1. **Introduction**

Personality, besides cognitive ability, has been recognized as a most important predictor of educational outcomes, such as academic performance (AP) and absenteeism. Among the five-factor mode (FFM) of personality (Costa & McCrae, 1992; Goldberg et al., 2006), conscientiousness, either the broad dimension or the narrow facets of it, has been found by a lot of studies to correlate strongly with AP (e.g., Poropat, 2009; McAbee & Oswald, 2013; Morris & Fritz, 2015), and absenteeism (MacCann, Duckworth, & Roberts, 2009). Although the relationship between the personality dimension and absenteeism at workplace has been thoroughly studied (e.g., Judge, Martocchio, & Thoresen, 1997; Hattrup, O’Connell, & Wingate, 1998; Vlasveld et al., 2013), only few studies can be found assessing this relationship at school. Also, a very large proportion of the research involving the relationship between conscientiousness and AP was conducted within a single country or culture. In a cross-cultural setting, few studies were conducted investigating and comparing the performance of conscientiousness as a predictor of AP and absenteeism. Data collected from 68 countries and regions for the 2012 Program of International Student Assessment (PISA 2012) has provided us insight into how conscientiousness predicts educational outcomes in different cultures.

In a cross-cultural setting, comparisons of correlations and mean scores will be meaningless and misleading if the measurement is lacking equivalence (Church, 2001; Drasgow & Kanfer, 1985). Without measurement equivalence (ME), it’s hard to know if the observed mean score difference is due to true cultural difference or the relationships that vary across groups between the latent variable and the observed score (Raju, Laffitte, & Byrne, 2002), and therefore, the differences may be attributed to inaccurate cultural stereotypes (Church, 2001).

We believed that by conducting comparisons across all 68 countries and regions, the differences would be overstated between countries in the same region or are influenced by the same culture, with the analyses being too complicated, while the results being scattering.

Therefore, in the current study, we used PISA 2012 data to investigate how well conscientiousness predicts math achievement and truancy across cultures. We first made 9 groups out of the 68 countries and regions, based on geographic proximity, degrees of economic development, and cultural influence such as prevalent philosophy (e.g., Confucianism in East Asia), and religions (Saucier et al., 2015), and then conducted the analyses using structural equation model using the partial invariant conscientiousness measurement, so that comparisons across groups were meaningful and justified.

* 1. *Conscientiousness and academic performance*

Conscientiousness is a personality construct that reflects individual differences in characteristics such as being diligent, organized, rule abiding, self-controlled, and responsible to others (Roberts, Lejuez, Krueger, Richards, & Hill 2014; Roberts, Jackson, Fayard, Edmonds, & Meints, 2009). According to Poropat (2009), conscientiousness is the FFM dimension that’s the most closely related to AP. Conscientiousness resembles the most Webb’s *w* factor (Webb, 1915), or as in Digman (1989), the willingness to achieve, both of which were found linked to AP (De Raad & Schouwenburg, 1996). Poropat (2009) also pointed out that due to the link between conscientiousness, sustain effort, and goal setting (Barrick, Mount, & Strauss, 1993), the dimension also contributes to a variety of AP-related behaviors, such as concentration on homework and following the requirement (Trautwein, Ludtke, Schnyder, & Niggli, 2006), and time management and effort regulation associated with learning (Bidjerano & Dali, 2007).

The importance of conscientiousness in predicting AP emerged from studies assessing personality traits’ prediction of AP. Early studies (e.g., Harris, 1940; De Raad & Schouwenburg, 1996) on this topic found inconsistent patterns, largely due to the lack of an overarching framework or paradigm, and flawed methodologies (Poropat, 2009). Therefore, Poropat (2009) meta-analyzed the relationship between personality and AP by adopting the FFM, which the author believed was both comprehensive and parsimonious. The FFM personality consists of 5 broad dimensions, which are Agreeableness, Conscientiousness, Emotional Stability, Extraversion, and Openness. In Poropat’s (2009) pivotal meta-analysis, conscientiousness stood out as the best predictor of AP. The correlation between conscientiousness and AP held even when intelligence was controlled for. When secondary academic performance was controlled for, conscientiousness contributed as much to the prediction of tertiary performance as did intelligence. McAbee and Oswald (2013) in their meta-analysis tested the assumption of previous meta-analyses that different personality measurements all tapped into the same personality trait and therefore predict outcomes in similar ways. What they found was that conscientiousness, regardless of the specific FFM personality tests used, consistently exhibited the strongest criterion-related validity for GPA prediction, compared with the other four personality traits. Apart from GPA and exam grades, conscientiousness was found to be related even more strongly to students’ grades of overall coursework (Morris & Fritz, 2015). In terms of prediction of performance on math or scientific subjects, Steinmayr and Spinath (2007), by using a German adolescent students sample, found conscientiousness being the only consistent predictor of German, math, and GPA. Similarly, using a Slovenia sample, Peklaj, Podleske, and Pečjak (2015) proved that conscientiousness was the most important predictor of math grades of secondary school students, both directly and indirectly. In 2016, Dumfart and Neubauer studied a sample of Australian eighth-grade students, and concluded that conscientiousness, among the Big Five, and other traits such as grit, self-efficacy, and self-discipline, was the single most powerful noncognitive predictor for grades on language subject, GPA, and scientific subjects, including math. Zhang and Ziegler (2016), in a sample of Chinese secondary school students, reported a significant mediator, a deep-learning approach, for the relation between conscientiousness and math achievement. The discoveries of the strong and consistent association between conscientiousness and AP, especially math grades, have led us to be curious about the similarities and differences of such a relationship across cultures, when equivalent measurements are used.

* 1. *Conscientiousness and absenteeism*

Many studies have looked at relationships between conscientiousness and employee absence form work, because work place absenteeism can be economically costly to organizations and employers. Not surprisingly, theses studies almost consistently support the conclusion that conscientiousness is a predictor of absenteeism from work (e.g., Vlasveld et al., 2013; Judge et al., 1997; Hattrup, O’Connell, & Wingate, 1998). But we were only able to find one study evaluating the correlation between conscientiousness and absenteeism among students. MacCann and colleagues (2009) recruited 291 American high school students, and was able to extract 8 facets based on their responses to Conscientiousness items from different models, which were Industriousness, Perfectionism, Tidiness, Procrastination Refrainment, Control, Cautiousness, Task Planning, and Perseverance. Among the eight facets, Industriousness demonstrated significant prediction of absenteeism. This study demonstrated the importance of using narrow facets rather than a broad dimension of personality, and that of taking various personality models based on the Big Five (MacCann et al., 2009).

The 5-item Conscientiousness scale in the PISA 2012 is called “Perseverance”, and covers items belonging to 4 of the 8 facets extracted in MacCann et al. (2009). These 4 facets are industriousness (“Exceed expectation”), perfectionism (“Continue to perfection”), procrastination refrainment (“Put off difficult problems”), and perseverance (“Give up easily”; “Remain interested”). Therefore, with PISA 2012, we would be able to assess the how narrow conscientiousness facets deriving from various models connect to school absenteeism in a cross-cultural setting.

* 1. *Measurement equivalence of PISA*

The Program for International Student Assessment (PISA) is a survey developed by the Organization for Economic Co-operation and Development (OECD), and conducted in 3-yearly cycles across more than 80 countries, aiming to evaluate education systems worldwide by testing the math, reading, and science skills and knowledge of 15-year-old students (https://www.oecd.org/pisa/aboutpisa/). During the past 16 years, PISA has had great impact on not only scholarly community, mass-media, and public, but also national and international policy makers (Kankaraš & Moors, 2014). However, as important as ME is to meaningful mean score comparisons acorss cultures (Church, 2001), studies using the PISA data rarely test it, but instead assuming that ME exists cross-culturally. But in fact, to test for ME in PISA measurements, we are not being paranoid.

According to Drasgow (1984), ME was obtained when participants from different groups had the same expected observed score as long as they were at the same latent trait level, and measurement non-equivalence is prevalent in cross-cultural personality measurement (Church, 2001), majorly due to three types of bias: construct bias, method bias, and item bias (Van de Vijver & Tanzer, 1997). Construct bias refers to definitions of behavioral examples of the measured construct not being completely the same across cultures. Method bias consists of bias existing in sampling (e.g., different education levels or familiarity with the procedure of assessment across cultures), instrument (e.g., different response styles acorss cultures), and administration (e.g., misunderstandings between the assessor and the participants). Item bias, also known as differential item functioning (DIF), exists at the item level, and likely be due to translations nonequivalence, or the inclusion of culturally nonequivalent items (Church, 2001). Any combination of these three types of bias will likely lead to misinterpretation of the test, what Dana (2000) called the interpretation bias (Church, 2001). Whenever any of these factors presents in the measurement, direct comparison acorss cultures is unjustified and any differences discovered among mean scores could be induced by test bias, instead of true cultural differences. Over the years, although ME in the earlier PISA (before 2012) has been questioned by some studies (e.g., Kankaraš & Moors, 2014; Goldstein, 2008; Yildirim, 2006), the tests were majorly on the achievement tests, rather than the personality measurements, and ME of PISA 2012 is barely studied. Therefore, we considered it necessary to first establish a satisfactory degree of ME over the personality measurement of conscientiousness (or perseverance, as named in PISA 2012), before any further analyses was carried out.

The most widely used approach to studying ME is Confirmatory Factor Analytic (CFA) mean and covariance structure (MACS), which examines whether a common factor model exists across groups (Raju et al., 2002) and focuses on testing three levels of measurement invariance, which are configural, metric, and scalar invariance (Vandenberg & Lance, 2000). According to Horn and McArdle (1992), configural invariance should be achieved before the other two types of measurement invariance can be teste. Configural invariance is the weakest type of ME, and is demonstrated when the measurement has same number of factors and the same pattern of salient loadings acorss cultures (Church et al., 2011). Metric variance refers to factor loadings being invariant across cultures (Church et al., 2011), and scalar invariance, the most stringent form of invariance of the three, implies that when items regress on latent variables, they have the same intercepts across groups (Church et al., 2011; Steenkamp & Baumgartner, 1998). Scalar invariance is necessary to meaningful cross-cultural comparisons of mean scores (Church et al., 2011).

In the current study, before assessing relationships of conscientiousness with AP and absenteeism, ME of the conscientiousness scale (i.e., perseverance) was first tested, because only with a reasonable degree of measurement equivalence, could we conduct further analyses.

* 1. *Grouping strategies and ecological fallacy*

Instead of comparing all 68 individual countries and regions that participated in PISA 2012, we grouped them into 9 major groups based on a combination of geographic proximity, economic development, and cultural background (Saucier et al., 2015). We discussed more about grouping in the Method section. We believe that this is a good approach to represent different cultures in PISA without overstating trivial differences between individual countries.

Ecological fallacy refers to the failure to recognized the possibility that correlations can be different at the individual and aggregated levels, and PISA is no stranger to this kind of discrepancy. In PISA 2003, the mean within-country correlation between math self-concept and math achievement was 0.4, while across countries, the correlation was -0.2 (Kyllonen & Bertling, 2013). Van de Gaer, Grisay, Schulz, and Gebhardt (2012) found in PISA 2006 the same pattern of reversal in within-school and across-school correlations between science self-concept and science achievement, and believed that the reference group effect was the cause. They concluded the Likert-type scale used in PISA actually reflected not only absolute values, but also relative ones. Respondents from different cultures tend to compare themselves to different references and standards when responding on a Likert-type scale, and the frames of reference are believed to be influences by norms within but not necessarily across cultures (Duckworth & Yeager, 2015). That is to say, ecological fallacy may not have to do with true cultural differences on the construct measured, but rather a type of bias (Dana, 2000) possibly induced by language, culture, economic development, or other factors that are not the purpose of the study, but contribute to the form of different references. Therefore, if the 9 major groups we have in the current study truly captured cultural similarities among countries in the same group, and discriminated countries influenced by different cultures, then we may expect the absence of reference bias at the level of major groups, which means the correlations between our variables should be in the same direction both within-country and within-group, while correlations in the opposite direction may appear across countries and across the 9 major groups, given the evidence that reference bias does exist in PISA (Van de Gaer et al., 2012).

To summarize, in the current study, we built our comparisons and interpretation of cross-cultural relationships of conscientiousness with math achievement and absenteeism on a grouping strategy that allows for cultural representation and analyses parsimony, and on measurement equivalence of the personality measurement, which used to be ignored or assumed by the PISA literature.