Reliability, Validity and Factor Analysis of the Persian Academic Delay of Gratification Scale

Mehdi Arabzadeh\textsuperscript{20}, Parvin Kadivar\textsuperscript{21}, Ali Delavar\textsuperscript{22}, Javad Kavousian\textsuperscript{23}

Abstract
The aim of this study was to examine the psychometric properties of Persian form academic delay of gratification scale, including reliability, validity and factor analysis. Iranian students from Karaj Iran (N = 600, 326 boys and 274 girls,) completed the academic delay of gratification scale. Test-retest reliability academic delay of gratification scale based on the results of performance two tests and its internal consistency by Cronbach's alpha coefficients were calculated and confirmed. The purpose of concurrent validity was used motivation strategy learning questionnaire to indicate the concurrent validity is good. In addition, results indicated that the separate confirmatory factor analysis (CFA) for the male and female samples showed a good fit with the data. However; the scale is a valid and reliable instrument to measure student's academic delay of gratification. Implications for cross-cultural research are discussed.

Keywords: Academic delay of gratification, Reliability, Validity, Factor structure

1. Introduction
In the now classic delay of gratification paradigm, children are given the choice between accepting an immediately available small reward (e.g., one marshmallow) and a larger reward (e.g., two marshmallows) if they wait a given period of time (Ayduk \textit{et al.}, 2000; Mischel and Ayduk, 2002). Research using this technique, sometimes called the marshmallow test (Goldman, 1995) has examined the long-term developmental correlates of delay and the situational determinants that promote or hinder children's ability to delay gratification. Mischel \textit{et al.}, (1988), for example, reported that adolescents, who as preschoolers were able to delay gratification, were more academically oriented and socially competent as adolescents than were those participants who had succumbed to the immediately available smaller reward as children. In a subsequent follow-up longitudinal study, when those adolescents became 30 years old (Ayduk \textit{et al.}, 2000; Mischel and Ayduk, 2002), they were more able to delay gratification, which helped them cope with stress and frustration (Ayduk, 1999). Studies of situational determinants have examined how children represent stimuli during the delay period. For example, having children imagine marshmallows as clouds resulted in greater delay of gratification than when the children focused on consummator interactions with the stimuli, such as imagining how the food would taste (Ayduk, 1999). Distracting children during the delay period by providing those toys, or even having them imagine playing with toys, also increased their ability to delay. The ability to delay gratification or what in common parlance is called willpower is clearly a necessity in life because giving in to impulses during the present moment can jeopardize the chances of having a more rewarding future. Psychologists working in different areas (e.g., clinical, social, developmental, personality, and educational) have long recognized the necessity for delay of gratification and have found that this is beneficial for a variety of outcomes (Faber & Vohs, 2004).

\textsuperscript{20} Ph.D. student education psychology, TMU
\textsuperscript{21} Ph.D. education psychology, TMU
\textsuperscript{22} Ph.D. education, ATU
\textsuperscript{23} Ph.D. education psychology, TMU
Delay of gratification (DOG) applies because increasing the chances of accomplishing academic objectives that require extended time and effort (e.g., performing well on a course exam, completing an undergraduate degree) often means foregoing the reward of a more immediately available attractive activity or outcome (e.g., surfing the web or meeting friends). Within the field of educational psychology, delay of gratification has also received some attention.

Academic delay of gratification refers to students’ postponement of immediately available opportunities to satisfy their impulses in order to pursue important goals that are temporally remote but ostensibly more important (Bembenutty & Karabenick, 1998). It has been identified as a key component of self-regulated learning and has been associated with success in school and other positive educational outcomes (Bembenutty, 2010; Zhang, Karabenick, Maruno, & Lauermann, 2011; King & Du, 2011).

That hypothesis was supported given that ADOG was directly related to students’ reported use of cognitive strategies of rehearsal, elaboration, organization, and critical thinking assessed by the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993). ADOG was also related to students’ reported use of metacognitive strategies (planning, monitoring, and self-regulation; Pintrich & De Groot, 1990). In sum, students more likely to delay gratification are also those who self-regulate in order to maximize the likelihood of academic success, as well as more adaptively motivated (Bembenutty & Karabenick, 1998).

To measure this construct, Bembenutty and Karabenick (1998) developed the Academic Delay of Gratification Scale (ADOGS) and tested it on an American sample. However, the validity of the ADOGS needs to be assessed in a different cultural context aside from the U.S. Recently, there has been a call for researchers in the educational setting to be more sensitive to the cultural context. Therefore, there is a need to test for the validity of instruments developed from the West when applied to different cultural contexts (Maneesriwongul & Dixon, 2004).

It might be possible that some instruments developed in the West might not work properly in non-Western settings, thus it is imperative to test the construct reliability and validity of any instrument developed in the West before they are used in a new cultural context. In this study, we wanted to test the cross-cultural validity of the ADOGS among a sample of Persian students. In addition, we also used a more stringent criterion for construct validation. Bembenutty and Karabenick (1998) only relied on Cronbach’s alpha estimates and Pearson correlations with other related variables to examine the psychometric properties of the ADOGS.

Many of the theories and measuring instruments that have long dominated the psychological literature are based on Western values and research that may not be relevant to non-Western contexts (Markus & Kitayama, 1991). Thus it is imperative that instruments developed from one cultural context be validated before they are applied to a new cultural context. The issue of the cross-cultural applicability must be addressed. While it is an acceptable practice to adapt foreign-developed measures, cross-cultural researchers conducting studies with individuals from different cultural groups need to consider whether the scores obtained are comparable. Equivalence and bias are important issues that need to be addressed before meaningful cross-cultural comparisons can be made.

First, the present study calculated Test-retest reliability and internal consistency by Cronbach’s alpha coefficients academic delay of gratification scale. The purpose of concurrent validity was used motivation strategy learning questionnaire, in our study, we assessed how academic delay of gratification relates to other important educational variables such as self-regulated learning. In addition, using confirmatory factor analysis to test the
2. Methodology
2.1. Participants
Students from a high school in a city Karaj (Iran) were invited to volunteer to participate in the study (274 female students and 326 male students completed the questionnaires). Forty-two percent of the students were first year and others were from second and third year. The average age is 16 (SD = 1.62).

2.2. Instruments
2.2.1. Academic Delay of Gratification Scale:
The Academic Delay of Gratification Scale for students (ADOG) is a 10-item designed by Bembenutty and Karabenick (1998). The students rated their preference for an immediately available attractive option versus a delayed alternative. An example of the items is: “Go to a party the night before a test for this course” OR “Study first and party only if you have time.” Students responded on a 4-point scale: Definitely choose A, Probably choose A, Probably choose B, and Definitely choose B. Responses were coded and averaged across items so that the scores ranged from 1 to 4, with higher values indicating a greater delay of gratification.

This has shown evidence of both validity and reliability in Bembenutty and Karabenick (1998) studies. They reported consistency by Cronbach's alpha coefficients between 0.68-0.85 and concurrent validity was used motivation strategy learning questionnaire, were calculated and confirmed.

First; we describe how the English version was translated to and backtranslated from Persian. A committee approach was used in translating the original English version of the ADOGS into Persian. In this approach, a team of bilinguals is constituted to translate the items in the scale from the source to the target language. The members of the translation committee; spoke Persian as a first language and English as a second language. Members were Ph.D. The committee members debated and negotiated on the basis and merits of the specific versions of the translation for each item until a consensus was formed for each translated item.

The translation committee convened at a later time, but this time to undertake a back-translation of the Persian translation into English. Backtranslation is a highly recommended technique by experts in cross-cultural research (Maneesriwongul & Dixon, 2004), and it involves translating back into the source language version in order to verify translation of the research instrument. However, in this study, the committee approach was maintained, where each member was again first asked to independently back translate the Persian items into English, after which the committee convened again to compare, discuss, and form a consensus on the back-translation. During the committee work, some of the Persian translations were adjusted or revised after the English back-translation indicated some conceptual differences with the original English version. After the translation procedure was done, a small scale pilot study was conducted to assess the appropriateness of the scales.

2.2.2. Motivated Strategies for Learning Questionnaire (MSLQ)
The original Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1993) was also adapted for use with Iranian high school student. Translations and back translations were conducted by five experts in Persian and English. The MSLQ consists of two parts: motivation and self-regulated learning strategies. This 55-item questionnaire consists of 10 different scales, including self-efficacy, test anxiety, cognitive and metacognitive strategies (five dimensions) and resource management (three dimensions). This scale has been used...
with middle school students with internal reliability coefficients for the various scales ranging from 0.52 to 0.93 (Duncan & McKeachie, 2005). Between 2000 and 2004, the Motivated Strategies for Learning Questionnaire was used in 58 studies, 19 of which were conducted internationally (Duncan & McKeachie, 2005).

In addition to translation into Persian, there were two further modifications. First, the questionnaire was reduced in length to 47 items from the original 55 items. Second, instead of using all subscales, self-efficacy, test anxiety, intrinsic goal orientation, and cognitive strategy and Metacognition strategy, were chosen. In general, the subscales selected were those with simpler terms and syntax and judged more familiar to the student population. The response scales for the MSLQ ranged from 1 (not at all true of me) to 5 (very true of me).

That hypothesis was supported given that ADOG was directly related to students’ reported use of cognitive and metacognitive strategies assessed by the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1993; VanderStoep, Pintrich, & Fagerlin, 1996). ADOG was also related to students’ reported high self-efficacy, intrinsic goal orientation and low test anxiety (Zhang, Karabenick, Maruno, & Lauermann, 2011).

In sum, students more likely to delay gratification are also those who self-regulated in order to maximize the likelihood of academic success, as well as more adaptively motivated (King & Du, 2011).

2.3. Procedure

The ADOG and MSLQ scales were administered to students by the first listed author with the assistance of classroom teachers. The instructions included: “Below is a series of choices between two alternative courses of action. Please read each set of statements carefully? There is no correct answer, just choose the closest one to your own situation without thinking too much. Your answer has no influence on your grade, and your class teachers and classmates cannot see your responses. Please give the true response which can really reflect your situation”.

The presenter then made certain before proceeding that all students understood the instructions and how to respond. Following the instructions, the students were given 25-35 min to complete the survey.

2.4. Data analysis

We first assessed the descriptive statistics of the ADOGS and its Cronbach’s alpha as a measure of internal consistency reliability. Then we calculated Test-retest reliability. 65 Volunteers were randomly selected after four weeks then were tested again.

Second, to assess Concurrent validity, we looked at the correlations of ADOG with five types of motivation strategy learning (self-efficacy, test anxiety, intrinsic goal orientation, cognitive, Metacognition strategy and MSLQ).

Finally, confirmatory factor analysis to assess whether the hypothesized structure of the academic delay of gratification construct which Bembenutty and Karabenick (1998) proposed was applicable to the Iranian cultural context. It is assumed that the model has one latent factor with 10 indicator variables. Confirmatory factor analysis assesses the extent to which items reflect the underlying constructs. Model fit was assessed by a combination of model fit indices. In this study, we focused on the chi-square statistic and other goodness-of-fit indices such as the Goodness-of-Fit Index (GFI), incremental fit index (IFI), comparative fit index (CFI), adjusted Goodness of fit index (AGFA), normal fit index (NFI), root mean square error of approximation (RMSEA), and chi-square/degrees of freedom ratio. Its generally accepted that in good measurement models, the GFI, IFI, NFI, and CFI will be above 0.90 while the RMSEA will be below 0.08 (Byrne, 2001). The chi-square/degrees of freedom ratio should also be non-significant. However, researchers have found that this is usually overly sensitive to
sample size differences. Thus we decided to focus on the other fit indices to assess the fit of the hypothesized.

Finally, Multi-group confirmatory factor analysis was also conducted to assess whether the instrument worked in the same way for male and female students.

3. Results
3.1. Descriptive Statistics and Cronbach’s alpha

Table 1

<table>
<thead>
<tr>
<th>variable</th>
<th>M</th>
<th>SD</th>
<th>Cranach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic delay of gratification</td>
<td>2.40</td>
<td>0.55</td>
<td>0.71</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>3.14</td>
<td>0.48</td>
<td>0.62</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>1.16</td>
<td>0.36</td>
<td>0.58</td>
</tr>
<tr>
<td>Intrinsic goal orientation</td>
<td>2.84</td>
<td>0.89</td>
<td>0.69</td>
</tr>
<tr>
<td>Cognitive strategy</td>
<td>2.38</td>
<td>0.57</td>
<td>0.74</td>
</tr>
<tr>
<td>Metacognitive strategy</td>
<td>2.72</td>
<td>0.92</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Results indicated that the ADOGS had an acceptable reliability in the Iranian sample. The coefficient alpha generated in this study was comparable to what previous studies found in a sample consisting mostly of Caucasian Americans ($\alpha = 0.70$) (Bembenutty & Karabenick, 1998) and in another sample which was dominated by ethnic minorities in the U.S. such as Hispanic, African American, and Asian ($\alpha = 0.84$) (Bembenutty, 2010). The Cronbach’s alpha for the other scales was also acceptable (See Table 1).

3.2. Test-retest reliability

Test-retest reliability academic delay of gratification calculated on 65 volunteers was randomly selected after four weeks then were tested again (See Table 2).

Table 2

<table>
<thead>
<tr>
<th>variable</th>
<th>Test-retest reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic delay of gratification</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Calculated Scale reliability coefficient of 0.74, this reliability indicates good test-retest reliability coefficient the academic delay of gratification scale (Table 2).

3.3. Concurrent validity

To assess Concurrent validity we examined the relationship of the academic delay of gratification with motivation strategy learning (self-efficacy, test anxiety, intrinsic goal orientation, cognitive, Metacognition strategy). Results indicate that scores on ADOGS are significantly correlated to self-efficacy, intrinsic goal orientation, cognitive, Metacognition strategy. The correlations of the ADOGS with test anxiety non-significant. These results provide evidence for the Concurrent validity of the Persian translation of ADOGS (See Table 3).
Table 3

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic delay of gratification</td>
<td>.63**</td>
<td>.61**</td>
<td>.49**</td>
<td>.08</td>
<td>.49**</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>.78**</td>
<td>.76**</td>
<td>.41**</td>
<td>.12</td>
<td>1</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>.15</td>
<td>.13</td>
<td>.9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intrinsic goal orientation</td>
<td>.22**</td>
<td>.18*</td>
<td>.18</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Cognitive Strategy</td>
<td>.41**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metacognitive strategy</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P≤.05, **P≤.01

3.4. Confirmatory Factor Analysis

Researchers conducted a CFA with one latent factor and 10 indicator variables. They hypothesized that all the items of the 10-item ADOGS scale would load onto one overall academic delay of gratification factor. Results indicated that the fit indices were adequate.

Table 4

<table>
<thead>
<tr>
<th>Goodness-of-fit Indices for the CFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>42.57</td>
</tr>
</tbody>
</table>

This resulted in a more acceptable model with most of the goodness-of-fit indices reaching the criteria for a well-fitting model. The factor loadings were also all significant for this model (See Table 4).

Researchers tested the fit of the CFA model obtained above separately for the male and female samples (Table 5).

Table 5

<table>
<thead>
<tr>
<th>Goodness-of-fit Indices for the Male and Female CFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
</tbody>
</table>

Results indicated that the separate CFA for the male and female samples showed a good fit with the data (See Table 5).

4. Discussion

The purpose of the study was to assess the reliability and validity of the Persian version of the ADOGS. In general, the results show that ADOGS can be considered a reliable and valid instrument to measure the academic delay of gratification construct in the Persian cultural setting. The ADOGS has been used in countries other than the U.S. such as with Korean (Kim, Chung, Lee, & Kwon, 2001) and Chinese students (Zhang et al., 2011). It has also been used with ethnic minorities in the U.S. (e.g., Bembenutty, 2010). However, a stringent exploration of its psychometric properties in other cultures has not been explicitly examined.
before as most of these previous studies did not focus on examining the statistical properties of the questionnaire.

The confirmatory factor analyses indicated that ADOGS has one underlying latent construct, i.e. academic delay of gratification which can be measured by 10 indicator items. The fit indices for this model were shown to be acceptable. The factor loadings were all significant and internal consistency reliability was also acceptable. This study therefore presents stronger evidence that the ADOGS can be used to assess academic delay of gratification, at least in the Persian context. The items which were used to operationalize academic delay of gratification in Bembenutty and Karabenick’s (1998) original study among American students are shown to be applicable also in the Persian context. It appears that students in different cultural contexts face similar challenges and options in terms of choosing between temporally remote rewards that are more important versus rewards that are easily accessible but not that beneficial in the long run.

To assess Concurrent validity we examined the relationship of the academic delay of gratification with motivation strategy learning (self-efficacy, test anxiety, intrinsic goal orientation, cognitive, Metacognition strategy. Results indicate that scores on ADOGS are significantly correlated to self-efficacy, intrinsic goal orientation, cognitive, Metacognition strategy .The correlations of the ADOGS with test anxiety non-significant. These results provide evidence for the Concurrent validity of the Persian translation of ADOGS. Test-retest reliability academic delay of gratification calculated after four weeks, Scale reliability coefficient indicates good test-retest reliability coefficient the academic delay of gratification scale.

Separate CFA conducted among the male and female students indicated acceptable fit. Results of the confirmatory factor analyses indicate that male and female students answered the ADOGS in a similar manner. This is important because most studies looking at gender and self-regulated learning assess differences using a mean difference method. Such a method would not be warranted if males and females answered the questions differently. Indeed, a few previous studies have shown that some psycho educational constructs may function differently between genders (Kim, Kim, &Kamphaus, 2010 for an example). The measurement of academic delay of gratification in cross-cultural settings is imperative for educational psychologists whose aim is to improve classroom learning. Academic delay of gratification has been consistently linked to positive learning outcomes (Bembenutty, 2010; Kinget al., 2011; Zhang et al., 2011), thus the proper measurement of this construct in different cultural settings is a fruitful research agenda. The Persian version of the ADOGS can provide educational psychologists in Iran with a tool that is quite easy to administer (10-item test) and that can reveal an important dimension of students’ self-regulated learning. Academic delay of gratification is an important construct because it has an impact on the learning experience and academic success of students. Students who are able to engage in academic delay of gratification are more likely to have better educational outcomes. It might be useful for future interventions to teach students delay of gratification.
References


