“Big data” versus “small data” in social sciences: A reflection of cross-cultural studies on structures of character strengths

Wenjie Duan, Wenlong Mu and He Bu
Wuhan University, China

Abstract
Culture significantly influences human cognition and behavior, which has become a major obstacle in cross-cultural studies. Comparing traditional sampling studies (i.e., small data research) with the novelty of millions of samples studies (i.e., big data research), we suggest that the results of the finely controlled, precisely sampled, and accurately analyzed theory-driven small-data research can be replicated by big data studies. This conclusion has been illustrated by recent studies on structures of character strengths that were conducted in both western and eastern countries. Therefore, big data studies that take into account both emic and etic components will be an important approach to conduct cross-cultural research. It facilitates the construction of theories and measures with cross-cultural consistency. Nevertheless, it should be noted that “small data” and “big data” studies are complementary and should not be treated as substitutes for one another.

Keywords
big data, cross-cultural research, character strength, combined emic-etic approach

Corresponding author:
Wenjie Duan, PhD, Associate Professor, Department of Sociology, Wuhan University, 299 Bayi Road, Wuchang District, Wuhan, Hubei 430072, China.
Emails: duan.w@whu.edu.cn; duan.w@outlook.com
Introduction

Comte (1880) once stated that the ultimate human science was the study of social behaviors of an organism with social characters in a cultural context (Yu, Peng, & Zheng, 2015). This statement not only indicated the sociality and biology of human beings, but also recognized that we should pay adequate attention to the cultural dependency of social behaviors. There is no doubt that culture plays an important role in social science research (DiMaggio, 1997). Individuals of various cultural contexts are dramatically different in social cognition as well as behavior. According to the definition of culture proposed by Weber and Swidler, Polavieja (2015) defined culture as “probabilistic tendency”, which means that individuals living in a given culture have a high probability of some specific behaviors. Furthermore, even though culture is a collective phenomenon, it manifests itself only through individual behaviors (Polavieja, 2015). However, many empirical studies have demonstrated that individuals under different cultures tend to show different characteristics, cognitive processes and typical behaviors (Dahlsgaard, Peterson, & Seligman, 2005; Heine & Buchtel, 2009; Schimmack, Radhakrishnan, Oishi, Dzokoto, & Ahadi, 2002). Accordingly, how to develop theories and measures with cross-cultural consistency, through traditional research paradigms and emerging technological approaches, becomes a very challenging but significant research question. Furthermore, the dialogues between West and East have been a great potential positive force in increasing international understanding (Lin & Palmer, 2016).

As an extremely complicated and dynamic concept, culture has challenged and will continue to challenge cross-cultural research. Traditional research paradigms employ theory-driven approaches, such as the etic approach, the emic approach, and the combined emic-etic approach; however, big data, as a new research approach, is data-driven. In this article, we aim to clarify the role of theory-driven small data and data-driven big data approaches in cross-cultural research. The empirical studies on the structures of character strengths in positive psychology will be taken as an example to reveal the probable effect of big data in cross-cultural research. We expect our findings can provide insight for future cross-cultural study.

Traditional paradigms of cross-cultural study

Etic approach and emic approach

The etic approach and emic approaches are common theory-driven research paradigms in modern cultural anthropology. In the area of cross-cultural research, the etic approach is a more common paradigm. The etic approach is to understand culture from the perspective of the outside observer and explain the reasons and results of culturally patterned behaviors with scientific standards (Harris, 2001). The etic approach emphasizes the universality and popularization of culture. It assumes that those theories as well as measures rooted in western culture, after being translated, can be applied directly to the target groups in non-western cultures, and vice versa. For instance, the subjective well-being theory and its main measure (i.e. Satisfaction with Life Scale, SWLS) was developed in the culture of individualism and considered as a construct with
cross-cultural consistency. As a result, it is widely used in different countries and regions of the world (Diener, Emmons, Larsen, & Griffin, 1985; Pavot & Diener, 1993; Pavot, Diener, Colvin, & Sandvik, 1991; Shevlin, Brunsden, & Miles, 1998).

A premise to conduct effective cross-cultural comparison in multicultural groups is to ensure the equivalence of concepts and measurements. The most frequently-used strategy for guaranteeing the equivalence is “translation and back-translation” approach (Brislin, 1970). After translations, the adapted theories, concepts and measures are recognized to be applied to other cultures (Brislin, 1970). It should be noted that the premise of this approach is that the results obtained from such theory and measures are comparable among groups from different cultures. However, each cultural environment is a cultural hybridity interwoven with different factors, processes, and traits (Duan, Bai, Ho, & Tang, 2012). Thus directly-translated concepts and the corresponding items may have totally different meanings to individuals from a different cultural background, which brings some difficulties to cross-cultural researchers explaining social phenomena and research results using this theory, paradigm and tool (Cheung, 2004). For example, Easterlin, Morgan, Switek, and Wang (2012) found that the scores of SWLS for Chinese in the past decades have been declining. Whether this result is reliable depends on whether Chinese well-being can be precisely measured by SWLS. Nevertheless, Ho and Cheung (2007) recognized that individual-oriented (“I”) life satisfaction indicators could not fully cover the subjective well-being of the Chinese people, because Chinese focus more on collective-oriented (“we”) well-being. When evaluating their well-being, they often take the well-being of “significant others” (family, friends, etc.) into account.

Therefore, another research paradigm opposite to the etic approach emerged which is called the emic approach. The emic approach requires researchers to understand present culture from the perspective of the locals and indigenes as much as possible, and then arrange and analyze the content of research based on the cognition and behavior of the locals (Malinowski, 1922). This approach emphasizes specificity of culture and assumes that theory and measures developed in a certain culture only apply to the members of this culture. If researchers want to develop an equivalent theory or measures in other cultures, then they must abandon the theory and hypothesis in the original culture and further develop relevant concepts or paradigms using local data (Lett, 1990). Due to the versatility of Chinese culture (Fan, 2000), Xing (2003) believes that researchers must develop theory and measures based on Chinese society and culture for accurately measuring the subjective well-being. Therefore, he developed the Subjective Well-being Scale for Chinese Urban Residents based on the emic approach.

It can be found that both etic approaches characterized by cultural generality and emic approaches characterized by cultural specificity have some disadvantages in developing a coherent cross-cultural theory. To compensate for their disadvantages, researchers have come up with a combined emic-etic approach.

**The combined etic-emic approach**

The comparability of theories and measuring tools based on the above two paradigms has been questioned and criticized widely by researchers (Hui & Triandis, 1985; Kankaraš & Moors, 2010). They demonstrate that some key issues, like the translation quality,
cultural relevance, measurement equivalence, and cultural validity, have not been well considered or evaluated, which is not conducive to construct a theoretical framework with cross-cultural consistency. In order to construct a theory with both internal validity and external validity (e.g., comprehensiveness, richness and balance), researchers have come up with a new way of integrating the above two approaches, which is called the Combined Emic-Etic Approach (Cheung, van de Vijver, & Leong, 2011; Leong, Leung, & Cheung, 2010).

Specifically, the Combined Etic-Emic Approach firstly requires the researchers to identify elements with intercultural comparability (e.g. conceptual categories or common items), which can be measured in both native and non-native contexts; then add elements with cultural specificity carefully and gradually; finally, this approach can help to develop a comprehensive theory and measures. Meanwhile, the new theory and measures will be compared with other related theories and tools based on an etic approach to make sure that elements with cultural specificity have incremental validity. In each step, it is necessary to ensure the conceptual equivalence, metric equivalence, functional equivalence and language equivalence (Ho, Rochelle, et al., 2014). Using this approach, Ho and Cheung (2007) added a Chinese characteristics dimension (i.e. “interpersonal well-being”) to the western subjective well-being theory and SWLS. They found that the modified Expanded SWLS (E-SWLS) showed high ecological validity and had feasibility and sensibility in Chinese sub-cultures (e.g., Hong Kong and Beijing). For instance, the E-SWLS can better describe and explain mental health status of individuals in Mainland China and Hong Kong, when satisfaction with life is considered as an indicator for positive mental health (Duan, Bai, Tang, et al., 2012; Duan, Ho, Bai, & Tang, 2013). Although most empirical studies in the field of cross-culture adopt theory-driven paradigms, the combined emic-etic approach does help researchers to develop theory and measures with cross-cultural consistency. In the following section, the studies on character strengths are summarized to illustrate this assumption.

**Method**

A comprehensive literature search was performed to select studies from the following database sources: Google Scholar searches, ERIC, and Web of Science. The keywords employed included: character strengths, factor structure, Values in Action Inventory of Strengths (VIA-IS), wellbeing, happiness, positive emotions, mental health, and academic performance. Additionally, the snowball method was also conducted to find relevant findings. The search covered literature published between 2000 and 2017.

**Review of the empirical studies of character strengths**

The Values in Action (VIA) classification developed by Peterson and Seligman (2004) is a systematic framework for studying character strengths. This classification comprises 24 character strengths that belong to six virtues, and each virtue has three to five strengths (Peterson & Seligman, 2004). The two-tier structure comprises 24 character strengths for the first tier and six core virtues for the second tier (Table 1). Peterson and his colleagues defined character strengths as “a family of positive characteristics shown...
in feelings, thoughts, and actions, each of which exists in degrees” (Park & Peterson, 2009, p. 3) and virtues as “a property of the whole person and the life that person leads” (Peterson & Seligman, 2004, p. 87). Both character strengths and virtues are positive qualities that are regarded as one of the three major themes in modern positive psychology (Seligman & Csikszentmihalyi, 2000). Character strengths and virtues are capacities for thinking, feeling, and behaving in different contexts.

Researchers of VIA (Peterson & Seligman, 2004) have stated that character strengths and virtues are the pathway to wellbeing and mental health. As indicated by the World Health Organization (2015) and some researchers (e.g., Keyes, 2005, 2009), mental health as a completed profile comprises positive and negative components. Character strengths have a significant positive relationship with the positive aspects of mental health outcomes, such as satisfaction with life (Buschor, Proyer, & Ruch, 2013; Park, Peterson, & Seligman, 2004; Peterson, Ruch, Beermann, Park, & Seligman, 2007), wellbeing (Brdar & Kashdan, 2010; Gillham et al., 2011; Park et al., 2004), happiness (Park & Peterson, 2006; Peterson et al., 2007), self-reporting general health (Otake et al., 2005; Shimai, Otake, Park, Peterson, & Seligman, 2006), positive emotions (Güsewell & Ruch, 2012), and the positive elements of life, such as academic success (Lounsbury, Fisher, Levy, & Welsh, 2009), quality of life (Proctor, Maltby, & Linley, 2011), orientation to work as a calling (Peterson, Park, Hall, & Seligman, 2009), and life meaning (Littman-Ovadia & Steger, 2010).

Among the 24 character strengths, Zest, Curiosity, Love, Hope, and Gratitude are identified as the most powerful predictors of happiness and satisfaction with life (Park & Peterson, 2006; Park et al., 2004; Proyer, Gander, Wyss, & Ruch, 2011; Shimai et al., 2006). Proyer et al. (2011) provided further clarification using data on 1087 German women (age ranging from 19 to 73 with an average of 44.10) who completed the online VIA-IS and Temporal Satisfaction with Life Scale. The Temporal Satisfaction with Life Scale was used to assess the past, present, and future satisfaction with life through 15 items. Correlations and stepwise multiple regression analyses showed that the character

Table 1. Values in action classification: 24 character strengths and six virtues.

<table>
<thead>
<tr>
<th>Virtues (2nd tier)</th>
<th>Character Strengths (1st tier)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Wisdom and knowledge</td>
<td>1 Creativity 4 Love of learning</td>
</tr>
<tr>
<td>II Courage</td>
<td>6 Authenticity 9 Zest</td>
</tr>
<tr>
<td>III Humanity</td>
<td>10 Kindness</td>
</tr>
<tr>
<td>IV Justice</td>
<td>13 Fairness</td>
</tr>
<tr>
<td>V Temperance</td>
<td>16 Forgiveness 19 Self-regulation</td>
</tr>
<tr>
<td>VI Transcendence</td>
<td>20 Appreciation of beauty and excellence 23 Humor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Character Strengths (1st tier)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Curiosity 5 Perspective 3 Open-mindedness 7 Bravery 8 Perseverance 11 Love and to be loved 12 Social intelligence 14 Leadership 15 Teamwork 17 Modesty 18 Prudence 21 Gratitude 22 Hope 24 Religiousness</td>
</tr>
</tbody>
</table>

Note. This VIA Classification table was summarized on the basis of Peterson & Seligman (2004).
strength of Appreciation of Beauty and Excellence contributed the most to past happiness and satisfaction with life, and Religiousness strength is the best predictor of future satisfaction with life. Proyer, Gander, Wellenzohn, and Ruch (2013) examined the self-reported psychological and physical wellbeing based on the responses of 440 adults. Most of the character strengths were positively related to satisfaction with life, physical fitness, and healthy behaviors (e.g., healthy eating and regular physical exercise). The mediation role of healthy behaviors was identified in the relationship between character strengths and wellbeing (Proyer et al., 2013). This finding is consistent with the Health Behavior Model (Armitage & Conner, 2000) that indicates the direct influence of strengths on health-related behaviors. In turn, this relationship facilitates mental wellbeing (Armitage & Conner, 2000; Cohen, Scribner, & Farley, 2000). Accordingly, significant associations between strengths and mental health were found.

Few studies have explored the relationship between character strengths and the negative aspects of mental health or mental illness. However, existing findings show that character strengths have a negative association with different illnesses. Physical illnesses may have a negative relationship with Appreciation of Beauty, Bravery, Curiosity, Fairness, Forgiveness, Gratitude, Humor, Kindness, Love of Learning, and Religiousness, and psychological disorders are significantly and negatively related to Love of Learning, Creativity, Curiosity, Appreciation of Beauty, and Gratitude (Peterson, Park, & Seligman, 2006). For instance, researchers have indicated that gelotophobia is mainly negatively related to Hope, Curiosity, Zest, and Love (Proyer & Ruch, 2009). Most of the abovementioned studies are cross-sectional investigations, and valid causal conclusions cannot be drawn. A longitudinal design is needed to reveal the temporal relationship between strengths and mental health.

Intervention studies have been conducted on the basis of character strengths to explore the potential temporal relationships between strengths and mental health. “Identifying and Using One’s Character Strengths” has been recognized as an effective strategy for increasing wellbeing and decreasing psychopathological problems among diverse populations (Duan & Bu, 2017b; Duan, Ho, Tang, Li, & Zhang, 2014; Duckworth, Steen, & Seligman, 2005; Proctor, Tsukayama, et al., 2011; Seligman & Csikszentmihalyi, 2000; Seligman, Steen, Park, & Peterson, 2005; Wood, Linley, Maltby, Kashdan, & Hurling, 2011). In a popular study conducted by Seligman et al. (2005), participants were asked to complete an investigation to obtain individual strength profiles and then use these character strengths in daily life. The participants were informed that using personal strengths would enhance their wellbeing and psychological health. They had to write a short essay to report their thoughts and feelings after using the strengths and the process. The participants consequently improved their happiness for up to six months.

Nevertheless, the character strength-based intervention conducted by Seligman et al. (2005) might have neglected the placebo effect (i.e., participants know the purpose of the intervention study). The placebo effect introduces psychological and physiological changes caused by the symbolic significance of the treatment (Finniss, Kaptchuk, Miller, & Benedetti, 2010; Linde, Fässler, & Meissner, 2011). For instance, previous studies on happiness promotion interventions that concealed the true purpose of the study from the participants (Lyubomirsky, Dickerhoof, Boehm, & Sheldon, 2011) produced relatively
weaker or less durable effects than those generated in the studies of Seligman. Therefore, the placebo effect, or the influence of knowing or not knowing the purpose of the study, should be controlled as an independent variable for examining the true effect of the character strength-based intervention.

To improve and examine the positive functions of character strengths reported in previous studies that were mainly conducted in western countries (Duckworth et al., 2005; Seligman et al., 2005), a strength-based intervention was conducted in Chinese universities in Mainland China for promoting the wellbeing of undergraduates (Duan et al., 2014). The placebo effect (i.e., knowing or not knowing the purpose of the study) was considered as an independent and grouping variable. The other procedures of intervention were the same as those in previous publications (Proctor, Tsukayama, et al., 2011; Quinlan, Swain, & Vella-Brodrick, 2012; Seligman et al., 2005), such as completing the inventory for obtaining character strengths, teaching the methods of using strengths, and writing short essays after using strengths. Overall, the participants increased their satisfaction with life from the baseline to the peak after the six-week intervention, and then they demonstrated a gradual downward trend (Duan et al., 2014). However, the experimental group sustained life satisfaction improvement compared with the control group, and the placebo effect gradually disappeared as time passed (Duan et al., 2014). More recently, Duan and Bu (2017b) constructed another Single Section Character Strengths-based Cognitive Intervention in the Chinese campus context, which was found to be effective in improving thriving (Duan, Guan, & Gan, 2016) and decreasing depression among freshmen. These intervention studies demonstrate that using character strengths is an effective approach for enhancing future mental health.

**Factor issues of character strengths**

*Theory-driven small-data research*

The 24 character strengths have been proposed as universal constructs that can be generalized to different countries or cultures (Peterson & Seligman, 2004). For example, researchers investigated 117,676 American, 17,057 British, 123 Kenyan Maasai, and 71 Inughuit, and reported that the positive qualities demonstrated cross-cultural consistency (Biswas-Diener, 2006; Linley et al., 2007; Park, Peterson, & Seligman, 2006). McGrath (2015) reanalyzed the cross-cultural similarity of the 24 strengths through a sample of 1,063,921 adults in 75 different nations who completed the VIA-IS online from 2002 to 2012. The results indicated that among the 24 character strengths, Honesty, Fairness, Kindness, Judgment, and Curiosity are the top five strengths with the highest consistency in terms of endorsement across all countries. The least endorsed character strengths are Self-regulation, Modesty, Prudence, and Spirituality (McGrath, 2015). However, these results may not apply to Asian countries because the sample size from Asia was relatively small.

Furthermore, the 24 character strengths have been proposed to be grouped into six categories (Table 1, 6 virtues) (Peterson & Seligman, 2004). Nonetheless, existing literature reveals that different structures can be obtained from the 24 character strengths in various cultures using different methods. These strength structures include five-factor
models (Azañedo, Fernández-Abascal, & Barraca, 2014; Peterson & Park, 2004; Peterson & Seligman, 2004; Ruch et al., 2010; Singh & Choubisa, 2010), four-factor models (Brdar & Kashdan, 2010; Macdonald, Bore, & Munro, 2008), three-factor models (Khumalo, Wissing, & Temane, 2008; Shryack, Steger, Krueger, & Kallie, 2010), and two-factor models (Park & Peterson, 2010). Table 2 summarizes these findings, including the researchers, countries, samples, methods, factor numbers, and corresponding labels in the past 11 years. Generally, eight of the aforementioned 14 studies revealed a five-factor structure for the 24 character strengths, two studies found a four-factor solution, three studies illustrated a three-factor solution, and only one study demonstrated a two-factor structure. These inconsistent structures of character strengths imply the justifiable examination of factorial invariance of the structure of these strengths. However, none of the above studies was conducted in any Asian country.

In view of this gap, the present author conducted a study in Mainland China to examine the factor structure of these 24 character strengths (Duan et al., 2011). In this study, 422 undergraduates in China were invited to complete the simplified Chinese version of VIA-IS online that was available on the website of the VIA Institute (http://www.viacharacter.org). The study aimed to explore the internal consistency and factor structure of the inventory (Duan et al., 2011). The Cronbach’s alphas of the 24 strengths (.64 – .84) and the 6 virtues (.88 – .94) were acceptable. However, the convergent and discriminant validities were poor. For instance, some strengths (e.g., Zest) did not correlate well with their corresponding virtues (e.g., Courage). Furthermore, the results of the confirmatory factor analysis (CFA) showed that none of the six-virtue CFA models had an acceptable goodness-of-fit index (e.g., CFI < .900, TLI < .900, and RMSEA > .080; the cut-off point is found in section 5.2.3). Accordingly, Duan et al. (2011) concluded that the structure of these character strengths among Chinese should be further examined.

The studies listed in Table 2 were conducted in various countries (e.g., Germany, United States, Australia, some countries in Africa, etc.) with different methods (e.g., principal component analysis, principal axis factoring, factor analysis, etc.). However, these different structures were extracted based on the same 24 strengths and 240 items. This operation may have disregarded the functional and conceptual equivalences for each strength and its corresponding items (Cheung et al., 2011; Leong et al., 2010). In other words, the strengths and items of the original VIA-IS may acquire different meanings in different cultures (Ho, Duan, & Tang, 2014; Ho, Rochelle, et al., 2014). A simplified Chinese VIA-IS has been developed and made available on the VIA Institute website (http://www.viacharacter.org). In a first attempt to explore the strength structure of this Chinese VIA-IS, Duan, Ho, et al. (2012) administered the VIA-IS in Simplified Chinese to 839 participants (365 males and 471 females) in Mainland China (mean age = 20.62, SD = 2.12; range 17–26). All participants were required to visit the VIA Institute website (http://www.viacharacter.org) to complete the simplified Chinese VIA-IS. Quantitative and qualitative methods were used to conduct the analyses.

In the quantitative phase of this study, the item-level exploratory factor analysis (EFA) showed that more than 50% of the items had a factor loading lower than 0.60 (mean loading = 0.56), and the current 10 items in most subscales of strength only
Table 2. Structures of strength in 14 studies (2004–2014).

<table>
<thead>
<tr>
<th>Researchers and Countries</th>
<th>Methods</th>
<th>Sample Characteristic and Size</th>
<th>Mean Age</th>
<th>Number of Factors</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGrath (2014b) in United States</td>
<td>PCA with Varimax &amp; PAF with Promax</td>
<td>Adults; 458,998</td>
<td>34.40</td>
<td>5</td>
<td>Interpersonal Strength, Emotional Strength, Intellectual Strength, Restraint Strength, Future Orientation Strength</td>
</tr>
<tr>
<td>Azanédo et al. (2014) in Spain</td>
<td>PCA with Varimax</td>
<td>Adults; 1,060</td>
<td>32.72</td>
<td>5</td>
<td>Emotional Strengths; Interpersonal Strengths, Strengths of Restraint, Theological Strengths, Intellectual Strengths</td>
</tr>
<tr>
<td>Littman-Ovadia and Lavy (2012) in Israel</td>
<td>PCA with Varimax</td>
<td>Adults; 635</td>
<td>27.00</td>
<td>5</td>
<td>Restraint Strengths, Interpersonal Strengths, Intellectual Strengths, Emotional Strengths, Theological Strengths</td>
</tr>
<tr>
<td>Ruch et al. (2010) in German</td>
<td>PCA with Varimax</td>
<td>Adults; 1,674</td>
<td>42.20</td>
<td>5</td>
<td>Emotional Strength, Interpersonal Strength, Strength of Restraint, Intellectual Strength, Theological strength</td>
</tr>
<tr>
<td>Peterson, Park, Pole, D’Andrea, and Seligman (2008) in United States</td>
<td>PCA with Varimax</td>
<td>Adults; 1,739</td>
<td>40.00</td>
<td>5</td>
<td>Interpersonal, Fortitude, Cognitive, Temperance, Transcendence</td>
</tr>
<tr>
<td>Peterson and Park (2004) in United States</td>
<td>PCA</td>
<td>Adults; 75,000</td>
<td>35.00</td>
<td>5</td>
<td>Conative Strengths, Emotional Strengths, Cognitive Strengths, Interpersonal Strengths, Transcendence Strengths</td>
</tr>
<tr>
<td>Brdar and Kashdan (2010) in Croatia</td>
<td>PCA, PAF, and MLA with Promax</td>
<td>Undergraduate; 881</td>
<td>20.90</td>
<td>4</td>
<td>Interpersonal Strengths, Fortitude, Vitality, Cautiousness</td>
</tr>
</tbody>
</table>

(continued)
Table 2. (continued)

<table>
<thead>
<tr>
<th>Researchers and Countries</th>
<th>Methods</th>
<th>Sample Characteristic and Size</th>
<th>Mean Age</th>
<th>Number of Factors</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macdonald et al. (2008) in Australia</td>
<td>PCA with Varimax and Oblique (Direct Oblimin)</td>
<td>Undergraduate; 123</td>
<td>21.50</td>
<td>4</td>
<td>Positivity, Intellect, Conscientiousness, Niceness</td>
</tr>
<tr>
<td>Shryack et al. (2010) in United States</td>
<td>PCA</td>
<td>Adult Twins; 332</td>
<td>49.00</td>
<td>3</td>
<td>Intellectual Strengths, Interpersonal Strengths, Temperance Strengths</td>
</tr>
<tr>
<td>Khumalo et al. (2008) in Africa</td>
<td>PCA with Varimax</td>
<td>Postgraduate; 256</td>
<td>22.50</td>
<td>3</td>
<td>Wisdom, knowledge, and Courage, Horizontal and vertical Relatedness, Temperance and Justice</td>
</tr>
<tr>
<td>Jónsdóttir (2004) in the Netherlands</td>
<td>PCA</td>
<td>Undergraduate; 523</td>
<td>21.00</td>
<td>3</td>
<td>Humanity, Development, Discipline</td>
</tr>
<tr>
<td>Park and Peterson (2010) in United States</td>
<td>PCA with Varimax</td>
<td>Adults; 47,369</td>
<td>N.A.</td>
<td>2</td>
<td>Virtues of Head, Virtues of Heart</td>
</tr>
</tbody>
</table>

Note. PCA = Principal Component Analysis; PAF = Principal Axis Factoring; MLA = Maximum Likelihood Analysis. All studies adopted VIA-IS as the measurement to assess the 24 strengths; N.A. = Not Available, i.e. no information provided.
explained no more than 35% of the total variance (Duan, Ho, et al., 2012). In the qualitative phase of the study, group interviews of another 40 participants were conducted to examine whether the items of the VIA-IS conformed to the social norms of China. The results indicated that 103 items did not conform to Chinese social norms and that another 27 items could be potentially ambiguous (Duan, Ho, et al., 2012). For instance, “I never tell outsiders bad things about my team” is used to measure Citizenship. However, this item may lack sensitivity for Mainland Chinese because it is a traditional social norm in a collective culture. Another item, “When I hear people say something mean, I make a protest,” may successfully assess the Bravery of Westerners. However, such a behavior is unacceptable in China, where people prefer a harmonious relationship and do not directly express disagreements on the opinion of others. After removing 144 items from the original 240-item pool, the remaining 96 items had a high degree of cross-cultural endorsement and sensitivity. The new instrument was called Chinese Virtues Questionnaire (CVQ-96). A follow-up second-order EFA revealed a clear three-factor structure, and the CFA further confirmed the stability of this three-factor structure. The three kinds of strengths were identified as Relationship, Vitality, and Conscientiousness (Duan et al., 2013; Duan, Ho, et al., 2012).

Ho et al. (2015) argued that CVQ is too lengthy to complete, especially for individuals with mental health issues or longitudinal research. Therefore, to develop a brief scale, the authors first created 36 items which were suitable to measure strengths in the Chinese culture and were based on positive elements or qualities selected from previous studies. Then, several explorations (i.e., EFA) were conducted, the results of which revealed the same three factors, which were assessed by 12 items, emerged from the population with a history of psychiatric problems (e.g., depression and anxiety). Moreover, the factors measured by the brief inventory can be generalized to normal undergraduate samples in Mainland China, as indicated by the CFA (Ho et al., 2015). Thus, the instrument, the Brief Strengths Scale-12 (BSS-12) (Ho et al., 2015), was developed with the purpose of assessing three kinds of strengths, namely, Interpersonal Strength, Intellectual Strength, and Temperance Strength. Overall, the brief inventory had satisfactory internal consistency (.76 – .84), content validity, and discriminant validity.

However, since the factor structure of this three-factor model was extracted from an Asian sample of clinical and undergraduate population, whether the factor structure is applicable to western samples and whether the metric and functional equivalences of the three-factor model achieved are applicable in different populations are unclear. Duan and Bu (2017a) conducted another research to address these issues by developing a Three-dimensional Inventory of Character Strengths (TICS). An Asian sample was adopted to develop the TICS (e.g. item selection and EFA), and the factor structure was then tested among western students for cross-cultural validation (e.g. CFA). After several attempts, 15 items (5 items per strength) were retained with five items per factor, and the internal reliabilities of the three subscales were higher than .74. To examine the measurement invariance of TICS across populations, a community sample and an inpatient sample were used, performing multi-group confirmatory factor analysis. Results showed that the number of factors and the factor loadings of the items of TICS in the medical and community populations can be equal. Therefore, this study developed a brief and psychometrically sound TICS to assess Caring, Inquisitiveness,
and Self-control and it demonstrates reasonable cross-population (medical vs. community) utility for assessing character strengths.

To sum up, three procedures were employed in the abovementioned studies to explore the structure of character strengths. First, the research procedures are finely controlled. The whole research process is under the guidance of the combined emic-etic approach. When researchers add or remove any of the components of the concepts and items, the conceptual equivalence, metric equivalence, functional equivalence and language equivalence are finely controlled. In addition, the quantitative analysis is used with the support of qualitative interviews to identify those inappropriate components which do not meet the norms of Chinese society and cognition. The quality of data collected via paper-and-pencil and online survey are also finely controlled by means of IP address, response integrity and response time. Second, the researched samples are precisely sampled. The samples cover western and eastern college students, ordinary workers in the factory, community samples, adolescents, individuals with a history of psychiatric problems, and medical populations. In addition, measurement invariance is also examined across age, gender, social economic status of family as well as education levels. Third, the collected data are accurately analyzed.

In addition to descriptive statistics, correlation analysis, regression analysis and other commonly used statistical methods, Exploratory Factory Analysis, Confirmatory Factory Analysis, Multi-group Confirmatory Factory Analysis and Exploratory Structural Equation Model are also used (Duan et al., 2011; Duan & Bu, 2017a; Duan & Ho, 2017b; Duan et al., 2013; Duan, Ho, et al., 2012; Ho et al., 2016; Zhang, Duan, et al., 2014). Finally, a three-dimensional character strengths model was obtained, which is theoretically meaningful, stable and of good applicability and sensibility (see Figure 1). Further studies found that the three-dimensional character strengths play an important role in alleviating depression (Duan & Bu, 2017b) and post-traumatic stress disorder (Duan, Guo, & Gan, 2015), as well as facilitating life satisfaction of individuals (Duan et al.,

Figure 1. Three-dimensional character strengths model.
2014), psychological well-being (Tang, Duan, Wang, & Liu, 2016), perceived pressure (Duan, 2016a; Duan, Ho, Siu, Li, & Zhang, 2015), pathological internet use (Zhang, Yang, et al., 2014), post-traumatic growth (Duan & Guo, 2015), and mindfulness (Duan, 2016b; Duan & Ho, 2017a).

Data-driven big-data research

Big data is recognized as the third revolution of modern social science (Guan, 2014; Luo & Luo, 2015). Nowadays, big data sampling based on a website could provide a precious resource unprecedented for researchers to observe and predict humans’ cognition and behavior. Abovementioned theory-driven small-data studies have explored many cases about how to develop a theory with cross-cultural consistency. Is the emergence of big data helpful to the cross-cultural study?

The studies on structures of character strengths based on big data of millions of samples offer a chance to compare the role between big data and small data in cross-cultural filed. The structures of character strengths have received widespread concern around the world since they were proposed in 2004. Since then, two network platforms have been created for research about character strengths. One is Authentic Happiness Center at the University of Pennsylvania (www.authentichappiness.sas.upenn.edu), and another is the VIA Institute on Character (www.viacharacter.org). Over the past 10 years, more than one million people have participated in this research through the two platforms. All of the participants had a high motivation, and they were attracted by the research which could help them understand themselves. Therefore, high quality data were collected. Meanwhile, poor-quality samples could be eliminated, referring to IP address, visit time and response time recorded by backstage supporters. Thus, the data quality could be further promoted. Nowadays, a growing number of social science researchers conduct their studies using the data collected from social media or the network (King, O’Rourke, & DeLongis, 2014).

McGrath (2014a) examined the character strength structure through the integration of psychological and cultural factors. In this research, three studies were operated with large and independent samples (a total of 1,070,549 cases from 75 countries and regions). Although different measurements of strengths (or good characters) were adopted, including VIA-IS, Essential–Natural–Uplifting Inventory, Perceptions of Personal Qualities, and International Personality Item Pool with VIA, the same three-component solution emerged across all of the datasets via Principal Components Analysis and Iterated Principal-axis Factor Analysis. McGrath labeled the three components as Caring, Inquisitiveness, and Self-control (McGrath, 2014a) and proposed that their three-factor structure substantively overlaps with the study of Duan, Ho, et al. (2012). That is, three factors similar to the ones proposed by Duan, Ho, et al. (2012) were obtained in different and independent studies, namely, Caring (Relationship in the system proposed by Duan et al.), Inquisitiveness (Vitality), and Self-control (Conscientiousness).

Although different labels were used to indicate the three strengths in different studies (Duan & Bu, 2017a; Duan & Ho, 2017b; Duan, Ho, et al., 2012; Ho et al., 2015; McGrath, 2014a; Shryack et al., 2010), their essence and content are the same:
Relationship Character Strength proposed by Duan, Ho, et al. (2012), Caring Character Strength by McGrath (2014a) and Duan and Ho (2017b), and Interpersonal Character Strength by Ho et al. (2015) and Shryack et al. (2010) all refer to the character strength that reflects love, concern, and gratitude of a person toward others; Vitality Character Strength proposed by Duan, Ho, et al. (2012), Inquisitiveness Character Strength by McGrath (2014a) and Duan and Ho (2017b), and Intellectual Character Strength by Ho et al. (2015) and Shryack et al. (2010) all refer to the character strength that reflects the curiosity and zest for creativity of an individual; Conscientiousness Character Strength proposed by Duan, Ho, et al. (2012), Self-control Character Strength by McGrath (2014a) and Duan and Ho (2017b), and Temperance Character Strength by Ho et al. (2015) and Shryack et al. (2010) all refer to the character strength that describes people who persist in achieving goals and exhibit self-control.

### Conclusion

Big data has played an important role and advanced substantial development in the field of health care, biomedicine, economy and electronic information engineering. Yet it remains rare in cross-cultural social sciences. This article takes character strengths as an example to illustrate the roles of traditional theory-driven small data studies and data-driven big data studies in developing theory and measures with cross-cultural consistency. The comparison further indicated that the fine controlled, precisely sampled, and accurately analyzed theory-driven-small-data studies can be replicated by big data ones. However, the following issues on usage of big data in cross-cultural studies should be noted.

The first issue is the quantity of big data. Big data refers to data of extremely large quantity, and they are impossible to be processed and analyzed in reasonable time to reach an understandable state of human cognition (Laney, 2012). Yet the magnitude or size of data is not considered as the sole criterion of big data (Li & Cheng, 2012). So far, there is no one widely accepted definition of big data (Luo & Luo, 2015). Limited by specific topic, there is a difference among data quantity involved with different topics. Thus, no one absolute magnitude can define the range of big data. Traditional studies mostly take samples from the population and extend the results to the wider population according to the statistical inference method. This operation makes the results inevitably influenced by the representation of samples, such as the abovementioned cultural differences of structure of character strengths. In the age of big data, due to the data on a certain specific topic from an online platform, direct research on the population can be performed. However, even though big data can be obtained from online platform, it might not reflect the population well. Users on the online platform cannot match the demographics of a general population, meaning that they are a nonrandom sample (Van Dijck, 2014). Thus, an important question of representation is raised when using an online platform as a “big data” model organism (Tufekci, 2014). Most commonly, we do have demographics (i.e., gender and age) of users on the online platform, which means that big data can be treated as a stratified sample and can reflect a specific population (Kern et al., 2016).

The second issue is the quality of big data. Researching whether in traditional or in big data, the quality of data is of vital significance. During the process of traditional
questionnaire survey and testing research (e.g. distrusting printed questionnaires, waiting for feedback for respondents, and then inputting and analyzing data), each step could cause an error and limit the efficiency. However, in the age of big data, techniques and devices, like networking platforms, cloud computing and real-time feedback, would automatically gather, store and analyze data, which contributes to the reduction of human error and costs as well as the improvement of efficiency. As for what has been discussed above, traditional methods spend up to six years to explore the structure of character strengths, while big data based on millions of samples greatly improves the efficiency. It should be noted that when big data is analyzed, the sampling process is also necessary to improve data quality. For example, Gong and Luo (2015) chose and analyzed the historical trend and evolutionary process of the maritime Silk Route and main trade products in the 19th century based on 8 million books in Google Books. The reason for choosing this period is because of the limitations of the research thesis and data integrity. Through this sampling process researchers could present results of higher quality. Therefore, big-data analysis in cross-cultural research should ensure high quality by rational sampling.

The third issue to note is the relationship between “big data” and “small data”. Through the above analyses, it is obvious that big data and small samples in fact represent two different research paradigms – the former is driven by data while the latter is theory-driven. Generally, data-driven research is considered as exploratory study while theory-driven research is confirmatory study. As explained earlier, the two mutually complement and verify each other, pushing the development of theory with cross-cultural consistency. For example, Weibo, Li, Li, Hao, Guan, and Zhu (2014) compared the personality traits based on users’ behaviors on Weibo and those from traditional self-reported surveys. They found that the correlation between the two reached 0.48 to 0.54 and suggested a consistency between personality traits based on online behaviors and traditional five personality theories.

In general, in cross-cultural studies big data with high quality can overcome the defects of the etic approach and the emic approach, both to satisfy the specificity and universality of culture. It will be an important means to conduct future cross-cultural studies and theorize about cross-cultural consistency.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: National Social Science Foundation of China (Grant No. 17CSH073); Wuhan University Humanities and Social Sciences Academic Development Program for Young Scholars “Sociology of Happiness and Positive Education” (Grant No. Whu2016019).

References


King, D. B., O’Rourke, N., & DeLongis, A. (2014). Social media recruitment and online data collection: A beginner’s guide and best practices for accessing low-prevalence and


Author biographies

Wenjie Duan, PhD, is an Associate Professor of Sociology and Social Work in the Department of Sociology at Wuhan University. His research focuses on the promotion of wellbeing and happiness in education, community, medical care, and elder care context, the development and evaluation of positive education, and the big-data based behavioral science and mental health. He is the author of 27 peer-reviewed SCI/SSCI publications. Recent publications include “Does being Mindful of Your Character Strengths Enhance Psychological Wellbeing? A Longitudinal Mediation Analysis” in Jounral of Happiness Studies (2017) and “Development and Initial Validation of a Short Three-dimensional Inventory of Character Strengths” in Quality of Life Research (2017).

Wenlong Mu is a graduate student majoring in Sociology in Department of Sociology, Wuhan University. His main fields of interest are social psychology and positive education. His recently publications include “Validation of a Chinese version of the Stress Overload Scale-Short and its use as a screening tool for mental health status” in Quality of Life Research (2017).

He Bu, MSW, is a graduate student majoring in Social Work in Department of Sociology, Wuhan University. Her main field of interest is positive psychology intervention. Her recently publications include “Randomized Trial Investigating of a Single-Session Character-Strength-Based Cognitive Intervention on Freshman’s Adaptability” in Research on Social Work Practice (2017).