All Good Things Come to Those Who Wait: Validating the Chinese Version of the Academic Delay of Gratification Scale (ADOGS)

Ronnel B King, The University of Hong Kong
Hongfei Du, The University of Hong Kong

Available at: https://works.bepress.com/ronnel_king/2/
All Good Things Come to Those Who Wait: Validating the Chinese Version of the Academic Delay of Gratification Scale (ADOGS)

Ronnel B. King  
*The University of Hong Kong, Hong Kong*

Hongfei Du  
*The University of Hong Kong, Hong Kong*

**Abstract**

The objective of this study was to examine the reliability and construct validity of the Chinese translation of the Academic Delay of Gratification Scale (ADOGS; Bembenutty & Karabenick, 1998). This scale measures the degree to which students postpone immediately available opportunities to satisfy their impulses in order to pursue important goals that are temporally remote but ostensibly more important. Chinese university students from Mainland China (N = 187) completed the ADOGS. Both within-network and between-network approaches to construct validation were adopted in the study. Responses to this questionnaire are shown to have good internal consistency reliability and support is provided for its construct validity in terms of its factorial structure and correlations with other educational outcomes. In addition, multi-group confirmatory factor analysis also indicated invariance of the instrument across males and females. Implications for cross-cultural research are discussed.

**Keywords:** academic delay of gratification, validation, China, self-regulated learning

Almost everyday people practice some sort of delay of gratification such as skipping sweets to keep a good figure, quitting smoking to have better health, shopping less to have more savings, and working hard to have a more secure future. 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; Metcalfe & Mischel, 1999).

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 & Karabenick, 1998; Zhang, Karabenick, Maruno, & Lauermann, 2011).

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 Chinese university students. In addition, we also used a more stringent criteria 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.

In the current study we adopt a more stringent criteria for construct validation utilizing both within-network and between-network construct validation approaches. Therefore the aim of the present study is to (a) assess the within-network construct validity of the Chinese translation of the ADOGS through confirmatory factor analysis, and (b) examine the between-network construct validity of the ADOGS by looking at its correlations with other theoretically-relevant constructs. An ancillary objective was to assess the equivalence of the Chinese version of the ADOGS across males and females.

Self-regulated Learning

Self-regulated learning refers to “self-generated thoughts, feelings, and actions that are planned and systematically adapted as needed to affect one’s learning and motivation” (Schunk & Ertmer, 2000, p. 631; Zimmerman, 2000). A key aspect of self-regulated learning is the experience of the self as an agent or origin in the learning. Self-regulated learning subsumes components such as goal-setting, learning strategies, self-monitoring, self-motivation, time management, help-seeking, self-efficacy, valuing of learning and also delay of gratification (Pintrich & De Groot, 1990).

Self-regulated learners know how to set personal goals for learning, concentrate on the classroom instruction, use effective strategies to organize, code, and rehearse information, establish a good working environment, manage time effectively, seek help when assistance is needed, and monitor their performance. More importantly, they know how to pursue goal-relevant behaviours and avoid off-task behaviours. Such forgoing involves delay of gratification in the academic domain, which Bembenutty and Karabenick (1998) referred to as academic delay of gratification. Pintrich (1999) argued that delay of gratification “is an important aspect of self-regulated learning as learners must often choose to focus their efforts on a learning activity in contrast to other more attractive options” (p. 342).

Academic Delay of Gratification

The academic delay of gratification construct is based on the work of Mischel and his colleagues (Mischel & Moore, 1973) who looked at effortful control among preschool children in what has now become known as the marshmallow test. In their experiments, they let children wait for two cookies (or other little treats) that they wanted and which they preferred to a smaller treat, such as one cookie. The children were then faced with the dilemma, because the experimenter left the room. They could either continue to wait for the experimenter to come back in order to get the larger treat or they could also
ring a bell to summon the experimenter at any time and immediately get the smaller treat at the expense of the larger reward. Children did not know how long the wait would be. If the child continued to wait, the experimenter would return spontaneously (after a maximum of about 20 minutes).

To attain greater specificity in the measurement of this construct, Bembenutty and Karabenick (1998) created the ADOGS to capture college students’ experiences of choosing between more important academic activities (e.g., “study a little every day for an exam in this course and spend less time with your friends”) or more immediately satisfying alternatives (e.g., “spend more time with your friends and cram just before the test.”). In this study, we will focus on assessing the psychometric properties of the ADOGS.

Recently, some researchers began to focus on delay of gratification among Chinese students (e.g., Yang & Wang, 2007; Yang, Xu, & Wang, 2003). However, almost all of these studies targeted children as participants in which the measurements or experimental manipulations cannot be used among adult samples. The ADOGS offers a viable alternative for measuring delay of gratification among adults in the academic setting. Hence, there is an urgent need to validate whether this instrument is applicable in the Chinese cultural context (Wong, 1999).

**Association between ADOG and Learning**

Academic delay of gratification has been found to be linked to students' motivational orientations (Bembenutty, 2010; Bembenutty & Karabenick, 1998; Zhang et al., 2011). For example, Bembenutty and Karabenick (1998) found that scores on the ADOGS correlated positively with intrinsic motivation, self-efficacy, and task value among university students. It was not significantly correlated with test anxiety. However, ADOGS was also found to be positively correlated with both self-perceived grades and actual final grades. In a more recent study, Zhang et al. (2011) found that among elementary school children, ADOGS correlated positively with intrinsic and extrinsic goal orientations, self-efficacy, task value, control beliefs, and grades. Aside from motivational orientations, ADOGS scores have also been found to correlate positively with a variety of adaptive learning strategies such as organization, elaboration, rehearsal, help-seeking, and effort regulation among others (Bembenutty & Karabenick, 1998; Zhang et al., 2011).

In this study, we tried to assess how academic delay of gratification would relate to goal orientations. Pintrich (2000) claimed that goal orientations are crucial to self-regulated learning. Goal orientations are conceptualized as the purpose of task engagement (Elliot & Church, 1997). The most common goals studied are: (1) mastery-approach goal, which refers to wanting to achieve to gain new knowledge and improve one’s competence; (2) performance-approach goal, which refers to wanting to achieve to outperform other students and demonstrate competence before others; and (3) performance-avoidance goal, which emphasizes the avoidance of showing incompetence relative to others (Elliot & Church, 1997).

The literature indicates that there is a good deal of converging evidence showing that mastery-approach goals are positively associated to self-regulated learning (Elliot, 2005). The relationship between performance-approach goals and self-regulated learning is more ambiguous. For example, some studies show
that performance-approach goals are positively related to self-regulated learning while some studies show that they are unrelated (Pintrich, 2000). Performance-avoidance goals have been consistently negatively related to self-regulated learning (Midgley & Urdan, 2001).

Bembenutty (1999) looked at the relationships between these three goal orientations and academic delay of gratification. He found that score on the ADOGS only correlated positively with mastery-approach goals. The correlations of the ADOGS with performance-approach and performance-avoidance goals were not significant. Therefore, in this study we also chose to look at the correlations of these goal orientations and grades to the scores in ADOGS as evidence of between-network construct validity.

The Role of Gender Differences

Recently, there has been a marked interest in gender differences in education with evidence pointing to declining levels of learning among males (Martin, 2003). Indeed, a good deal of data reflects significant differences in self-regulated learning as a function of gender (Meece, Glienke, & Askew, 2009). Gender differences have not been explicitly explored with regard to the academic delay of gratification construct, thus we decided to explore whether the ADOGS would be answered in a similar manner by males and females (see Bembenutty, 2009 for an exception).

Although previous studies on gender and academic delay of gratification have been conducted, they focused on mean-level differences and not on factorial equivalence (Bembenutty, 2007; Bembenutty & Karabenick, 1998). Assessment of gender differences in the present study provides an opportunity to explore the generality of the academic delay of gratification construct across males and females, which previous studies have neglected. It is common practice in educational research to pool together data from students of different genders. However, combining datasets together would only be warranted if invariance in terms of factor structure can be shown. Previous studies have given inadequate attention to the investigation of the equivalence of educational constructs to students of different genders, thus we wanted to investigate whether ADOGS has invariant factor structure across male and female students.

Cross-cultural Adaptation in the Chinese Context

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.

While adaptation of educational and psychological tests is a practice common to researchers, caution should not be taken lightly. Researchers
should carefully attend to reporting evidence of the reliability and validity of instrument if it is to be used in other cultural settings other than where it was initially developed.

The Present Study

The present study adopts a construct validation approach to the empirical assessment of the Chinese translation of the ADOGS (Marsh, 1997, 2002). Studies that adopt this approach can be classified as within-network or between-network studies. Within-network construct validation, also called internal construct validation refers to the examination of the factor structure and factor correlation matrix. It typically involves statistical techniques such as exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and reliability analysis. On the other hand, between-network or external construct validation approach entails examining patterns of relationships between the scales and other theoretically related constructs utilizing statistical techniques like correlational or regression analysis (Marsh, 1997). Few studies adopt this dual approach to validity thus providing relatively limited input into understanding the constructs being investigated.

In contrast, the present study uses both approaches. First, we conduct a within-network study using confirmatory factor analysis to test the factor structure of the ADOGS followed by multi-group confirmatory factor analysis to assess the invariance of the instrument across males and females. Consistent with the construct validation approach, it is not only important to address validity within an instrument (within-network validity) but it is also imperative to explore the possible differential relationships between the four types of goals and a set of theoretically relevant measures (between-network validity).

In our study, we assessed how academic delay of gratification relates to other important educational variables such as goal orientations and grades. We hypothesize that ADOGS would correlate positively with mastery goals (which is similar to intrinsic orientation to Bembenutty and Karabenick’s [1998] original study) and self-perceived grades. We expect it to be not significantly correlated with performance-approach and performance-avoidance goals.

Method

Participants

Students from a university in Guangzhou, Mainland China were invited to volunteer to participate in the study (105 female students and 82 male students completed the questionnaires). Seventy-three percent of the students were third year students and others were from first, second, and fourth year. The average age is 20.53 ($SD = 1.75$). Five students failed to report their age. All the students were Han Chinese.

Translation Procedure

First, we describe how the English version was translated to and back translated from Chinese. A committee approach was used in translating the original English version of the ADOGS into Chinese. In this approach, a team
of bilinguals is constituted to translate the items in the scale from the source to the target language. There were two members of the translation committee; the first member spoke English as a first language and Chinese as a second language while the second member spoke Chinese as a first language and English as a second language. Both members were Ph.D. candidates in fields related to psychology and education in a university where English was the medium of instruction.

The two members were tasked to develop a Chinese version of the ADOGS that closely expresses the psychological constructs originally expressed in English. This forward translation task was first done independently by each of the two-committee members. After the individual committee members completed their translations, the group met to compare and to discuss their outputs. 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 Chinese translation into English. Back-translation 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. Typically, different bilingual translators undertake the forward translation and back-translation. However, in this study, the committee approach was maintained, where each member was again first asked to independently back translate the Chinese 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 Chinese 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.

Instruments

**Academic Delay of Gratification Scale.** The Chinese translation of the ADOGS was given to the students to assess their academic delay of gratification. The ADOGS is a 10-item instrument where students rated their preference for an immediately-available attractive option versus a delayed alternative. For example, students were asked to choose between 2 options: “Study a little every day for an exam in this course and spend less time with your friends” or “Spend more time with your friends and cram just before the test.” 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 with higher scores indicating a greater preference for academic delay of gratification.

**Achievement goals.** The mastery goal (e.g., “One of my goals in class is to learn as much as I can.”), performance-approach goal (e.g., “One of my goals is to look smart in comparison to the other students in my class.”), and performance-avoidance goal (e.g., “It’s important to me that I don’t look stupid in class.”) subscales taken from the Chinese version of the Patterns of Adaptive Learning Survey (PALS, Midgley et al., 2000) were also administered to the
students as measures of between-construct network validity. The Chinese translation of the PALS has been validated in previous studies (Tao, 2000).

**Grade point average.** In Mainland China, most of universities adopt a five-category grading system to note student’s level: A (excellent, 85-100%); B (good, 75-84%); C (satisfactory, 65-74%); D (pass, 60-65%); G (failure, 0-60%). Students are familiar with this grading system and know which category their grades belong to. Considering that category G contains a too wide interval of grading, we divided the whole system into ten categories. Numbers from 1 to 10 refer to ten grading categories from the best to the worst. Specifically, 1 refers to top 10% in class, 2 refers to grades between top 10% to top 20% and the final 10 refers to bottom 10%. Participants were asked to select the number to indicate their achievement in class.

**Statistical Analysis**

We first assessed the descriptive statistics of the ADOGS and its Cronbach’s alpha as a measure of internal consistency reliability. Then we conducted 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 Chinese 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), Tucker-Lewis Index (TLI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and chi-square/degrees of freedom ratio. It is generally accepted that in good measurement models, the GFI, IFI, TLI, and CFI will be above .90 while the RMSEA will be below .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 (Anderson & Gerbing, 1988; Huang & Michael, 2000). Thus we decided to focus on the other fit indices to assess the fit of the hypothesized model.

Multi-group confirmatory factor analysis was also conducted to assess whether the instrument worked in the same way for male and female students. For the multi-group CFA, we followed a forward, stepwise approach, also called sequential constraint imposition (Dimitrov, 2010). In the present study, three levels of invariance were tested: configural invariance, measurement invariance, and structural invariance (Byrne, 2001). First, we tested for configural invariance which tests whether the number of factors and pattern of indicator-factor loadings is identical across males and females. The configural invariance model (Horn, McArdle, & Mason, 1983) is the model in which the same pattern of fixed (zero) and free factor loadings is specified for each of the two groups (males and females) and is considered the “minimal condition for factorial invariance” (Marsh, 1993, p. 851). It provides the basis for comparison with all subsequent models (Hong, Malik, & Lee, 2003).

Second, we tested for measurement invariance where the factor loadings were constrained to be equal. This was done after we have established the
configural invariance. Note that we only tested for invariant factor loadings in this stage and did not constrain the indicator intercepts to be equal because we were not interested in testing for differences in the latent means across the two versions.

Third, we tested for structural invariance where equality constraints were placed on the factor variances and covariances for the male and female groups. We did not test for equality of error variances and covariances across the two groups because such a test is “considered to be excessively stringent” (Byrne, 2001, p. 202). The classical approach in arguing for evidence of invariance is based on χ² difference (Bentler & Chou, 1987; Bollen, 1989); however, from a more practical perspective Cheung and Rensvold (2002) claimed that it is more reasonable to base invariance decisions on a difference in CFI. In line with this, we followed Cheung and Rensvold’s (2002) criteria indicating that a decrease of 0.01 in the comparative fit index (CFI) is evidence for lack of invariance.

Finally, as a test of between-network construct validity, we looked at the correlations of ADOGS with three types of achievement goals (mastery-approach goals, performance-approach goals, and performance-avoidance goals) and GPA.

### Results

#### Descriptive Statistics

**Table 1**

*Descriptive Statistics and Reliability Coefficients for the Various Scales*

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>M</th>
<th>SD</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic delay of gratification</td>
<td>1.00 - 4.00</td>
<td>2.32</td>
<td>0.45</td>
<td>0.70</td>
</tr>
<tr>
<td>Mastery-approach goals</td>
<td>1.00 – 4.00</td>
<td>2.25</td>
<td>0.44</td>
<td>0.79</td>
</tr>
<tr>
<td>Performance-approach goals</td>
<td>1.00 – 4.00</td>
<td>1.27</td>
<td>0.50</td>
<td>0.81</td>
</tr>
<tr>
<td>Performance-avoidance goals</td>
<td>1.00 – 4.00</td>
<td>1.36</td>
<td>0.56</td>
<td>0.80</td>
</tr>
<tr>
<td>GPA</td>
<td>1.00 – 10.00</td>
<td>4.21</td>
<td>2.45</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Note:** Higher scores indicated a greater endorsement of the item.

Results indicated that the ADOGS had an acceptable reliability in the Chinese sample. The coefficient alpha generated in this study was comparable to what previous studies found in a sample consisting mostly of Caucasian Americans (α = .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 (α = .84) (Bembenutty, 2010). The Cronbach’s alpha for the other scales were also acceptable (See Table 1).

#### Confirmatory Factor Analysis

Next, we conducted a CFA with one latent factor and 10 indicator variables. We 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 not that adequate. Based on the modification indices, we decided to correlate the error terms for items 1 and 4 (MI = 12.062). Items 1 and 4 both referred to forgoing social events in order to study.
example, in item 1, students were asked to choose between going to a favourite concert, play or sporting event and studying for the test. In item 4, students were asked whether they would prefer joining a party or studying for an exam tomorrow. Both items had somewhat similar content which justified the correlation of the error terms. This resulted in a more acceptable model with most of the goodness-of-fit indices reaching the criteria for a well-fitting model (See Table 2). The factor loadings were also all significant for this modified model. More importantly, the difference in chi square between the two models was significant (change in $\chi^2 = 20.654$, change in $df = 2$, $p < .001$). This indicates that correlating these two error terms resulted in a significant improvement in the fit.

Table 2  
**Goodness-of-fit Indices for the CFA**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>$df$</th>
<th>$\chi^2/df$</th>
<th>$p$</th>
<th>RMSEA</th>
<th>GFI</th>
<th>IFI</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original hypothesized model</td>
<td>71.23</td>
<td>35</td>
<td>2.04</td>
<td>.000</td>
<td>.08</td>
<td>.93</td>
<td>.86</td>
<td>.81</td>
<td>.85</td>
</tr>
<tr>
<td>Final model</td>
<td>50.57</td>
<td>33</td>
<td>1.53</td>
<td>.007</td>
<td>.06</td>
<td>.94</td>
<td>.91</td>
<td>.87</td>
<td>.90</td>
</tr>
</tbody>
</table>

*Note.* RMSEA = root mean square error of approximation, GFI = goodness-of-fit index, IFI = incremental fit index, TLI = Tucker-Lewis Index, CFI = comparative fit index.

**Multigroup Confirmatory Factor Analysis**

Before conducting a multigroup CFA, we first tested the fit of the final CFA model obtained above separately for the male and female samples. Results indicated that the separate CFAs for the male and female samples showed a good fit with the data (See Table 3).

Table 3  
**Goodness-of-fit Indices for the Male and Female CFAs**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>$df$</th>
<th>$\chi^2/df$</th>
<th>$p$</th>
<th>RMSEA</th>
<th>GFI</th>
<th>IFI</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>42.89</td>
<td>34</td>
<td>1.26</td>
<td>.14</td>
<td>.06</td>
<td>.90</td>
<td>.93</td>
<td>.89</td>
<td>.92</td>
</tr>
<tr>
<td>Females</td>
<td>40.90</td>
<td>34</td>
<td>1.20</td>
<td>.19</td>
<td>.04</td>
<td>.93</td>
<td>.95</td>
<td>.93</td>
<td>.94</td>
</tr>
</tbody>
</table>

*Note.* RMSEA = root mean square error of approximation, GFI = goodness-of-fit index, IFI = incremental fit index, TLI = Tucker-Lewis Index, CFI = comparative fit index.

Next, we conducted multi-group confirmatory factor analysis to test the equivalence of the measure across male and female students. There were 79 males and 105 females in this analysis (3 students failed to identify their genders, thus we did not include them in the analysis). Results indicated that the configural model where no invariance was imposed had good fit. After finding support for configural invariance, we then tested for measurement invariance by holding the factor loadings invariant across the two language versions. Results in Table 4 indicate that the factor loadings for the two measures were invariant given that the drop in CFI was $< 0.01$. In addition, the chi-square difference test between the two models was not significant (change in $\chi^2 = 9.45$, change in $df = 9$, $p > .05$) further supporting our conclusion that factor loadings were invariant for males and females. We then tested for a third model where we held both factor loadings and factor variances and covariances invariant. Again we saw evidence of invariance because the change in CFI was $< .01$. The chi-square difference
test was also not significant (change in $\chi^2 = 0.184$, change in $df = 1$, $p > .05$). Overall, we found evidence for invariant factor loadings and invariant factor variances and covariances. This indicated that male and female students in the sample answered the questionnaire in a similar manner.

Table 4

| Multigroup Confirmatory Factor Analysis across Males and Females |
|-------------------|---|---|---|---|---|---|---|---|
|                   | $\chi^2$ | df  | $\chi^2$/df | p   | RMSEA | GFI | IFI | TLI | CFI | Change in CFI |
| Configural invariance (no invariance imposed) | 83.81 | 68 | 1.23 | .09 | .04 | .92 | .94 | .91 | .93 | --- |
| Measurement invariance (invariant factor loadings) | 93.26 | 77 | 1.21 | .10 | .03 | .91 | .93 | .92 | .93 | .002 |
| Structural invariance (invariant factor variances and covariances) | 93.45 | 78 | 1.20 | .11 | .03 | .91 | .94 | .92 | .93 | .004 |

**Between Construct Network Validity**

To assess between-network construct validity we examined the relationship of the academic delay of gratification with different goal orientations (mastery-approach goals, performance-approach goals, and performance-avoidance goals) and self-perceived GPA. Results indicate that scores on ADOGS are significantly correlated to mastery goals and GPA only. The correlations of the ADOGS with performance-approach and performance-avoidance goals were non-significant. These results provide evidence for the between-network validity of the Chinese translation of ADOGS (See Table 5).
Table 5
Correlation of the ADOGS with Other Theoretically Relevant Constructs

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic delay of</td>
<td>.25***</td>
<td>.04</td>
<td>.13</td>
<td>.22**</td>
</tr>
<tr>
<td>gratification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastery goals</td>
<td>---</td>
<td>.15</td>
<td>.12</td>
<td>.17*</td>
</tr>
<tr>
<td>Performance-approach</td>
<td>---</td>
<td>.64**</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance-avoidance</td>
<td>---</td>
<td>.17*</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reports of GPA</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * p < .05, ** p < .01, *** p < .001.

Discussion

The primary objective of the study was to assess the validity of the Chinese 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 Chinese 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.

Within-network Construct Validity

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 (after correlating two error terms) 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 Chinese 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 Chinese context. It appears that college 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.

Between-network Construct Validity

In terms of correlations with other variables, we found that scores on the ADOGS were positively correlated with mastery-approach goals but not significantly correlated with performance-avoidance goals. These results corroborate evidence in the literature indicating that mastery-approach goals are positively linked to self-regulated learning, while performance-avoidance goals are negatively linked to it (for review see Pintrich, 2000b). Interestingly, we
found that performance-approach goals were not significantly correlated to scores on the ADOGS. As earlier research has suggested, the nomological network of performance-approach goals is somewhat more ambiguous compared to mastery and performance-avoidance goals. The literature is clear that mastery-approach goals produce positive outcomes while performance-avoidance goals are linked to negative outcomes (for review see Maehr & Zusho, 2009). However, with regard to performance-approach goals, the evidence is more mixed. Some studies indicated that it is linked to positive outcomes such as grades, while other studies have found that it is not significantly correlated with learning strategies (Maehr & Zusho, 2009).

The results of the present study indicate that performance-approach goals are not significantly related to academic delay of gratification. This implies that wanting to achieve to demonstrate your competence or to show that you are smart before others does not seem to be related to delaying gratification in the academic setting. These results corroborates the thesis of the traditional mastery goal hypothesis camp which argued that performance-approach goals are not beneficial for learning outcomes (Midgley, Kaplan, & Middleton, 2001; see Harackiewicz, Barron, Pintrich, Elliot, Thrash, 2002 for a contrasting opinion).

In terms of grades, we found that academic delay of gratification was positively correlated with grades. This corroborates the findings of Bembenutty and Karabenick (1998) who found that American students’ scores on ADOGS were positively related to GPA. This makes theoretical sense because students who are able to delay their gratification more are also more likely to get higher grades. They are capable of engaging in behaviours that would foster achievement such as refusing to party with friends in order to prepare for a test. This also corroborates previous research showing that delay of gratification is closely linked to academic achievement (Mischel, Shoda, & Peake, 1988).

**Gender Invariance**

Separate CFAs conducted among the male and female students indicated acceptable fit. Results of the multigroup 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 (see Kim, Kim, & Kamphaus, 2010 for an example).

Thus far, our study is the first to provide more stringent evidence that the academic delay of gratification construct is invariant across genders. This provides valuable information to both researchers and practitioners who are interested in studying self-regulated learning and in designing interventions to improve this capacity. A great deal of data analysis in self-regulated learning routinely pools male and female students. One important feature of a dataset that determines whether analysis should be disaggregated by gender is the extent to which the factor structure underpinning the dataset is invariant across male and female students.

The present analyses showed that the factor structure of the ADOGS is in fact invariant across gender providing support for the common practice of
aggregating male and female students in subsequent analysis. The invariance of the factor structure for ADOGS also holds some implications for educational interventions justifying the use of similar interventions for students of different genders.

**Educational Implications**

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; Bembenutty & Karabenick, 1998; Zhang et al., 2011), thus the proper measurement of this construct in different cultural settings is a fruitful research agenda. The Chinese version of the ADOGS can provide educational psychologists in China 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.

**Limitations, Future Research, and Conclusion**

This research has some limitations. First, the sample size on which the current study was based is small. Second, we only tested for the correlation of ADOGS with goal orientations and self-reported grades. Future research could test the relationship of ADOGS to a wider range of psycho-educational constructs similar to what Bembenutty and Karabenick (1998) in their study who looked at how ADOGS related to various kinds of learning strategies (e.g. rehearsal, elaboration, organization, and critical thinking) and a wider range of motivational constructs such as self-efficacy, intrinsic motivation, and extrinsic motivation. Third, we used self-report GPA rather than GPA obtained from the school records. However, the use of self-report grades is common in educational research (see Bembenutty, 1999 for an example). Fourth, we did not take into account the nested nature of the data. Future research with a larger sample size and a bigger scope could incorporate multilevel modelling to determine classroom effects.

In sum, the results of this research support the internal consistency reliability and construct validity of the Chinese version of the ADOGS developed in this research. An advantage of the present study was the use of both within-network and between-network approaches to construct validation. This presents a stronger case for the validity of the instrument. Future research is needed to assess the relationship of ADOGS to other outcome measures in both Western and non-Western settings.
References


## Appendix

请仔细阅读每个题目中的两个关于学习兴趣的陈述，根据你的实际情况，选择其中一个，并将字母 A 或 B 写在空格内，然后再选择后面的数字用来表达赞成的程度，在相应的数字上打“√”。

<table>
<thead>
<tr>
<th>题号</th>
<th>题项</th>
<th>同意 A</th>
<th>非常赞成</th>
<th>比较赞成</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A. 宁愿参加运动会、聚会或去听音乐会，而少学一点东西，即使明天考试也没有关系; B. 为了明天考试取得好成绩，决定在家里学习</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>A. 为了准备考试，每天都复习一些，减少与朋友在一起的时间; B. 花更多的时间与朋友在一起，对于考试则是临时才复习</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>A. 为了参加一个非常感兴趣的娱乐活动而逃课; B. 直到课程结束后才去参加娱乐活动</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A. 在考试的前一天晚上，去参加聚会，若有时间才复习; B. 学习第一，若有时间才去参加聚会</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>A. 喜欢在大部分时间学习课程中自己感兴趣的内容，即使学好这部分内容并不意味着自己学好了整门课程; B. 学习老师布置的所有内容，从而使自己更可能在这门课程上获得好成绩</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>A. 天气好的时候，出去玩，之后再自己抄一下笔记; B. 即使外面天气很好，自己也会坚持上课，确保不会错过某些知识</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>A. 喜欢在愉快的环境学习，即使环境令人分心; B. 喜欢在安静的地方，复习功课，以便学得更好</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>A. 下课后会马上去做自己喜欢做的事，即使可能没有理解考试会考到的内容; B. 下课后留在班级，问老师那些考试中出现的自己不懂的内容</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>A. 宁愿选择一个有趣的老师，即使他/她的讲课水平不高; B. 宁愿选择一个不太有趣的老师，但是他/她的讲课的水平很高</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

### About the Authors

**Ronnel B. King** is a Ph.D. candidate in educational psychology at the University of Hong Kong. His research interests are in achievement motivation and academic engagement. Email ronnel.king@gmail.com

**Hongfei Du** is a PhD candidate in psychology at the University of Hong Kong. His research primarily focuses on self-esteem and culture. Correspondence concerning this article should be addressed to Hongfei Du, Faculty of Education, The University of Hong Kong, Hong Kong. E-mail: hongfei@hku.hk