

Test and Measurement: Final Project



PSYC-8316P-1C: Test & Measurement

Professor: Dr. Rochelle Michel

November 16, 2012

Abstract

This Test and Measurement: Final Project consists of two sections. The first section contains *Test Construct and Development*, including test construct, test specifications, and test items development. The second section contains *APA Results Section*, including statistical analyses of *validity analysis, reliability analysis, and factor analysis* interpretations. The test construct that I would like to measure is the impact of tsunami disaster in Japan 2011, and its effect on posttraumatic growth among Japanese children and adolescents. The construct would be assessing tsunami victims on their experiences of posttraumatic, growth, and grief inventories. In a real life situation, I am involved with the disaster relief in Japan, including medical attention, temporary housing, and finding caretakers for children, adolescents, and seniors. The test items are consisting of 50 items, however, after item analysis and factor analysis elimination, there will be approximately 25 items retained: 15 related to grief experience and trauma history, 15 related to growth, and 20 related to posttraumatic experience. Because we are assessing children and adolescents, so I prefer using open-end and closed-ended questions. Statistical analysis was conducted using SPSS 20 software, including *Item Reliability Analysis, Validity Coefficient Analysis, and Factor Analysis*. The file that we used to illustrate the statistical procedures is called the MoneyData.sav, the dataset contains participants of $N = 1146$. The data file includes a six-item scale to measure financial Risk-taking tendencies: R1: I'd rather run my own business than work for someone else; R2: I don't mind risking large amounts of money if there is a good chance I can come out ahead; R3: I'm willing to take real chances to get ahead financially; R4: I get bored unless I'm taking some risks with my career; R5: Being too conservative with your investments can cause financial problems; R6: Running a business is something that I think of as interesting and exciting.

Test and Measurement: Final Project*First Section: Test Construct and Development:**Test construct, test specifications, and test items*

The research topic of this test construct is “The impact of 2011 tsunami on posttraumatic growth in Japanese adolescent victims.” The test specification that I intended to measure is psychological disorder, and the specific construct of the measurement is Posttraumatic Growth. The behavior associated with the construct would be: The impact of tsunami disaster in Japan, 2011, and its effect on posttraumatic growth among Japanese adolescents. The construct would be assessing tsunami victims on their experiences of posttraumatic, grief, and growth inventories, as well as self-reported questionnaire of trauma history. The process for initially generating items would be considering my personal involvement with the disaster relief in Japan, including medical attention, temporary housing, and finding caregivers for children and adolescents. Therefore, I will interview victims and non-victims from the database of our hospital and temporary housing, the test items would be developed based upon the literature and peer-reviewed journal articles, and DSM-IV criteria for PTSD.

The test items are consisting of 50 items, however, after item analysis and factor analysis elimination, there will be approximately 25 items retained: 15 related to grief experience and trauma history, 15 related to growth, and 20 related to posttraumatic experience. Because we are assessing children and adolescents, so I prefer using open-end and closed-ended questions. According to Klein (2005), both open-ended and close-ended questions can offer straightforward answers when making decisions, especially in clarifying conflict findings at the early stage of development of the research (Kline, 2005). This is the reason why I chose the open-ended and close-ended response formats (e.g., dichotomous responses: Yes/No), starting with the first open-

ended question and then progress to the dichotomous questions such as truth/false or yes/no responses. One of the examples from my developed test items: “prior to this disaster, have you ever lost any family members? (Yes/No), and “if so, which one of your family member(s) that you lost?” as a follow-up question, and so on.

According to Anastasi and Urbina (1997), the general purpose of defining the appropriate target population is to ensure the characteristics of the population of interest would serve the goal of the study (Anastasi & Urbina, 1997). The three most important considerations of sampling methodology are: 1) defining the target populations, 2) requirements for recruiting participants for the sample, and 3) finding sufficient amount of participants for the size of the sample (Anastasi & Urbina, 1997). First, defining the target populations: as the topic indicates, “The impact of 2011 tsunami on posttraumatic growth in Japanese adolescent victims,” the population of interest is the victims of the tsunami and nuclear disasters, that is, people who lost their homes and family members to the disasters. Second, requirements for recruiting participants for the sample: the sample of the population interest is the adolescent victims between 12 to 17 years of age, female and male. The sample contains the control group (non-victims) and the experimental group (victims) to evaluate the impact of posttraumatic experience on growth. Each group has 100 participants of 50 males and 50 females, with total of 200 participants, comparing victims vs. non-victims.

The information that I anticipate to collect through my test instrument is to conduct a structured interview and survey for assessing posttraumatic symptoms, grief experiences, and growth inventory, for measuring the impact of 2011 tsunami on posttraumatic growth in Japanese adolescent victims. According to Kline (2005), when conducting self-reported questionnaires, surveys, telephone interviews, or in-person interviews, several factors should be

taken into account, such as respondent's home environment, the safety of the interviewer, interruptions during the interview, and the presence or influence of the family members or roommates (Kline, 2005). Therefore, in order to make the test result more reliable and valid, it is ideal to have the survey and interview at designated and easy accessible locations, in this case, the designated location for conducting my interview is at the temporary housing and medical facility that provided to the victims.

The validity of my developed test items are based on the references of clinical approaches that have been accepted and proven by the mental health professionals and organizations, such as Posttraumatic Stress Disorder (PTSD) inventory for children and adults, trauma symptom checklist for children and adolescents, and growth inventory (Briere, 2005; Saigh, 2004). The diagnosis of PTSD has been developed into a wider range for assessments, categorizing by three major areas: posttraumatic symptoms (intrusive re-experiencing, avoidance numbing, autonomic hyper arousal, and stress-total), posttraumatic impairment of psychosocial functioning, and associated features of PTSD, including trauma experience prior and after the posttraumatic incident, the actual symptom of PTSD, and the residual PTSD symptoms associated features, such as trauma-specific dissociation, suicidal, and substance abuse (Briere, 1996; 2005).

The Children's PTSD Inventory is designed for the identification and the assessment of posttraumatic stress disorder (PTSD) in children and adolescents (Saigh, 2004). The test instrument for this inventory designed as a structured interview questionnaire, consisting of an answer sheet, a scoring sheet, and a manual, while the item set consists of seven sections closely derived from DSM-IV-TR diagnostic criteria for PTSD: exposure, situational reactivity, re-experiencing, avoidance/numbing, increased arousal, significant distress or impairment, and stress total (DSM-IV-TR; as cited in Saigh, 2004). The influence of reliability of the measure of

children's PTSD inventory on the magnitude of its validity must yield on the clinical contexts of the assessment, in which including the diagnosis of the disorder, the symptom of the disorder, the treatment planning, and the predictions of outcome evaluations (Briere, 2005; Saigh, 2004).

The data outcomes for participants in both the control group and the experimental group would be obtained through administrator for comparison, and the outcomes of their test results would continually being measured. The follow-up survey would be conducted every 6 month, up to 3 years, this way we can find out adolescents who were at 15 to 17 years of age at the time of the disaster, and what they might have become 3 years later into the adulthood, in regards to whether or not a long-term impact will be the outcome of our evaluation. The data collection method that I have chosen is "Theoretical or Purpose Sampling," in which the sample is used in theory building (Kline, 2005). In this case, the theory that I am building for is "how PTSD impact growth." According to Kline (2005), the procedure using by data collectors for "Theoretical or Purpose Sampling" is conducting interview and survey to measure an individual for a certain phenomenon (how PTSD impact growth). The advantage of this method is the convenience for me to obtain sample and collect data from adolescent victims who living in the temporary housing around the disastrous areas, as well as finding useful information from the results of the interviews and surveys in understanding the phenomenon and supporting the theory building (Kline, 2005). The disadvantage of this method might be the "theme saturation" (Kline, 2005), to which the participants memorized the same questions being asked by data collectors, and offers no further information, when this situation occurred, the sampling would be disrupted while there is no new information being obtained for supporting the theory (Kline, 2005); therefore, considering revising the questionnaire and survey should be taken into account.

As described previously, in order to ensure that the sample represents the population of interest, I utilize a real life situation, which is my personal involvement of the disaster relief in Japan, the supports that we offered are including medical attention, temporary housing, and finding caretakers for children, adolescents, and seniors. (Note: multicultural concern would not be applied to this research, because the target population of interest is Japanese adolescent victims from the tsunami regions). Therefore, we recruit adolescents who live in the temporary housing by the tsunami regions, asking for both the volunteers and pay participants, for those who lost parents or legal guardians, we asked social service professionals to oversee the process, whereas for those who have parents and legal guardians would be asked for parental/guardian consents. Third, finding sufficient amount of participants for the size of the sample, because the sample of the age group is the adolescents between 12 to 17 years of age. The sample size ($N = 200$) may be critical; therefore, we might consider going beyond to other tsunami relief locations, in order to obtain sufficient amount of adolescent victims. Furthermore, because the tsunami relief involves health and public safety, in which the contaminations associated with the exposures of Nuclear Radiation should be taking into serious precautions, professional analysis and examination for hazardous contaminants and environments should also be addressed in specific standards for public safety.

This test construct may also utilize the Item response theory (IRT) and classical test theory (CTT). IRT is a group of psychometric models measuring the applicability of each model in a particular situation and individual differences according to the test items, while the CTT model is primarily focused on (the sum of the raw test scores) test-level statistics such as correlation coefficient and reliability coefficient, (George & Mallery, 2011). First, the CTT test-level statistics can be used to calculate the reliability coefficient of the internal consistency

between the average score of a sample (observed indicator: adolescent victims) from a target population (hypothetical indicator: tsunami victims). Second, the IRT item-level statistics can be used to measure logistic function and relation between the latent trait (posttraumatic growth) and the item responses (the impact of tsunami). The test construct for my research is the combination of open-ended and closed-ended questionnaires and interviews (e.g., dichotomous item: Yes/No). In a clinical assessment, IRT is used to investigate the relation between the latent trait and the item responses, items scored dichotomously are referred to as binary items, each correct answer received a “1” point, and the incorrect answer received a “0” point (Yes=1; No=0). The item response curve (IRC) is for binary items, the category response curve (CRC) is for polytomous items, and the two-parameter logistic (2PL) allows differing binary items in two ways: discrimination and severity (Reise & Waller, 2009). Finally, the two-parameter logistic (2PL) model can measure the discrimination and severity of my developed test items, which helps me to figure out the severity of each test items (e.g., posttraumatic, grief, and growth inventory, and how these items indicate the impact of tsunami disaster).

Second Section: APA Results Section

Statistical Analysis: validity, reliability, and factor analysis interpretation

Statistical analysis was conducted using SPSS 20 software, including *Item Reliability Analysis, Validity Coefficient Analysis, and Factor Analysis*. The file that we used to illustrate the statistical procedures is called the MoneyData.sav, the dataset contains participants of $N = 1146$. The data file includes a six-item scale to measure financial Risk-taking tendencies:

R1: I'd rather run my own business than work for someone else

R2: I don't mind risking large amounts of money if there is a good chance I can come out ahead

R3: I'm willing to take real chances to get ahead financially

R4: I get bored unless I'm taking some risks with my career

R5: Being too conservative with your investments can cause financial problems

R6: Running a business is something that I think of as interesting and exciting

In *Item Reliability Analysis*, the *Cronbach's Coefficient Alpha* was conducted to estimate the reliability of these six items (R1 to R6), the alpha coefficient is: 0.79, the scale standard deviation is: 5.02, and the scale's standard error of measurement is: 2.28. According to *corrected item-total correlation*, R3 has the highest value of: 0.664, whereas R5 has the lowest value of: 0.285; therefore, we suggested to eliminate item R5 in order to improve *Cronbach's Coefficient Alpha* for the scale.

In *Validity Coefficient Analysis*, *Pearson Correlation Coefficient* was conducted to estimate the validity between a predictor scale and criterion measure. The scales that I am interested in measuring are the Risk-taking scale, along with the expected correlation between Risk-taking and Need-luxury, and Risk-taking and Lifestyle. The validity hypothesis is that for those who have the tendency to take financial risks in investments and careers usually demands more luxury in lifestyle.

Measuring the expected correlation between Risk-taking and Need-Luxury:

H₀: There is no significant correlation between Risk-taking and Need-Luxury.

H₁: There is a significant correlation between Risk-taking and Need-luxury.

We reject the null, conclude that there is a significant correlation between Risk-taking and Need-luxury, $r = 0.107$, $p < 0.01$ (sig, 2 tails)

Measuring the expected correlation between Risk-taking and Lifestyle:

H₀: There is no significant correlation between Risk-taking and Lifestyle.

H₁: There is a significant correlation between Risk-taking and Lifestyle.

We reject the null, conclude that there is a significant correlation between Risk-taking and Lifestyle, $r = 0.254$, $p < 0.01$ (sig, 2 tails)

The statistical analysis of *Pearson Correlation Coefficient* between a predictor scale and criterion measure was conducted to calculate the expected correlation between Risk-taking scale versus Need-luxury and Lifestyle. The MoneyData.sav dataset contains participants of $N = 1146$. As the results indicated, there is a positive correlation between Risk-taking and Lifestyle, $r = 0.254$, $p < 0.01$, as well as a positive correlation between Risk-taking and Need-luxury, $r = 0.107$, $p < 0.01$. Although both Lifestyle and Need-luxury indicated a positive relationship with Risk-taking scale; however, there is a stronger relationship between Risk-taking and Lifestyle than Risk-taking and Need-luxury.

In *Factor Analysis*, *Exploratory Factor Analysis* was conducted to investigate whether or not the factors are useful and can be replicated, and to characterize the factors by assigning them with a name or label. Factor analysis is a data-reduction technique used to define the number of factors underlying correlations among the variables. Three major steps required to conduct a factor analysis are *correlation matrix*, *factor extraction*, and *factor rotation*. *Exploratory factor analysis* is a technique used to define how items group together, and *confirmatory factor analysis* is used to investigate whether these items group together are consistent with the theory that we developed or the results that we expected. In *Exploratory Factor Analysis*, the *Principle Axis Factoring* extraction with *Direct Oblimin* rotation were used to estimate the likely number of factors, as well as to examine a solution that assumes that the factors are correlated with one another. The *Factor Analysis* shows that there are three distinct scales, with 6 items each, and total of 18 test items (variables), including Lifestyle scale (L1 to L5, L6R), Dependence scale

(D1 to D6), and Risk-taking scale (R1 to R6). As the results indicated, there are six factors with *eigenvalues* larger than 1.0 and they accounted for 64% of the total variance, and the first three factors are accounted for the highest percentage of variance (Factor 1: 19.94%, Factor 2: 13.62%, and Factor 3: 11.03%). The Scree Plot indicated that these three distinct factors should be retained, that is, Financial risk-taking tendencies (Factor 1), Tendency to depend on others for financial support (Factor 2), and Desire for a luxurious lifestyle (Factor 3). According to the *Pattern Matrix*, each item has a high loading in only one factor, and for the *Factor Correlation Matrix*, the three factors that have been identified correspond to Factor 1 and Factor 3 has a value of (- 0.293); Factor 2 and Factor 3 has a value of (- 0.122); Factor 1 and Factor 2 has a value of (- 0.016).

References

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Briere, J. (2001). Detailed Assessment of Posttraumatic Stress.
- Briere, J. (2005). Trauma Symptom Checklist for Young Children.
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). College Park, MD: ERIC Clearinghouse on Assessment and Evaluation. Retrieved from the Walden Library Databases.
- George, D., & Mallery, P. (2011). *IBM SPSS statistics 19: A simple guide and reference* (12th ed.). Boston: Pearson.
- Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage. Kline, T., Psychological testing: A practical approach to design and evaluation. Copyright 2005 Sage Publications Inc. Books. Used with permission from Sage Publications via the Copyright Clearance Center.
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, 5, 27–48. Retrieved from the Walden Library databases.
- Saigh, P. (2004). Children's PTSD Inventory: A Structured Interview for Diagnosing Posttraumatic Stress Disorder.

Appendix

Construct Validity is usually accumulated over time and represents the accumulation of many validity studies; it indicates whether the test appropriately measures a theoretical construct. The construct I selected is “Children's PTSD Inventory.” The measure of this construct includes a structured interview for diagnosing Posttraumatic Stress Disorder (PTSD), and trauma symptom checklist for children and adolescents between 7 and 17 years of age. This Inventory consists 3 sections: answer sheet, scoring sheet, and the manual (Saigh, 2004). Literatures and peer-reviewed studies supporting the construct validity of Children’s PTSD Inventory, by closely follow the DSM-IV-TR diagnostic criteria for PTSD, with the total of 5 sections of PTSD inventory: 1) Exposure and Situational Reactivity, 2) Re-experiencing, 3) Avoidance and Numbing, 4) Intrusive and increased Arousal, 5) Significant Distress or impairment (Saigh, 2004; Briere, 2005). The influence of reliability of the measure of children’s PTSD inventory on the magnitude of its validity must yield on the clinical contexts of the assessment, in which including the diagnosis of the disorder, the symptom of the disorder, the treatment planning and outcome evaluations, and the prediction of the outcome of the test construct, these requirements are imperative for test validity (Saigh, 2004; Briere, 2005).

Test Item Development

As described previously, the test construct that I would like to measure is the impact of tsunami disaster in Japan 2011, and its effect on posttraumatic growth among Japanese children and adolescents. The construct would be assessing tsunami victims on their experiences of posttraumatic, growth, and grief inventories. In a real life situation, I am involved with the disaster relief in Japan, including medical attention, temporary housing, and finding caretakers for children, adolescents, and seniors. The test items are consisting of 50 items, however, after

item analysis and factor analysis elimination, there will be approximately 25 items retained: 15 related to grief experience and trauma history, 15 related to growth, and 20 related to posttraumatic experience. Because we are assessing children and adolescents, so I prefer using open-end and closed-ended questions.

Grief Experience & Trauma History:

1). Prior to this disaster, have you ever lost any family members?

Yes / No

2). Which one of your family member(s) that you lost (death)?

3). Do you feel lost and helpless for the future, after the tsunami disaster?

Yes / No

4). Do you have any family members missing from this disaster?

Yes / No

5). If so, which one of your family member(s) are missing?

6). Do you have close relationship with your parents?

Yes / No

7). Do you have close relationship with your siblings? If so, are they your brother or sisters?

8). Do you feel uncomfortable and insecure living in this temporary housing?

Yes / No

9). If you don't feel comfortable and secure living in this temporary housing, where do you prefer to live?

10). I am worry about my personal safety living in this temporary housing.

Truth / False

11). Prior to this disaster, have you ever experienced any natural disaster?

Yes / No

12). If so, what types of natural disaster have you experienced?

13). Prior to this disaster, have you ever lost any loved ones to death?

Yes / No

14). If so, what was the cause of his or her death?

15). I feel all along and abandoned by my family members and loved ones.

Truth / False

Posttraumatic Indications:

16). I believed that the lost of my parents, family members, or loved ones are only temporary, and therefore, they will come back to me in the future.

Truth / False

17). I often experiencing death anxiety, in which the same disaster may occur once again, and when it happens, I might not be alive this time.

Truth / False

18). Do you feel guilty that you are the only survivor in your family?

Yes / No

19). After the disaster, have you ever experiencing anger, hostility, or loss of control of your temper?

Yes / No

20). When you loss of control of your temper, what do you usually do to release your anger?

21). After the disaster, have you ever attempted to commit a suicide?

Yes / No

22). If so, how often do you have suicidal attempt?

23). I believed that I am suffering from depression.

Truth / False

24). Do you experiencing social isolation, to which the peer at school and people in your circle do not treat you like the same person before the disaster?

Yes / No

25). Do you know anyone who suffering from the same disaster as you? If so, who are they?

26). Do you have trouble sleeping at night?

Yes / No

27). I frequently experienced horrifying nightmares.

Truth / False

28). I have hard time distinguish between reality and fantasy.

Truth / False

29). I frequently re-experiencing the tsunami disaster, and the feeling are intrusive.

Truth / False

30). I frequently experiencing autonomic hyper-arousal feelings.

Truth / False

31). If so, what do you usually do to release the hyper-arousal feelings?

32). My psychosocial functioning is impaired.

Yes / No

33). I often feel avoidance and numbing when interacting with other people.

Yes / No

34). Have you ever drink alcoholic beverages? If so, do you remember how many drinks that you have had so far?

35). Have you ever tried any illicit drugs? If so, what kind of drugs have you tried?

Growth Inventory:

36). Besides psychological distress caused by the disaster, is there any physical discomfort or illness that you have?

Yes / No

37). If so, what kind of illness do you have?

38). Are you willing to seek for professional help and support during the time of need?

Yes / No

39). Do you trust those people who desire of helping you through this tough time?

Yes / No

40). Do you think that this tragic event (tsunami disaster) will have a negative impact on your adult life?

Yes / No

41). Are you romantically involved with someone?

Yes / No

42). Is there any changes of your future plans after the disaster; such as what would you like to be when you grow up (e.g., teacher, doctor, policeman, soldier, singer, actor, or president, etc.)?

Yes / No

43). What would you like to be when you growth up?

44). Is there any changes of your future plans in pursuing your academic goals?

Yes / No

45). What is your academic goals?

46). Do you feel like your dreams for the future are shattered because of this disaster?

Yes / No

47). Do you somehow believe that you will overcome this tragedy, and become a stronger person?

Yes / No

48). Do you somehow believe that this tragic event would serve you as going through trial, and makes you more matured compare to the same age peers?

Yes / No

49). Do you have faith in yourself that you can grow up to be a very successful person?

Yes / No

50). Would you like to help others who are in the same situation as you?

Yes / No