This study examined the psychometric properties of the Korean version of the Beck Depression Inventory (K-BDI) for a Korean high school-aged population within the framework of item response theory (IRT). The specific goals were as follows: First, this study tested for the presence of essential unidimensionality of the K-BDI. Second, this study investigated for the evidence of the unequal discriminating ability of the K-BDI items with respect to depression severity. Third, this study examined the correctness of the ordering of item categories in measuring depressive severity. Finally, this study investigated the effectiveness of the K-BDI as an entire test for screening for depression among Korean high school-aged population. With the permission of Beck, a K-BDI was developed by the researcher using a back translation method (Brislin, Lonner, & Thorndike, 1973). For this study, 1497 students attending three senior high schools in a large metropolitan city in South Korea were recruited. The generalized partial credit (GPC) model (Muraki, 1992, Muraki, 1997) was chosen to depict the data and the PARSCALE computer program version 2.2 of this model (Muraki & Bock, 1991) was used for parameter estimation. The IRTGRAPH (Child & Schlumpf, 1999) was used to transform parameters into the graphs of item category response functions and item information functions. The results indicated that the K-BDI possessed essential unidimensionality. Items having greater discriminating ability were 1, 4, 5, 7, 15, and 17 while items possessing less discriminating ability were 18, 19, and 21. The third category for most items might not contribute much to the identification of depression severity among this population. The ordering between the third and the fourth categories were reversed for all 21 items. The K-BDI as an entire instrument offers greater information in distinguishing among individuals at $\theta = 2.11$, which corresponds roughly to the raw total score of 30. Accommodations of cut-off scores were discussed. Recommendations for revising the instrument conducting IRT-based analyses were also discussed. Limitations of this study and suggestions for future research were discussed.

Key words: K-BDI, item response theory, discriminating ability, response category effectiveness, test information, cut-off score
The Beck Depression Inventory (BDI) is one of the most widely used instruments in research and clinical settings for measuring depression in clinical and nonclinical populations of both adults and adolescents. While the majority of studies on the psychometric properties of the BDI were conducted within the framework of classical test theory (CTT), the BDI also has received attention over the past two decades from researchers using item response theory (IRT) models (Bouman & Kok, 1987; DeRoos & Allen-Meares, 1992; Frick, Rehm, & Thien, 1997; Grayson, Henderson, & Kay, 1987; Keller & Kempf, 1997; Mehta, 1996; Santor, Ramsay, & Zuroff, 1994; Santor & Ramsay, 1998). Studies using IRT have conducted more refined item analyses than those allowed with CTT and have resulted in recommendations for revising the instrument. Beck and his colleagues (Beck, Steer, & Brown, 1996) published a second edition of the BDI, incorporating the results of their IRT investigations along with their clinical and research experiences (Dozois, Dobson, & Ahnberg, 1998).

Issues regarding dimensionality, unequal discriminating ability of items and tests at different levels of a trait, and response category (option) effectiveness are critical to the precision of every psychological measurement. The application of IRT methods has proven useful in addressing these psychometric issues for many psychological measures, including measures of depression. There has been much confusion regarding the dimensionality of the BDI due to complex conceptual and methodological issues (Beck, Steer, & Garbin, 1988; Bouman & Kok, 1987; Tanaka & Huba, 1984). Additionally, there is evidence that some BDI items and response categories are not sensitive to changes in depression severity (Mehta, 1996; Santor, Ramsay, & Zuroff, 1994). For example, some symptoms or items may be good indicators of depression at severe levels, but not as good at moderate levels. IRT methods have proven more successful in identifying and resolving such issues than other methods.

Research and clinical practice in the area of adolescent depression have been hindered by the lack of culturally appropriate measures for adolescent depression. In Korea, the BDI has been used widely in research and clinical settings for both adults and adolescents. Validity studies using versions developed on CTT have been conducted with both Korean adults and adolescents (Lee, Lee, Park, Son, Jeong, & Hong, 1995; Park, Kim, Kim, & Jeon, 2000). However, little has any Korean version of the instrument been subjected to IRT analysis to investigate the IRT-based psychometric properties when used with Korean adolescents.

This study, therefore, subjected a Korean version of the BDI (K-BDI) to IRT analysis to provide a population-specific, culturally appropriate version of this instrument for use with Korean high school-aged adolescents in research and clinical settings. Because this study was exploratory in nature, no formal hypotheses were tested. However, four research questions were investigated: First, does the K-BDI present a degree of essential unidimensionality that would justify IRT analysis? Second, do the 21 items on the K-BDI have differential discriminating ability across different levels of depression severity, and, if so, which have more and which have less...