Validity and reliability of the happiness scale for adolescents in Yogyakarta

Sintia Gumara Dini

Department of Psychology, Faculty of Educational Sciences, Universitas Negeri Yogyakarta,
Jl. Colombo No. 1, Depok, Sleman, Yogyakarta 55281, Indonesia
sintia.gumara2015@student.uny.ac.id

Abstract

Happiness is a subjective concept because every individual has a different benchmark. Nolen-Hoeksema (1998) argues that adolescents have a higher level of depression than adults, who tend to show lower levels. For that reason, there needs to be a scale to measure happiness in adolescents. This research aims to construct a happiness scale for adolescents in Yogyakarta to measure the level of happiness of adolescents in Yogyakarta in a valid way because not all measurement tools used in western cultures can be implemented in the east. This research uses a constructed psychological measurement tool. This research involves 120 adolescents aged 16-18 years in Yogyakarta. In the factorial validity test or factor analysis, there are several items separated from the initial indicators that form new indicators. These, however, are still able to uncover the aim of constructing the scale. From the 60 items proposed, five items were rejected, and 55 others were considered appropriate. Reliability testing yields a result of 0.936, meaning that the scale has a sufficient level of reliability.

Keywords: *happiness, validity, reliability*

Introduction

In life, every individual experiences different events. Each will give a different assessment of each event experienced, even though the events experienced by an individual and another individual may be similar or identical. Additionally, the attitudes that each displays toward these events are also different. Diener, Suh, and Oishi (1997) mention that internal experiences received by each can be described by the term subjective well-being.

Subjective well-being is a person's perception towards experiences in his/her life, encompassing cognitive and affective evaluation of life and representation of psychological welfare (Compton, 2005). According to several psychology experts, subjective well-being is a scientific term used for happiness or contentment. Carr (2004) provides a similar definition of subjective-well being and happiness, which is the positive psychological condition characterized by high levels of satisfaction toward life, high levels of positive emotion and low levels of negative sentiment.

Happiness is a subjective concept because each has different benchmarks. Further,

Aristoteles (in Adler, 2003) states that happiness is something that creates a joyful experience. Seligman (2005) explains that happiness is a concept that refers to a positive emotion that is felt by the individual as well as positive activities that have no emotional component. Meanwhile, happiness, according to Biswas, Diener, and Dean (2007), is the quality of overall life of human life, something that makes life better as a whole, such as good health, high creativity or high income. Based on the definitions of happiness outlined, it can be concluded that happiness is a positive feeling that is marked by joyful experiences felt by individuals that cause the individual's life and health to improve.

According to Park (2004), subjective well-being or happiness has long been considered a core component in the good of life. Individuals with higher levels of subjective well-being or happiness tend to have outstanding qualities (Diener, 2000). These individuals can control their emotions better and face many events in life with a better attitude. Meanwhile, individuals with a low level of subjective well-being or happiness tend to view their lives inferior and consider events

occurring as unpleasant. Therefore, negative emotions emerge, such as anxiety, depression, and anger (Myers & Diener, 1995).

Based on the afore-mentioned discussion, it can be concluded that not all individuals have high levels of subjective-wellbeing or happiness. Nolen-Hoeksema (1998) argues that adolescents have high levels of depression and adults have lower levels of depression (in Ehrlich & Isaacowitz, 2002). Arnett (1999) also states that teenagers also feel self-conscious and confused two or three times more often than their adults and also have a tendency to be awkward, lonely, anxious and ignored. Arnett (1999) also mentions that adolescents also experience mood disorders that are more often than pre-adolescents.

According to Sarwono (2003), this condition is caused by several events in the adolescent stage, i.e., the transition from adolescence to adulthood. During this stage, teenagers will experience changes and receiving demands from adulthood. Kurniawan (1998) explains that adolescents experience several changes in their lives, such as physical and physiological, emotional, mental, social and moral. These changes demand teenagers to undergo big changes in their lives with regards to attitude and behavior according to their developments. Moreover, according to Arnett (1999), Stanley Hall considers the adolescent sage as the 'storm and stress' stage. According to history and the theories combined with contemporary research, the core of storm and stress is a portrayal of the adolescent stage that is the most difficult.

Meanwhile, a psychosocial figure named Erik Erikson states that the main task of adolescents is overcoming identity crisis or identity confusion. During this stage, teenagers are demanded to find a personal identity and avoid role confusion as well as identity confusion. However, Kidwell, Dunham, Bacho, Pastorino, and Portes argue that adolescents who actively search for their identities tend to display doubt, thinking disorders, impulsive behavior, and conflicts with their parents (in Rice, 1999). Meanwhile, the process of identity formation does not only occur during the adolescent stage; rather it takes a very long time.

Seligmen (2009) explains that in reality, individuals want positive emotions, but go through infinite imaginary shortcuts to feel happy such as drugs, free sex, shopping, masturbation, chocolate, and television. Happiness obtained through these means are temporary. Thus the individuals do not feel happiness on the whole. These methods cause negative effects, such as addiction or even depression.

Puspitorini (2012) describes happiness as a mental state or feeling of joy and peace, both physically and emotionally, defined to improve self-function. Happiness causes individuals to have a healthy personality. A positive mood can cause individuals to be objective when responding to something, tolerant, not be defensive, generous and lateral or able to solve problems in a creative way (Seligman, 2005).

From the afore-mentioned discussion, it can be concluded that the definition of happiness used in this research is happiness that can be attained if the individual can achieve their targets or fulfill their needs, such as material, freedom, health, education, knowledge, and friendship. This individual will feel accomplished when all of his/her goals can be achieved. The satisfaction which is obtained from this is a sign that the individual has achieved happiness. Therefore, the individual's happiness increases as his/her satisfaction increases.

According to Carr (2004), the factors that influence happiness are, among others: (1) personality, (2) culture, (3) marriage, (4) social support, (5) friendship, (6) health, (7) religion and spirituality, and (8) cooperation. From the interviews with teenagers aged 16-18 in Yogyakarta, the researcher finds that there is an inability to solve problems in these adolescents. The inability of these teenagers to solve problems gives rise to dissatisfaction, influencing the life they are currently living. The more the issues that cannot be resolved, the lower the happiness felt by these teenagers. Another issue is digression shown by these adolescents to obtain happiness. Many of them even commit these acts as an escape from their problems and inabilities to face them.

Currently, many measurement tools can be used to measure happiness across cultures. Considering that most of these tools were developed in western cultures, there may be differences in the result when implemented in the east (Taggart et al., 2013). This is what eventually motivated this research in validity and reliability of happiness scales on adolescents in Yogyakarta. This research is also aimed to verify information about the level of happiness in teenagers in Yogyakarta.

Method

The procedures were undergone to determine the standardized testing scale for happiness in adolescents in Yogyakarta, include: (1) determining concept or theory, (2) choosing scale format, (3) item formulation, (4) item testing, (5) instrument testing, (6) item difference testing, (7) factorial validity test, (8) revision on instrument or final scale formulation, (9) reliability testing, and (10) final instrument.

The construction of a happiness scale starts by determining the concept or theory to be uncovered by the researcher. In deciding the concept or theory, the researcher uses Carr's theory to determine the aspects to be included in the instrument, to measure the level of happiness of adolescents in Yogyakarta.

The scale format used in this instrument is the Likert scale, consisting of (1) strongly disagree, (2) disagree, (3) neutral, (4) agree, and (5) strongly agree. The subjects were asked to give their most honest responses based on their experiences or feelings.

The researcher designed a large number of items, but must still conform to the indicators that have been outlined, because, during the process of validation, some of the items will be eliminated. Periantalo (2014) advises that researchers make two, three, even four times more items than the final. The availability of these items will ease the researcher in finding good items and eliminating the bad ones. In this instrument, the researcher formulated 60 items, producing a final number of 55 items.

The researcher used a Likert scale that used the word 'favorable.' This word supports or favors the subjects in this research.

The subjects of this research are teenagers aged 16-18 in Yogyakarta. The number of subjects is 120 people. Subject selection was made randomly using the 'google forms' application to provide the subjects with ease in filling the survey.

The data analysis technique used in this research to analyze happiness scale is correlation technique. This method correlates the item scores with the total scale score. The items that fulfill the criteria are those items that have positive correlation or have correlation value above 0.2. In construct validity, the analysis was conducted using factor analysis. Factor analysis was used to confirm the factors that form the scale. The items that fulfill the criteria will correlate positively with similar components or aspects and have low correlation with other components or aspects.

Findings and Discussion

Item Validity Testing

Validity is affiliated with accuracy and suitability (Anastasi & Urbina, 1997). Therefore, the concept of validity is the absolute requirement that must be fulfilled in the construction of a measurement tool. For that reason, a study of validity is a very important point to continue to maintain the scientific credibility of education and psychological evaluation.

In the process of validating this instrument, the researcher used 8 expert judgments to measure Aiken's coefficient that has a range of 0-1. The results of validity testing analysis on the items in the instrument constructed by the researcher are presented in Table 1.

Based on the analysis of Aiken's validity as presented in Table 1, it can be concluded that there are no items with a value of lower than 0.5. This means that all of the items in this instrument is adequate. Therefore, no items were eliminated during the process of validity.

Table 1. The results of validity testing analysis on the items in the constructed instrument

No.	∑	V
1	21	0.875
2	19	0.792
3	19	0.792
4	19	0.792
5	23	0.958
6	21	0.875
7	22	0.917
8	21	0.875
9	21	0.875
10	21	0.875
11	24	1
12	20	0.833
13	22	0.917
14	23	0.958
15	23	0.958
16	22	0.917
17	22	0.917
18	20	0.833
19	20	0.833
20	18	0.75
21	22	0.917
22	20	0.833
23	20	0.833
24	19	0.792
25	24	1
26	22	0.917
27	23	0.958
28	22	0.917
29	21	0.875
30	22	0.917
31	21	0.875
32	21	0.875
33	23	0.958
34	22	0.917
35	22	0.917
36	23	0.958
37	22	0.917
38	23	0.958
39	22	0.917
40	19	0.792
41	19	0.792
42	20	0.833
43	22	0.917
44	24	1
45	22	0.917
46	23	0.958
47	23	0.958
48	20	0.833
49	22	0.917
50	22	0.917
51	17	0.708
52	23	0.958
53	23	0.958
54	23	0.958
55	23	0.958
56	20	0.833
57	19	0.792
58	22	0.917
59	23	0.958
60	22	0.917

Instrument Testing

During the instrument testing process, from the 122 sets of the data received, 120 data were selected from 120 subjects and two sets of data from the other subjects were considered corrupt due to incorrectly filled out survey forms. Using these sets of data, the researcher conducted reliability testing analysis, difference testing, validity testing or factor analysis. Analyses on this instrument were not conducted in parallel. They were conducted separately based on the number of aspects that are in this instrument, which are seven aspects. Therefore, the researcher can decide which items are adequate or fulfill the criteria, and which ones are not adequate. The items that fulfill the criteria persist and the items that do not are eliminated. Consequently, the researcher can determine the reliability of the instrument.

Difference Testing

The first step in the analysis process is difference testing. Difference testing which was used on the items in this instrument was correlation technique with a total score. The results of the difference testing on the instrument which was constructed by the researcher on each aspect are clearly presented in Table 2 until Table 8.

Table 2. The results of the difference testing on the constructed instrument on the aspect of personality

No.	Personality
1	0.493
2	0.336
3	0.386
4	0.332
5	0.122
6	0.452
7	0.375
8	0.483
9	0.419
10	0.423
11	0.293
12	0.239
13	0.432
14	0.361

Reliability: 0.757

Table 3. The results of the difference testing on the constructed instrument on the aspect of culture

No.	Culture
15	0.325
16	0.290
17	0.242
18	0.311
19	0.407
20	0.515

Reliability: 0.614

Table 4. The results of the difference testing on the constructed instrument on the aspect of social support

No.	Social Support
21	0.300
22	0.156
23	0.248
24	-0.023
25	0.446
26	0.372
27	0.491
28	0.183
29	0.364
30	0.460
31	0.536
32	0.556

Reliability: 0.706

Table 5. The results of the difference testing on the constructed instrument on the aspect of friendship

No.	Friendship
33	0.452
34	0.223
35	0.187
36	0.595
37	0.467
38	0.801
39	0.793
40	0.698
41	0.736
42	0.759
43	0.779
44	0.704

Reliability: 0.891

Table 6. The results of the difference testing on the constructed instrument on the aspect of health

No.	Health
45	0.555
46	0.469
47	0.499
48	0.421
49	0.575
50	0.532
51	0.501

Reliability: 0.785

Table 7. The results of the difference testing on the constructed instrument on the aspect of religion and spirituality

No.	Religion and Spirituality
52	0.623
53	0.621
54	0.716

Reliability: 0.805

Table 8. The results of the difference testing on the constructed instrument on the aspect of cooperation

No.	Cooperation
55	0.285
56	0.649
57	0.633
58	0.590
59	0.543
60	0.599

Reliability : 0.793

The data presented in Table 2, Table 3, Table 4, Table 5, Table 6, Table 7, and Table 8 are the results of difference testing analysis based on each aspect. The items that fulfil the criteria correlate positively or have a correlation value above 0.2 can support the function of the scale. The higher the correlation value, the more adequate the item. Those data also show the reliability of the scale on each aspect, such as (1) Reliability of Personality is 0.757, (2) Reliability of Culture 0.614, (3) Reliability of Social Support is 0.706, (4) Reliability of Friendship is 0.891, (5) Reliability of Health is 0.785, (6) Reliability of Religion and

Spirituality is 0.805, and (7) Reliability of Co-operation is 0.793.

Factorial Validity

Factorial validity is a construct valid-ity, which is a validity based on field data. These data are used to find the relationship with the theory used to construct the scale. Azwar (2013) explains that factor analysis is a complex mathematical procedure used to see the relationship between variables. The factor analysis results on the instrument constructed by the researcher can be seen in Table 9.

Table 9. Results of factor analysis on the constructed instrument

Aspect/Factor	Item
Personality	
1	1, 2, 3, 11, 12
2	4, 13, 14
3	6, 7, 8, 9, 10
Culture	
1	15, 16
2	17, 18, 19, 20
Social support	
1	21, 23, 29
2	25, 26
3	27, 30, 31, 32
Friendship	
1	33, 34, 37
2	36, 38, 39, 40, 41, 42, 43, 44
Health	
1	45, 46
2	47, 48, 49, 50, 51
Religion and Spirituality	
1	52, 53, 54
Cooperation	
1	55, 56, 57
2	58, 59, 60

Based on Table 9, it can be concluded that each aspect has differing numbers of fac-tors or indicators. Personality has three indi-cators, culture has two indicators, social sup-port has three indicators, friendship has two indicators, health has two indicators, religion and spirituality has one indicator, and cooper-ation has two indicators.

Final Scale Construction

The construction of the final scale can be done if the factor analysis process has been conducted, because factor analysis produces data about the most adequate items. The items with high correlation with another fac-tor and a low correlation with another factor will be combined into one factor. In the preliminary instrument, there were six indica-tors of personality, three indicators of culture, five indicators of social support, five indica-tors of friendship, three indicators of health, two indicators of religion and spirituality, and four indicators of cooperation. However, af-ter conducting factor analysis, several changes were made. The data presented in Table 10 are the factor analysis results of the 60 items formulated by the researcher. Five items were not adequate or did not fulfil the criteria, having a correlation value of $\leq$ 0.2, which indicates negative correlation, thus was eliminated.

Table 10. Summary of the results of factor analysis on the constructed instrument

Aspect	Component/ Indicator	Item
Personality	Responsibility	1, 2, 3, 11, 12
	Socio-emotion	4, 13, 14
	Emotional stability	6, 7, 8, 9, 10
Culture	Social similarity	15, 16
	Values or beliefs	17, 18, 19, 20
Social support	Emotional support	21, 23,29
	Appreciative support	25, 26
	Social network support	27, 30, 31, 32
Friendship	Conflict and betrayal	33, 34, 37
	Familiarity and problem solving	36, 38, 39, 40, 41, 42, 43, 44
Health	Physical health	45, 46
	Mental, social and economic health	47, 48, 49, 50, 51
Religion and spirituality	Faith and worship	52, 53, 54

| Cooperation | Similarity of target, role and responsibility | 55, 56, 57 |
| | Tolerance and mutual benefit | 58, 59, 60 |

Reliability

The main idea behind reliability is the achievement of consistency in every measurement. Reliability is considered a score property, rather than a test property, which means that the score obtained from the analysis of a test can vary depending on the factors that influence it (subject method and characteristic) (Feldt & Brennan, 1989). Reliability is as important as validity in every measurement tool. A tool that is reliable can maintain consistency of measurement in a period and between raters (Robert, 2006).

Final reliability is determined from the remaining adequate items that correlate positively, which are 55 items. Table 11 presents the result of reliability analysis in instrument designed by the researcher.

Table 11. Result of reliability analysis in the designed instrument

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.935	0.936	55

Based on Table 11, it is determined that the internal consistency of the reliability of the scale is 0.936. This shows that the instrument has a good reliability.

Conclusion

The results of difference or correlation testing on each item yield a total score that shows that 5 items were eliminated due to having a negative correlation or a correlation score ≤ 0.20, bringing the total number of items in the instrument to 55 items. During the factorial validity test or factor analysis test, several items were separated from the original indicator or preliminary factor and combined with another indicator or factor, but these items can still reveal the indicators and aspects in this scale. The reliability testing score is 0.936, meaning that the scale has a good reliability.

References

Anatasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New York, NY: Macmillan.

Arnett, J.J. (1999). Adolescent storm and stress, reconsidered. *American Psychology*, 54(5). 317-326.

Carr, A. (2004). *Positive psychology. The science of happiness and human strengths*. New York, NY: Brunner Routledge.

Diener, E., Suh, E., & Oishi, S. (1997). *Recent findings on subjective well-being*. Retrieved on June 29, 2004, from http://www.psych.uiuc.edu/ -diener/hottopic/paper1.html, 1997.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. H. Linn (Ed.), *Educational measurement* (3rd ed.). New York, NY: Macmillan.

Kurniawan, I. N. (1998). Kecenderungan berperilaku delinkuen pada remaja ditinjau dari orientasi religius dan jenis kelamin. *Psikologika*, 6, 55-65.

Myers, D. G. & Diener, E. (1995). Who is happy? *Psychological Science*, 6, 10-19. Retrieved on June 21, 2004, from http://www.acsu.buffalo. edu/-kashdan/happy.pdf

Robert, P. & Priest, H. (2006). Reliability and validity in research. *Journal of Nursing Standard*, 20, 41-45.

Seligman, E. (2005). *Authentic happiness*. Bandung: PT. Mizan Pustaka.

Taggart, F., Friede, T., Weich, S., Clarke, A., Johnson, M., & Stewart-Brown, S. (2013). Cross cultural evaluation of the Warwick-Edinburgh mental well-being scale (WEMWBS) – A mixed methods study. *Health and Quality of Life Outcomes*, 11(7), 1-12 .doi: https://doi.org/10.1186/1477-7525-11-27