESEM. In the context of psychometric studies, ESEM allows for the rigorous examination of the underlying structure of responses to a scale based on exploratory factor analysis (EFA) modelling. Items are not restricted to load on one and only one factor as is required when using CFA. Factor analysis using ESEM provides more accurate estimates of factor loadings and between-factor correlations when compared with CFA when cross-loadings are significantly different from zero. Additionally, unlike EFA, ESEM permits investigators to estimate the standard errors of factor loadings, to include correlated residuals that reflect common methods factors (e.g. aspects of item or response format) and to examine the invariance of measurement instruments (Marsh et al., 2009). These characteristics make ESEM an important tool to study the properties and mean structure of instruments when the dimensionality of a measure is not well-known (technical details can be found in Asparouhov & Muthén, 2009).

In the context of ESEM, longitudinal invariance is assessed with a series of nested models similar to those used for the examination of multigroup and longitudinal invariance with CFA (Meredith & Teresi, 2006). In this study we first examined whether the items loaded on the same latent factors over time (i.e. configural invariance). We were interested in examining whether an unrestricted longitudinal model would show similarities in the pattern of item-factor associations. A second model tested the metric invariance which required that the factor pattern and factor loadings be invariant across the four measurement occasions (i.e. the meaning of
items were the same across waves). The third model assessed whether the factor loadings and item intercepts (i.e. indicator means) were invariant over the multiple measurement occasions (i.e. scalar invariance). Support of scalar invariance indicates that there is no differential item functioning over time. The last model examined the invariance over time of the item residual variances (i.e. strict invariance). Strict invariance requires that loadings, means and residual variance be equivalent across waves. Equality of item residual variances implies equal reliabilities across waves and that factor means and covariances can be compared over time (Gregorich, 2006; Meredith & Teresi, 2006).

The analyses were conducted in three steps. First, we examined the dimensionality of the MARS-A10 using ESEM. Three exploratory models were estimated to determine whether a one-, two- or three-factor model better represented the data. Based on interviewers’ feedback that participants had some difficulties answering the first two items (i.e. item 1: ‘I only use my [NAME OF MEDICINE] when I need it’; and item 2: ‘I only use it when I feel breathless’) and because of the similarity in their wording, the residual variances of these items were allowed to covary in all models. Second, we examined the longitudinal invariance of the MARS-A10 following the steps delineated above. These longitudinal models included both within- and between-wave correlated residual variances. The only freely estimated within-wave was that between the residual variances of item 1 and item 2. We also estimated the correlations of each item’s uniqueness assessed on multiple occasions (i.e. between-wave). This decision follows recommendations that failure to include these between-wave correlations may bias the estimation of the model parameters (Marsh & Hau, 1996). Finally, we examined the composite reliability (i.e. internal consistency) of the full scale and the subscales across waves (Raykov, 1997).

The covariance coverage for the longitudinal data (i.e. proportion of subjects for whom there is available data) ranged between 0.45 and 0.93. The incomplete data were assumed to be missing at random (MAR). MAR assumes that the probabilities of missing values can be predicted from variables that are not missing, such as those assessed at the first wave (Little & Rubin, 2002). In longitudinal studies, such as the current one, scores for missing cases show a MAR pattern as they are correlated with their own scores from earlier waves.

The Mardia test of multivariate skewness and kurtosis revealed that the multivariate distribution of the data at each wave showed important departures from normality (skewness ranged between 37.14 and 40.45; kurtosis ranged between 189.71 and 201.38, all p’s < 0.01). Because of this, parameters of the ESEM models were estimated using full-information maximum likelihood (FIML) with robust standard errors (T2* statistic: Yuan & Bentler, 2000) in Mplus 5.21. We chose this procedure because parameter estimation with all cases (complete and incomplete) yields less-biased parameter estimates than those obtained with complete cases (i.e. analysis using listwise or pairwise case deletion) (Schafer & Graham, 2002). Latent factors were rotated with geomin rotation which is the default option in Mplus.

Model fit was evaluated based on various indices of exact and close fit: robust $\chi^2$, the comparative fit index (CFI), the Tucker–Lewis fit index (TLI), the root mean squared error of approximation (RMSEA) and the standardised root mean square residual (SRMR). As Marsh et al. (in press) have indicated; it is unclear the extent to which traditional cut-off values for model fit used in a CFA context are relevant for ESEM models. Thus, frequently used cut-offs for these various fit indices
To ascertain lack of measurement invariance we considered the following information when comparing nested models (Chen, 2007): non-significant \( \chi^2 \) difference test (we followed the procedure described by Satorra & Bentler, 2001); differences in CFI (decrements \( \leq 0.005 \)) and changes in RMSEA (increases \( \leq 0.01 \)).

### Results

**Factor structure of the MARS-A10**

To determine the dimensionality of the MARS-A10, we tested and compared three models (Table 1). The model examining a one-factor structure (model 1) had an inadequate fit to the data. The three-factor model did not produce reliable results due to identification problems (i.e. correction factor value was greater than 1). Model 2 which examined a two-factor structure had an excellent fit. These results showed that items 1, 2, 8 and 9 loaded on a factor that reflected non-adherence based on experiential and somatic changes (standardised loadings ranged between 0.55 and 0.76, all \( p's < 0.01 \)). Although the loadings for items 3, 4, and 6 on this factor were not statistically significant, their sizes were substantial enough (range 0.11–0.18) to suggest that constraining them to be zero might lead to biases in parameter estimation. Items 3, 4, 5, 7 and 10 were significantly associated with factor 2 (range of standardised loadings range 0.38–0.86, all \( p's < 0.05 \)). The content of these items suggests that the construct assessed by this factor was non-adherence due to the intentional avoidance of medication. The error correlation between item 1 and item 2 was significant \( r = 0.62, p < 0.01 \). The between-factor correlation was \( r = 0.67, \)

---

Table 1. Overall fit of models examining the dimensionality of the MARS-A10 and its invariance over time.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( c )</th>
<th>( p )</th>
<th>RMSEA (90% CI)</th>
<th>CFI</th>
<th>TLI</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor structure baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1: One factor</td>
<td>89.88</td>
<td>34</td>
<td>1.602</td>
<td>&lt;0.01</td>
<td>0.077 (0.058, 0.096)</td>
<td>0.915</td>
<td>0.888</td>
<td>0.074</td>
</tr>
<tr>
<td>Model 2: Two factors</td>
<td>37.69</td>
<td>25</td>
<td>1.489</td>
<td>0.05</td>
<td>0.043 (0.002, 0.069)</td>
<td>0.981</td>
<td>0.965</td>
<td>0.034</td>
</tr>
<tr>
<td>Multigroup comparisons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3: Unconditional longitudinal</td>
<td>847.25</td>
<td>616</td>
<td>1.238</td>
<td>&lt;0.01</td>
<td>0.036 (0.030, 0.041)</td>
<td>0.936</td>
<td>0.919</td>
<td>0.062</td>
</tr>
<tr>
<td>Model 4: Equality of loadings (factor means freed)</td>
<td>867.45</td>
<td>664</td>
<td>1.280</td>
<td>&lt;0.01</td>
<td>0.032 (0.026, 0.038)</td>
<td>0.944</td>
<td>0.934</td>
<td>0.067</td>
</tr>
<tr>
<td>Model 5: Item intercepts</td>
<td>898.71</td>
<td>688</td>
<td>1.269</td>
<td>&lt;0.01</td>
<td>0.032 (0.026, 0.038)</td>
<td>0.942</td>
<td>0.934</td>
<td>0.067</td>
</tr>
<tr>
<td>Model 6: Equality of item errors</td>
<td>953.54</td>
<td>718</td>
<td>1.287</td>
<td>&lt;0.01</td>
<td>0.033 (0.027, 0.039)</td>
<td>0.935</td>
<td>0.929</td>
<td>0.088</td>
</tr>
<tr>
<td>Model 7: Equality of item errorsb</td>
<td>938.74</td>
<td>715</td>
<td>1.289</td>
<td>&lt;0.01</td>
<td>0.033 (0.027, 0.038)</td>
<td>0.938</td>
<td>0.933</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Notes: \( c = \) non-normality scaling correction factor.

\( \text{aRescaled chi-square.} \)

\( \text{bErrors for 'forget' item freely estimated.} \)

(i.e. probability of \( \chi^2 \) value \( > 0.05 \), CFI and TLI \( > 0.90 \), RMSEA \( < 0.05 \) and SRMR \( < 0.08 \)) were used only as guidelines (Marsh et al., in press). Additionally, we assessed the relevance of the models based on the size and the pattern of parameter estimates. To ascertain lack of measurement invariance we considered the following information when comparing nested models (Chen, 2007): non-significant \( \chi^2 \) difference test (we followed the procedure described by Satorra & Bentler, 2001); differences in CFI (decrements \( \leq 0.005 \)) and changes in RMSEA (increases \( \leq 0.01 \)).
p<0.01, indicating that despite the important overlap each factor tapped different aspects of intentional non-adherence ($r^2=0.45$). The correlations between factor scores estimates and the respective factors (i.e. factor determinacies) were high for the two factors ($\rho_{\text{factor 1}} = 0.92$ and $\rho_{\text{factor 2}} = 0.94$).

**Longitudinal invariance**

To test whether the two-factor model was invariant over time, we followed the procedures delineated in section ‘Method’. First, we examined the pattern of factor loadings with an unconstrained longitudinal model. We were interested in examining the similarity of the parameter estimates and the pattern of the factor loadings. Table 2 shows the unstandardised factor loadings of each item across the four waves for the two MARS-A10 factors. As can be seen, both the factor structure and factor loadings show an important degree of stability over the four assessment points.
The chi-square test for this model was significant; however, the other indices revealed an adequate close fit of the model to the data (see Table 1, model 3). Model 4 tested metric invariance by constraining the factor loadings to be equal across waves. The close fit of this model to the data was good and did not differ from that of model 2 ($\chi^2_{\text{diff}}(48) = 33.77, p = 0.94; \Delta\text{CFI} = 0.012; \Delta\text{RMSEA} = -0.004$). This supported the assumption that the meaning of items over the different assessment points was equivalent. We then examined scalar invariance by constraining item intercepts to be equal across waves. The close fit of this model (cf. Table 1, model 5) was similar to the less parsimonious model testing metric invariance ($\chi^2_{\text{diff}}(24) = 31.23, p = 0.15; \Delta\text{CFI} = -0.002; \Delta\text{RMSEA} = 0.000$). Support for scalar invariance indicates that changes in the means can reasonably be interpreted as changes in the constructs and not in confounding factors.

The results for model 6 (cf. Table 1) indicated that strict invariance was not tenable ($\chi^2_{\text{diff}}(30) = 51.03, p < 0.01; \Delta\text{CFI} = -0.007; \Delta\text{RMSEA} = 0.001$). Based on the modification indices, we decided to remove the equality constraint for the variance error of item 5 (i.e. I forget to take it). The close fit of this new model (i.e. model 7) was indistinguishable from the fit of model 5, providing support for partial strict invariance ($\chi^2_{\text{diff}}(27) = 38.68, p = 0.07; \Delta\text{CFI} = -0.004; \Delta\text{RMSEA} = 0.001$). Table 3 shows the factor loadings, uniqueness and intercepts obtained in this last model.

### Table 3. Factor structure, unstandardised loadings and intercepts from the model testing partial strict invariance (model 7).

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Factor 1 Loading</th>
<th>Factor 2 Loading</th>
<th>Uniqueness</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I only use my [NAME OF MEDICINE] when I need it</td>
<td>1.06**</td>
<td>0.22**</td>
<td>0.93</td>
<td>3.83</td>
</tr>
<tr>
<td>2. I only use it when I feel breathless</td>
<td>1.26**</td>
<td>0.00</td>
<td>0.84</td>
<td>3.89</td>
</tr>
<tr>
<td>3. I decide to miss out a dose</td>
<td>0.01</td>
<td>0.86**</td>
<td>0.49</td>
<td>4.21</td>
</tr>
<tr>
<td>4. I try to avoid using it</td>
<td>0.22*</td>
<td>0.70**</td>
<td>0.53</td>
<td>4.30</td>
</tr>
<tr>
<td>5. I forget to take it*</td>
<td>-0.18*</td>
<td>0.48**</td>
<td>0.88</td>
<td>4.24</td>
</tr>
<tr>
<td>6. I alter the dose</td>
<td>0.01</td>
<td>0.39**</td>
<td>0.59</td>
<td>4.64</td>
</tr>
<tr>
<td>7. I stop taking it for a while</td>
<td>0.02</td>
<td>0.81**</td>
<td>0.44</td>
<td>4.40</td>
</tr>
<tr>
<td>8. I use it as a reserve, if my other treatment doesn’t work</td>
<td>0.61**</td>
<td>0.31**</td>
<td>0.91</td>
<td>4.31</td>
</tr>
<tr>
<td>9. I use it before doing something which might make me breathless</td>
<td>0.90**</td>
<td>-0.23*</td>
<td>1.42</td>
<td>4.05</td>
</tr>
<tr>
<td>10. I take it less than instructed</td>
<td>-0.06</td>
<td>0.94**</td>
<td>0.58</td>
<td>4.29</td>
</tr>
</tbody>
</table>

Notes: *Error for this item variable correspond to Wave 1. The error terms for this item in subsequent waves are: Wave 2 = 0.97; Wave 3 = 0.74; Wave 4 = 0.57.

* $p < 0.05.

** $p < 0.01.$

The stability of the MARS-A10 over time was determined based on two types of evidence: changes in latent factor means and factor correlations over time. Table 4 shows the latent means of the two latent factors over the four measurement points.
Table 4. Means, variances, factor determinacies and inter-wave correlations for the MARS-A10 latent factors.

<table>
<thead>
<tr>
<th></th>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
<th>Wave 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>0.58</td>
<td>0.71</td>
<td>0.71</td>
<td>0.68</td>
</tr>
<tr>
<td>Factor 2</td>
<td>–</td>
<td>0.44</td>
<td>0.37</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Note: These estimates were obtained from model 7. All correlations are significant at an alpha level = .01.

(for wave 1, the factor means were set at 0 and their variances were set at 1). The results revealed that factors’ means (Z tests ranged between 0.27 and 0.84, p’s > 0.40) and variances (Z tests range −0.88 and 0.60, p’s > 0.38) were stable. These results demonstrated that adherence behaviour over time for this group remained the same throughout the 12-month period.

Inter wave factor correlations showed the typical pattern of associations for repeated measures; scores closer in time are more strongly correlated with one another than scores from waves farther apart (Table 4). The inter wave factor correlation range for factor 1 (0.57, 0.79) and factor 2 (0.57, −0.79) indicated that there is an important degree of stability for inter individual differences in non-adherence over time. Overall, these analyses showed that non-adherence to asthma treatment was greatly stable over time both intra and inter-individually.

Reliability over time

We estimated the reliability of the MARS-A10 and its subscales using the method for congeneric measures proposed by Raykov and collaborators (Raykov 1997; Raykov & Shrout, 2002) for congeneric measures. This method assumes that each individual item measures the same latent variable(s) with different scales (i.e. item loadings are not equal), with different degrees of precision (i.e. different item intercepts) and with different amounts of error (Graham, 2006). In addition, this procedure takes into consideration the presence of correlated errors as in the present case. We estimated the reliability of the full 10-item scale and both subscales (i.e. factors) using the loadings and errors from the last model. These analyses showed
that the reliability for the full scale ranged between 0.88 and 0.89 across the four waves (differences were due to the different error estimates of item 5 and the different estimates of the error correlation between item 1 and item 2). The reliability of the subscales was estimated in two different ways. First, consistent with common practice, we only used the items with the highest loadings on each factor. These results showed that the reliability for factor 1 based on four items ranged between 0.73 and 0.75. The reliability for factor 2 including the six items with the highest loadings ranged between 0.83 and 0.85. Because of the several cross-loadings, we also estimated the subscales reliability using the items with loadings statistically different from zero (i.e. six items for factor 1 and nine items for factor 2). These analyses showed that the reliability for factor 1 ranged between 0.73 and 0.75 and the reliability for factor 2 ranged between 0.78 and 0.79.

Discussion
In this study, we examined whether the MARS-A10 showed adequate psychometric properties in a longitudinal framework. Specifically, we provided important evidence demonstrating that the MARS-A10 was largely invariant over time. These findings have practical and theoretical implications for the study of medication non-adherence and illness self-management. The evidence supporting invariance in the structure, meaning, intercepts and item uniqueness indicate that score changes in the MARS-A10 over time can be correctly attributed to changes in non-adherence behaviour. This is critical for longitudinal studies interested in examining factors that contribute to changes in non-adherence and for interventions that focus on non-adherence either as an outcome or as a mediator. Results showing that factor determinacies were high for all waves indicate that factor scores can be a good alternative for studies when structural modelling is not used. This is especially relevant because given the presence of cross-loadings, sums may not be the best way to obtain total scores for each individual. The composite reliabilities for the full scale and the subscales were stable over time and their sizes were acceptable for research purposes (Nunnally & Bernstein, 1994).

The unexpected underlying factor structure of the MARS-A10 requires closer examination. Contrary to what one might have expected, the structure of the items did not fall into the intentional–unintentional non-adherence categories. One of the limitations of the version of the MARS-A10 used in this study is that unintentional non-adherence was only assessed with one item. This may have led to the identification problems we encountered when testing the three-factor model. It is also possible that self-presentation bias may have led participants to feel more comfortable reporting ‘forgetting’ than deliberately missing doses; resulting in the observed commonality among items assessing both intentional and unintentional non-adherence (Horne, Cooper, Gellaitry, Date, & Fisher, 2007).

These results suggest that the conceptualisation of intentional non-adherence may need to be reconsidered because it is not a unidimensional construct. The factor structure revealed two types of intentional non-adherence: (1) stimulus/symptom driven non-adherence, and (2) an intentional avoidance or reduction of medications. Studies examining the impact of commonsense beliefs on non-adherence behaviour indicate that the use of somatic stimuli as an assessment tool for determining the effectiveness and/or noxiousness of medical treatment is widespread among patients
with chronic conditions. For example, in their seminal work among hypertensive patients Meyer, Leventhal, and Gutmann (1985) found that the main reason given by patients who did not adhere to medical treatment was that it did not help relieve their symptoms. Similarly, those who adhere to medical treatment did so because they felt that their symptoms had improved. A study with asthmatic inpatients also showed that illness beliefs consistent with an acute model (i.e. asthma only manifests in symptomatic exacerbations) led to poorer adherence than a chronic model (Halm, Mora, & Leventhal, 2006). A recent study examining adherence to highly active anti-retroviral therapy revealed that HIV and treatment-related symptoms led participants to skip doses (Cooper, Gellaitry, Hankins, Fisher, & Horne, 2009).

The second factor reflects a deliberative behaviour that may reflect the patient’s weighing of different types of information (Wroe, 2002). This type of non-adherence may be more susceptible to the influence of beliefs about the necessity or importance of using the treatment, and the concerns about the treatment’s side or negative effects (Clifford et al., 2008; Horne, Cooper, & Fisher, 2008; Horne & Weinman, 2002). The current results showed that although both types of non-adherence overlapped they were largely independent. Future research needs to determine the extent to which these two factors are influenced by the same or different antecedents and whether they affect health outcomes in different manners.

The evidence that intentional non-adherence is a multidimensional construct poses interesting challenges for future research. A first critical issue relates to the conceptualisation of intentional non-adherence. Several studies have argued that intentional non-adherence is the result of a decision-making process that involves the weighing of pros and cons of medical treatment (Clifford et al., 2008; Wroe, 2002). However, the present data suggest that there may be at least one additional process involved in intentional non-adherence similar to that described by the commonsense model of self-regulation for the development of representations of health threats. Specifically, it is likely that patients determine the effectiveness of their treatment (or lack thereof) by comparing their somatic changes against illness and treatment prototypes (e.g. if symptoms subside then the treatment works). The feedback from this action and the meaning assigned to it by patients will motivate them to either keep using their treatment, alter the doses, or stop the treatment altogether.

A related challenge concerns the measurement of adherence with self-reports: is it possible to assess intentional non-adherence without the inclusion of the cues used for evaluating treatment efficacy, the time for observing effects or the weighing of pros and cons of a given treatment? Ideally, the measurement of non-adherence should reflect the pattern of treatment use (e.g. number of doses missed and number of days of perfect adherence in the past week), followed by questions about the reasons for non-adherence (e.g. did the patient miss the dose because she or he was feeling well?). However, because of social desirability and recall bias the results of these questions tend to misestimate non-adherence and the reasons underlying patients’ behaviours (Berg & Arnsten, 2006). The inclusion of instructions and questions that normalise non-adherent behaviour tend to provide better information to validly distinguish poor adherers (Wilson, Carter, & Berg, 2009). However, most of these questions include in the same items both reports of non-adherence and their causes which lead some participants to feel uncertain about how to respond if the question does not apply directly to them (e.g. ‘I only use it when I feel breathless’) (Horne & Hankins, 2008). We argue that the solution to this conundrum will depend on the purpose of the assessment. If the goal is to ascertain the prevalence of
medication use, then the questions that assess pattern of use will be more appropriate if measures are taken to prevent biases. Accurate estimates may be difficult with self-reports, and a combination of techniques may be needed to achieve this goal. For clinical purposes, the most relevant issue is to detect those patients who are poor adherers rather than the exact pattern of non-adherence, thus predictive validity will become the most important criterion to select the assessment tool. The MARS-A10 is a good choice for practitioners who deal with asthmatic patients as it has been validated against electronic monitoring (Cohen et al., 2008). Additionally, we have demonstrated that the MARS-A10 can be validly and reliably used for follow-ups to assess changes in patients’ behaviours. Finally for substantive research, both approaches have the potential to be useful if there is a strong theoretical framework guiding the study. The decision should be made on a case-by-case basis. For example, if adherence rates are high for a particular condition, separating non-adherence patterns from causes could lead to null results because of reduced variability (i.e. ceiling effects) and not because the underlying reasons associated with non-adherence are inconsequential (e.g. being symptom free). The inclusion of reasons, on the other hand, may create rather than assess cognitions if there is a poor understanding of the phenomenon (Ogden, 2003).

The evaluation of the present results must be tempered by the properties of this study. Because our sample consisted mainly of inner city Latino patients with an important burden of asthma morbidity, the generalisability of our findings to different settings may be limited. For example, among asthma patients with less severe asthma, non-adherence may not show the stability observed in this study. Patients with mild to moderate asthma may be more likely to believe that their asthma is an acute condition (Halm et al., 2006) and thus show a cyclical pattern of non-adherence consistent with symptom exacerbations. Future studies examining the trajectory of non-adherence over time need to include patients with different degrees of disease severity. Although the structure of the MARS-A10 is theoretically sound for intentional non-adherence, it is still unclear whether unintentional non-adherence is unidimensional or not. In the present study, only one item assessed unintentional non-adherence and it loaded together with items reflecting intentional medication avoidance. Future research will need to further explore the nature of unintentional non-adherence and develop more items to assess it. Additionally, as we did not originally expect to find two factors underlying intentional non-adherence, we did not plan on the examination of their discriminant validity. Our understanding of adherence behaviour will benefit from further research examining whether or not the two facets of intentional non-adherence represent different and distinct pathways linking psychological factors and health outcomes. Finally, it is possible that the mechanisms underlying missing data were not completely MAR and that unobserved non-adherence at the time of measurement may have played a role. However, because the putative cause of missingness (i.e. unobserved non-adherence) can be predicted by prior assessments, the potential bias resulting from violations of the MAR assumptions will have minimal impact on statistical conclusions (Collins, Schafer, & Kam, 2001; Schafer & Graham, 2002).

Acknowledgements
This study was funded by grants from the National Institute on Aging (RO1 HS09973 and 5R24AG023958), and the Agency for Healthcare Research and Quality (K08 HS013312).
We would like to thank Carmelen Chiusano, MPA, Jessica Lorenzo, MPH, Nicky O’Connor, MPH, Jessica Segni, Amy Badler and Julian Baez for their assistance on various aspects of the research reported in this article.

References


