**Validity and Reliability of Happiness Scales**

**Deni Cahyono**

1Department of psychology, 17 August 1945 University of Surabaya, Surabaya, East Java, Indonesia

**ABSTRACT**

This study aims to develop a happiness scale by applying the Content Validity Index (CVI) method and testing the validity and reliability of the instrument. The research subjects were 53 respondents selected through a purposive sampling technique, and the data were collected using Google Forms. The CVI test results showed a CVR score for the clarity aspect of 0.897, the relevance aspect of 0.903, and the sufficiency aspect of 0.886. These values indicate that the developed scale has good construct validity. Item validity was conducted using Ms. Excel, showing that 17 items have validity scores above 0.3, while the scale's reliability reaches 0.83 for 17 existing items. Based on the results of the validity and reliability tests, it can be concluded that the happiness scale developed has good validity and reliability.

**Keywords:** Happiness scale, CVI, validity, reliability, Cronbach's Alpha

1. **INTRODUCTION**

The scientific study of happiness started from philosophy to modern psychology. In his work 'Nicomachean Ethics' (350 BC), Aristotle put forward the concept of eudaimonia; in this case, eudaimonia is a concept where happiness is considered the primary goal of life. According to Aristotle, happiness can be achieved through virtue and self-actualization. Meanwhile, humanistic psychologist Maslow (1954) developed a hierarchy of needs that states that self-actualization is the pinnacle of happiness. In the modern era, Seligman (2000) introduced the concept of measurable happiness that can be increased through positive interventions.

According to Seligman (2011), happiness is a concept that refers to positive emotions and positive activities favored by individuals. Meanwhile, Irianto & Subandi (2015) stated happiness is a concept that describes an individual's condition when directing his feelings to something positive and utilizing his positive character to interpret the events he lives in everyday life. Ng (2022) defines happiness as positive feelings that are more dominant than negative feelings. Encyclopaedia Britannica explains happiness in a narrow sense as the emotional well-being that a person experiences when good things happen at a particular moment. In a broader sense, it is defined as a positive evaluation of one's life and achievements. Happiness can be distinguished from negative emotions (such as sadness, fear, and anger) and other positive emotions (such as affection, joy, and interest).

Recent research on happiness suggests that social, emotional, and psychological factors significantly influence individual well-being. A study by Seligman et al. (2011) revealed that the practice of virtues, such as gratitude and being kind to others, can substantially increase happiness levels. In addition, research by Diener and Seligman (2005) highlighted the importance of quality social relationships, where individuals with strong social support tend to report higher happiness levels. Recent research has also emphasized the critical role of work-life balance, with one study suggesting that flexibility in work schedules can increase employees' life satisfaction and reduce stress (Krekel et al., 2022). In addition, research by Helliwell et al. (2022) confirmed that a sense of purpose in life and engagement in meaningful activities contribute significantly to long-term happiness. These findings suggest that happiness results from economic conditions, interpersonal relationships, and a more profound quality of life.

Measuring happiness plays a vital role in understanding individual and societal well-being. Firstly, happiness measurement provides deep insight into a person's quality of life beyond economic figures such as income and education level. This allows researchers and policymakers to evaluate the impact of social and economic policies on people's well-being (Diener, 2000). In addition, happiness data can be used to design intervention programs that are more effective in improving mental and emotional health, thus helping individuals cope with stress and increase life satisfaction (Ryan & Deci, 2001). In the work environment, the level of employee happiness is directly proportional to productivity, so this measurement is essential for companies to create a positive work atmosphere (Helliwell & Wang, 2011). Happiness measurement helps create a more inclusive and supportive community, focusing on collective well-being and encouraging further research in psychology and sociology.

The Theory of Happiness, developed by Martin Seligman (2011), is an essential framework in positive psychology that explains the critical components of individual happiness. According to Seligman, individual happiness consists of three elements: Positive Emotion, Engagement, and Meaning. Positive emotions encompass happy experiences that improve quality of life, while engagement refers to experiences of ‘flow’ where individuals are fully engaged in their activities. Meaning gives individuals purpose in their lives, which makes them feel connected to something bigger than themselves.

1. **METHODOLOGY**

In this study, the researcher determined the research steps as follows:

1. Determining the Concept to be Measured

The concept must have a solid theoretical basis so that the scale has good concept validity (Anastasi & Urbina, 1997). In this study, the concept of happiness refers to Martin Selingman's definition, which states that happiness is a state in which a person experiences positive emotions and satisfying activities. According to Seligman, authentic happiness comes from positive self-assessment and the development of fundamental strengths associated with positive emotions.

1. Operational Definition

Once the main concepts are identified, the researcher must create clear operational definitions of each component or aspect of the concept. These operational definitions will determine how the variables are precisely measured (DeVellis, 2016). According to Martin Seligman, happiness has three main components: Positive Emotion, Engagement, and Meaning. Positive emotions include pleasant experiences such as joy, love, and gratitude, which provide happiness in everyday life. However, these emotions alone are not enough. Engagement occurs when one is fully engaged in a challenging yet fulfilling activity, often called flow, where time seems to fly by. Meaning is finding a purpose greater than oneself, such as contributing to society or following spiritual values. Deep happiness is achieved by combining these three components.

1. Item construction

Items should represent aspects of the defined concept. In constructing items, it is crucial to ensure that the language used is clear, unambiguous, and easily understood by respondents (DeVellis, 2016). In this case, the researcher compiled 18 items. Each indicator is represented by six items: three favorable and three unfavorable. The blueprint for this happiness scale is as follows:

**Tabel 1. Blueprint Happines Scale**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Indicator** | **Favorabel** | **Unfaforabel** | **Sum** |
| 1 | Positif Emotion | 1,2,3 | 4,5,6 | 4 |
| 2 | Enggangement | 7,8,9 | 10,11,12 | 4 |
| 3 | Meaning | 13,14,15 | 16,17,18 | 4 |

1. Determining the Rating Scale

Select the type of rating scale to be used, such as a Likert scale, semantic differential scale, or other appropriate format. Likert scales are commonly used in psychological scales because they allow respondents to indicate their degree of agreement or disagreement with a statement (Likert, 1932). This happiness scale uses the Likert scale method with the following information:

**Tabel 2: Scoring system Happines Scale**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Answer* | | | |
|  | Strongly agree | Agree | Disagree | Strongly disagree |
| *Unfaavorabel* | 1 | 2 | 3 | 4 |
| *Favorabel* | 4 | 3 | 2 | 1 |

1. Conduct a Content Validity Test

To ensure that the scale covers all relevant aspects of the measured concept, conduct content validation by asking experts in the relevant field to review the scale items. This ensures that the items reflect essential aspects of the measured concept (Haynes et al., 1995). In the content validity test, the researcher used professional judgment. The professionals involved in this study were ten professionals with a background in master's degree.

1. Collecting Trial Data

After the scale is completed, the next step is to test the scale on a small representative sample. The goal is to test the feasibility of the items and the suitability of the scale format and to detect possible problems before being used on a larger scale (DeVellis, 2016). The trial was conducted with 53 respondents, with the respondents' age above 20 years. The data collection method is by using Google Forms.

1. Testing Validity and Reliability

After the data is collected, a reliability test is carried out to determine the consistency of the scale. The commonly used reliability test is Cronbach's alpha coefficient test, which assesses the extent to which the items in the scale have internal consistency (Cohen & Swerdlik, 2018).

1. **ANALYSIS**
2. CVI test Result

**Table 3: Result CVI Analysis**

|  |  |  |  |
| --- | --- | --- | --- |
| **Aspect** | **CVR** | **P** | **Aiken Ideal Value** |
| Clearity | 0.879 | 0.04 | 0.75 |
| Relevancy | 0.903 | 0.04 | 0.75 |
| Suficiency | 0.886 | 0.04 | 0.75 |

For clarity, it is known that the CVR value is 0.897; in this case, the value is above the required ideal value. In terms of clarity, this scale is of good quality. In contrast, The CVR value of the relevancy aspect is 0.903, with a value of significance level of p = 0.4. The ideal Aiken index that must be achieved is at least 0.75. So, for the relevancy aspect, this scale is of good quality. As for the sufficiency aspect, it has a value of 0.886. This indicates that the sufficiency aspect also has good relevance.

1. Validity Test Results

Researchers use MS Excel to process data and conduct a validity test. The correlation validity test determines how much a measurement instrument can measure what should be measured. This validity is often done through correlation analysis between the score on each item in the instrument and the total score. This approach is known as internal validity or item-total correlation validity. The higher the correlation coefficient between the item and the total score, the better the item's validity, which indicates that the item is consistent with the overall test in measuring the same construct. The correlation coefficient can be calculated using the Pearson Product Moment method for interval or ratio data and Spearman Rank for ordinal data. Testing the validity of items through correlation is essential to maintain the quality of the instrument so that it can provide valid and reliable results in research (Azwar, 2012). The following are the results of the validity test calculation.

**Tabel 4: Item Validation Analysis results**

|  |  |  |  |
| --- | --- | --- | --- |
| **Item** | **Corelation** | **Standart** | **Noted** |
| 1 | 0,59 | 0,3 | Valid |
| 2 | 0,58 | 0,3 | Valid |
| 3 | 0,60 | 0,3 | Valid |
| 4 | 0,47 | 0,3 | Valid |
| 5 | 0,51 | 0,3 | Valid |
| 6 | 0,38 | 0,3 | Valid |
| 7 | 0,53 | 0,3 | Valid |
| 8 | 0,24 | 0,3 | Not Valid |
| 9 | 0,74 | 0,3 | Valid |
| 10 | 0,55 | 0,3 | Valid |
| 11 | 0,48 | 0,3 | Valid |
| 12 | 0,76 | 0,3 | Valid |
| 13 | 0,49 | 0,3 | Valid |
| 14 | 0,41 | 0,3 | Valid |
| 15 | 0,49 | 0,3 | Valid |
| 16 | 0,46 | 0,3 | Valid |
| 17 | 0,38 | 0,3 | Valid |
| 18 | 0,49 | 0,3 | Valid |

From the calculations above, it is known that one of the 18 items is canceled, namely item number 8.

1. Reliability test results

Based on the reliability results that have been carried out previously, researchers conducted two reliability tests. The first test with complete items totaled 18 items, while the second test used 17 items; one item was eliminated because the reliability score was less than the predetermined standard. The results for Cronbach's Alpha show how well the items in the instrument are related and measure the same construct. Cronbach's Alpha values range from 0 to 1, with higher values indicating better reliability. As a rule, alpha values above 0.7 are considered sufficiently reliable, although values above 0.6 are considered adequate in some contexts (Ghozali, 2016).

**Tabel 5: Item Validation Analysis results (18 item)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 229,12 | 52,00 | 4,41 | 5,51 | 0,00 | 1,36 |
| Columns | 133,10 | 17,00 | 7,83 | 9,80 | 0,00 | 1,63 |
| Error | 706,51 | 884,00 | 0,80 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 1068,73 | 953,00 |  |  |  |  |
|  |  |  |  |  |  |  |
| **Cronbach's Alpha** | **0,82** |  |  |  |  |  |

**Tabel 5: Item Validation Analysis results (17 item)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 231,15 | 52,00 | 4,45 | 5,88 | 0,00 | 1,36 |
| Columns | 102,98 | 16,00 | 6,44 | 8,52 | 0,00 | 1,66 |
| Error | 628,55 | 832,00 | 0,76 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 962,67 | 900,00 |  |  |  |  |
|  |  |  |  |  |  |  |
| **Cronbach's Alpha** | **0,83** |  |  |  |  |  |

The reliability results show scores of 0.82 and 0.83, indicating that this happiness-measuring instrument is highly reliable.

1. **RESULTS AND DISCUSSION**

CVR is used to evaluate content validity, which is the extent to which the items in the instrument measure all relevant aspects of the construct being measured. This measuring instrument's CVR score is in the good category. This indicates that the items in the measuring instrument have been assessed as relevant by experts in accordance with the desired measurement objectives. This means that the instrument has good content relevance, vital measurement accuracy, and consistency in measurement results, so it can be trusted to provide valid and reliable data collection results.

High score validity indicates that the measuring instrument can accurately measure what should be measured. If the validity is high, the results from the measuring instrument reflect the phenomenon to be measured so that the resulting data is relevant and accurate for analysis purposes.

High reliability means that the measuring instrument is consistent in its measurements. Reliability refers to the stability and consistency of the results obtained when the measuring instrument is used repeatedly under the same conditions. If the reliability is high, the instrument can produce similar scores over time or among different raters, ensuring reliable measurement results.

1. **REFERENCES**
2. Anastasi, A., & Urbina, S. (1997). Psychological testing (7th ed.). Upper Saddle River, NJ: Prentice Hall.
3. Azwar, S. (2012). Reliabilitas dan validitas. Yogyakarta: Pustaka Pelajar.
4. DeVellis, R.F. (2016) Scale Development: Theory and Applications. Vol. 26, Sage, Thousand Oaks.
5. Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. American Psychologist, 55(1), 34–43. https://doi.org/10.1037/0003-066X.55.1.34
6. Diener, E., & Seligman, M. E. (2002). Very happy people. Psychological science, 13(1), 81–84. https://doi.org/10.1111/1467-9280.00415
7. Ghozali, Imam. (2016). Aplikasi Analisis Multivariete Dengan Program. IBM SPSS 23 (Edisi 8). Cetakan ke VIII. Semarang: Badan Penerbit. Universitas Diponegoro
8. Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological Assessment, 7, 238-247.http://dx.doi.org/10.1037/1040-3590.7.3.238
9. Helliwell, J. F., Huang, H., Norton, M., Goff, L., & Wang, S. (2023). World happiness, trust, and social connections in times of crisis. In World Happiness Report 2023 (11th ed., Chapter 2). Sustainable Development Solutions Network.
10. Irianto & Subandi. (2015). Studi Fenomenologis Kebahagiaan Guru di. Papua. Gadjah Mada Journal of Psychology,1 (3): 140-166
11. Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, 22 140, 55.
12. Maslow, A. H. (1954). Motivation and personality. Harper & Row.
13. Ng, Y. K. (2022). Happiness: Concept, measurement and promotion. Springer Nature.
14. Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. Annual Review of Psychology, 52, 141–166. https://doi.org/10.1146/annurev.psych.52.1.141
15. Seligman, M. E. P. (2000). Learned optimism: How to change your mind and your life. Pocket Books.
16. Seligman, M. E. P. (2011). Flourish: A visionary new understanding of happiness and well-being. Free Press.