Journal of Psychology and Behavioral Science
June2019, Vol. 7, No. 1, pp. 74-83
ISSN: 2374-2380 (Print), 2374-2399 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jpbs.v7n1a9
URL: https://doi.org/10.15640/jpbs.v7n1a9

Why Do Researchers and Educators Still Use the Rosenberg Scale? Alternative New Concepts and Measurement Tools for Self-Esteem

Katsuyuki Yamasaki[1]

Abstract

Rosenberg's Self-Esteem Scale (RSES) has been the most frequently and widely used tool in self-esteem research and education. Until recently, findings on self-esteem have steadily accumulated using this scale. However, in recent years, more studies began to indicate various faults of the RSES, particularly in that it does not measure what Rosenberg (1965) demonstrated as the concept of desirable self-esteem. In this paper, first, it isdiscussed why the RSES should not be used in terms of validity. Thereafter, recent new concepts of self-esteem are introduced, including concepts by Deci and Ryan (1995), Kernis (2003), and Yamasaki et al. (2017). Next, according to Yamasaki et al. (2017), it is suggested that desirable self-esteem testing is needed using certain non-conscious measuring tools because self-report (i.e., consciously answered) questionnaires cannot capture desirable components of self-esteem. Regarding new assessment tools, newly developed implicit association tests (IATs) that are completely distinct from previous ones are discussed in terms of their measurementof desirable self-esteem. Finally, it is discussed how programs to enhance desirable self-esteem should operate, given their effectiveness. Thus, the current paper emphasizes the necessity to disseminate new concepts, measuring tools, and programs to correct wrong avenues on which researchers and educators have been treading.

Keywords: self-esteem, Rosenberg's Self-Esteem Scale, validity, measurement, education

1. Introduction

Self-esteem has been widely and frequently investigated in the domains of psychology and education. Even now a great amount of research covers this topic. In self-esteem research, valid and reliable psychological assessment tools are essential since it is a genuine psychological construct. Regarding assessment tools, the Rosenberg's (1965) Self-Esteem Scale (RSES) is the questionnaire that has been utilized most frequently worldwide.

Although it has been almost 50 years since the RSES was developed, it is still the most frequently used measure for self-esteem. Notwithstanding that many limitations have been discovered with respect to the RSES, the trend of research on self-esteem using the RSES has not declined. Researchers may consider that the limitations do not deserve note. Moreover, although the limitations that have been indicated are not ones to be dismissed, more serious limitations that cannot be revised exist in the RSES. This review indicates such fundamental limitations that researchers can never ignore, suggesting new types of assessment to overcome them, along with new conceptualizations of self-esteem.

2. What Does the Rosenberg Self-Esteem Scale Measure?

2.1 The Validity of the Rosenberg Self-Esteem Scale

[1] Department of Psychology and Educational Science, Naruto University of Education, Naruto, Tokushima, Japan 772-8502. e-mail: ky341349@naruto-u.ac.jp. office phone: +81-88-687-6240

The RSES suffers from an unclear developmental process. Although the RSES was originally a Guttman scale, most of the previous studies have utilized it as a Likert scale. To our knowledge, it is uncertain when the scale started to be used as a Likert scale, or when the order of the question items in the scale arose (see Table1). According to Gravie (1979), it seems likely that Rosenberg himself identified the present common format of the scale in 1979 at the latest, agreeing on the Likert modification.

Rosenberg (1965) defined self-esteemas a positive or negative attitude toward the self, and divided self-esteem into feeling "very good" and "good enough." The "very good" type is characterized by the feeling that one is superior to others, while the "good enough" type is depicted as follows: (1) "the individual respects himself, considers himself worthy; he does not necessarily consider himself better than others, but he definitely does not consider himself worse" (p. 31); (2) "he does not feel that he is the ultimate in perfection but, on the contrary, recognizes his limitations and expects to grow and improve" (p.31). He developed the RSES to measure the "good enough" type of self-esteem as a desirable trait that is adaptive and healthy. In contrast, the "very good" type of self-esteem is an undesirable trait that is nonadaptive and unhealthy. The 10 items in the RSES are shown in Table 1. However, when scrutinizing the contents of the items in terms ofcontent validity, it seems likely that most of the items do not represent the "good enough" type (Yamasaki, Yokoshima, & Uchida, 2017). To put it concretely, among the items in the RSES, items 4, 7, 6, and 9 cannot discriminate between "very good" and "good enough." For items 6 and 9, persons with "very good" self-esteem would answer with higher ratings than those with "good enough" self-esteem. Moreover, no items could assess the second characteristic above, although only item 10 might slightly touch on this characteristic. In line with these considerations, it is highly possible that the RSES includes no content validity, indicating that the RSES does not measure what Rosenberg defined regarding desirable self-esteem. Regarding the lack of content validity, Kernis (2003) also indicated that the RSES could not discriminate his optimal self-esteem (secure high self-esteem) that is similar to the "good enough" type and fragile high self-esteem that is similar to the "very good" type.

Table 1. Ten Items and Instructions for Rosenberg's Self-Esteem Scale (RSES) (adapted from Rosenberg, 1965)

Below is a list of statements dealing with your general feelings about yourself. Please indicate how strongly you agree or disagree with each statement.*
1. On the whole, I am satisfied with myself.
2. At times I think I am no good at all.
3. I feel that I have a number of good qualities.
4. I am able to do things as well as most other people.
5. I feel I do not have much to be proud of.
6. I certainly feel useless at times.
7. I feel that I'm a person of worth, at least on an equal plane with others.
8. I wish I could have more respect for myself.
9. All in all, I am inclined to feel that I am a failure.
10. I take a positive attitude toward myself.

* Four alternatives for answering: Strongly Agree, Agree, Disagree, Strongly Disagree

2.2Other Studies on the Validity of the RSES

The RSES includes internal consistency and test-retest reliability (Donnellan, Tresniewski, & Robins, 2015), but the findings concerning validity are complicated. In prior studies, the validity of the RSES has been examined in terms of convergent and divergent validity. Regarding convergent validity, the concurrent validity has been most frequently examined. For instance, examinations with the Self-Liking and Self-Competence Questionnaire (Tafarodi & Swann, 2001) by Falk, Heine, Takemura, Zhang, and Hsu (2015), Self-Perception Profiles for Adolescents (Harter, 1985) by Hagborg (1993), and Coopersmith Self-Esteem Inventory (Coopersmith, 1967) by Griffiths et al. (1999) were conducted. However, in these studies, the measures to compare with the RSES includeinsufficient validity in which adaptive and nonadaptive components of self-esteem are not distinguished.

Although the Coopersmith Inventory seems to examine validity in terms of the adaptive component, the participants in Falk et al.'s (2015) study were not elementary graders for which the inventory was developed, and it was unclear where they extracted the 25-item version instead of the original 58-item one.

With respect to convergent validity,relationships with health and adaptation status have often been examined. Although positive relationships have been reported (e.g., Blascovich & Tomaka, 1991; Donnellan, Tresniewski, & Robins, 2015), such positive relationships cannot provide evidence forvalidity because previous results concerning these relationships are inconsistent. Another method to confirm validity is via evaluation by others, which is a promising method in providing different views ofvalidity distinct from self-reports. Few studies have adopted this method to examine validity, but Demo (1985) reported extensive research, revealing that the results of the two among the methods used (i.e., Q-sort and assessment using only one item by peers) were positively correlated with the RSES. However, Demo himself described that the validity of the Q-sort method is unclear, and the one-item evaluation by others was just an evaluation using a 5-point Likert scale from 1 = "low self-esteem" to 5 = "high self-esteem", which was low in precision because the word "self-esteem" would be interpreted differently by individuals.

Next, concerning divergent validity, negative correlations with depression (e.g., Sinclair, Blais, Gansler, Sandberg, Bistis, & LoCicero, 2010; Wegener et al., 2015) and neuroticism (Robins, Tracy, Trzesniewski, Potter, & Gosling, 2001; Watson, Suls, & Haig, 2002) were reported. However, like positive correlations with happiness, the contents of the items used in measuring depression and neuroticism are reversely similar to those of the items used in the RSES, which suggests that those findings cannot provide evidence of validity. Moreover, although Blascovich and Tomaka (1991) considered no significant correlations with academic records by Reynolds (1988) and no relations with sex, age, marital status, and birth order by Fleming and Courty (1984) as evidence for divergent validity, those exact relationships with self-esteem are still undetermined.

Taken together, it can be concluded that the RSES lacks validity. So far as the contents of the RSES include both adaptive and nonadaptive components, it was predictable. However, although most of the above considerations suggestthe invalidity of the RSES, they do not directly show that the RSES lacks validity. Therefore, we need to refer to some recent studies that directly reveal a lack of validity for the RSES.

2.3 Empirical Research Showing Lack of Validity for the RSES

Izuma, Kennedy, Fitzjohn, Sedikides, and Shibata (2018) examined the relationships between neural activities in reward-related brain regions and scores on the implicit association test (IAT) for self-esteem versus the RSES. Data analyses were conducted using multi-voxel pattern analysis on functional MRI data. Neural activities were driven by responses to participants' own facial photos. Results showed that the RSES scores were not correlated with neural signals in the larger reward-related brain regions, although IAT scores were positively correlated with them. Activation of reward-related brain regions when watching one's own face is regarded to be a natural nonconscious response to oneself. The finding that this activation was not associated with the scores of the RSES suggests that the RSES does not measure true affirmative attitude toward oneself. The IAT for self-esteem is depicted in detail later in this article.

Another direct examination of the validity of the RSES was conducted using a nomination method by school teachers, targeting the children's version of the RSES (Yokoshima, Uchiyama, Uchida, & Yamasaki, 2019). Homeroom teachers at elementary schools nominated the highest and lowest two or three children in the homeroom class for each of the following three characteristics: (1) they are confident in themselves, and not competitive with other children; (2) they are confident in themselves, and competitive with other children; and (3) they are not confident in themselves. No. 1 represents a core characteristic of the high "good enough" type, and No. 2 shows the high "very good" type of self-esteem. The children in No. 3 are neither high in the"good enough" category nor high in the"very good" category,andthis representslow self-esteem. If the RSES represents not only "good enough" self-esteem but also "very good" self-esteem, the children nominated in Nos. 1 and 2 would equally show higher scores than those in No. 3. If the RSES represents either "good enough" self-esteem or "very good"self-esteem, the scores in Nos. 1 and 2 would be different. Results showed that the scores of the children nominated in Nos. 1 and 2 were not significantly different, and that they were higher than those in No. 3, which suggests that the RSES includes both the "good enough" and "very good" types of self-esteem.

Moreover, Yokoshima, Kaya, Uchida, and Yamasaki (2018) conducted a study to examine the effectiveness of a prevention program that can enhance the "good enough" type of self-esteem and decrease the "very good"type. The program was implemented in 4th-grade elementary school children each week for one month.

Results revealed that the scores of the RSES did not significantly change from pre- to post-intervention, although the scores of the "good enough" type of self-esteem that was measured using the IAT significantly increased. This finding also suggests that the RSES includes both the "very good" and "good enough" types of self-esteem. As the program included the power to enhance adaptive self-esteem and decrease nonadaptive types, the scores of the RSES did not change due to the balanced effects. Additionally, Jordan, Spencer, Zanna, Hoshino-Browne, and Correll (2003) compared the scores of self-esteem by the IAT and the RSES in terms of narcissism and defensive behaviors (in-group bias and dissonance reduction). Results showed that participants high in explicit SE (i.e., the RSES) but low in implicit self-esteem showed the highest levels of narcissism and defensiveness. This finding suggests that the RSES measures nonadaptive characteristics of self-esteem.

From the examination regarding the construct validity of the RSES above, it could be concluded that it is not clear what the RSES measures, since the scores of the RSES include both adaptive and nonadaptive components of self-esteem. At present, we cannot clarify to what extent each of the two types of self-esteem is included.

Also, apart from validity, various limitations have been indicated about the RSES, including the factor structure where more than one factor has been indicated (e.g., Kaplan & Pokorny, 1969; Richardson, Ratner, & Zumbo, 2009), although the RSES was originally developed with one factor. Furthermore, different numbers of Likert options (four to six) and different translations of the items even in one language exist at present. With these variations, a great amount of research has been conducted worldwide, and despite such variations, all the findings have been regarded as concerning self-esteem by Rosenberg. In particular, these confounded conditions with respect to the RSES become serious problems in cross-cultural studies. The RSES has been translated into different languages and utilized in many countries. However, there are no guarantees that different versions of the RSES measure the same self-esteem construct. Gnambs, Scharl, and Schroeders (2018) produced data showing a best fit for the model in which the RSES consists of two specific factors for positively and negatively worded items in addition to a general factor. However, they found that the general factor loadings were invariantly higher across samples from the United States and other highly individualistic countries than from less individualistic countries. Thus, they doubted the usefulness of the RSES for cross-cultural comparisons that can lead to seriously biased conclusions.

3. New Classification of Two Types of Self-Esteem

To put the above descriptions simply, Rosenberg (1965) divided two types of self-esteem, "good enough" and "very good",with the "good enough" type being a healthy and adaptive self-esteem. However, his questionnaire cannot measure the "good enough" type of self-esteem. More precisely, the scale includes both types of self-esteem, but the ratio between the two types is unclear, suggesting that what the RSES assesses cannot be clarified. This conclusion casts doubts on the effectiveness of previous studies using the RSES.

Moreover, several other researchers questioned what the RSES measures, and they developed two other new types of self-esteem to divide self-esteem into the adaptive and nonadaptive types more clearly than Rosenberg (1965). For instance, Deci and colleagues (e.g., Deci & Ryan, 1995) distinguished true and contingent self-esteem. People with high true self-esteem feel intrinsic satisfaction with themselves that is relatively stable. They do not reflect on their own self-esteem, being unconscious of it. In more detail, in terms of their cognitive evaluation sub-theory in the self-determination theory (e.g., Ryann & Deci, 2000, 2002), high true self-esteem consists of high autonomy, competence, and relatedness that are all associated with high intrinsic motivation. In contrast, those with high contingent self-esteem are sensitive to achieving criteria that are often set by themselves in comparison with other people. In this meaning, contingent self-esteem is consciously salient to theindividual. True self-esteem is associated with positive outcomes in terms of health, adjustment, and performance, while contingent self-esteem is associated with negative outcomes.

Another distinction was developed by Kernis (2003). He distinguished secure and fragile high self-esteem. Secure self-esteem is not contingent on external criteria and is positively associated with a wide range of adjustment and well-being indices. It is congruent with nonconscious (implicit) self-worth. He named this self-esteem "optimal self-esteem". In contrast, fragile high self-esteem is contingent on external criteria and exhibits an unstable characteristic that is influenced by successes and failures in achieving contingent criteria.

It is discrepant with nonconscious self-worth and nonadaptive. Moreover, he proposed "authenticity" as the theoretical basis of optimal self-esteem that is characterized as the unobstructed operation of one's true self in one's daily enterprise.

Thereafter, in recent years, Yamasaki et al. (2017) developed another new distinction of autonomous and heteronomous self-esteem. They considered that high autonomous self-esteem consists of high levels of self-confidence, confidence in others, and intrinsic motivation. In contrast, high heteronomous self-esteem is characterized by low levels of these three components. The former is regarded as a healthy and adaptive self-esteem, while the latter is unhealthy and nonadaptive. They underscored that all three components need to be high for autonomous self-esteem, and that all three should be low for heteronomous self-esteem. Furthermore, they proposed that autonomous self-esteem cannot be consciously measured using self-report questionnaires, although heteronomous self-esteem can be. Namely, autonomous self-esteem needs to be nonconsciously (implicitly) assessed using methods other than self-reports. Including Rosenberg (1965), the above developers of new concepts of self-esteem suggested that adaptive and healthy self-esteem functions nonconsciously, but no groups clarifiedits nonconscious characteristics except for Yamasaki et al. (2017). The reason would be that in their ages, unlike at present, nonconscious characteristics of human behavioral and psychological functions were not revealed empirically well.

4. How Can Healthy and Adaptive Self-Esteem Be Assessed Nonconsciously?

Yamasaki et al. (2017) posited that since adaptive self-esteem such as autonomous self-esteem is an abstract concept, it is difficult to express it concretely in question items listed on self-reports. Therefore, when it is expressed abstractly such as "are you confident in yourself?", people often refer to concrete behaviors or outcomes in memories that can be resources for answering the question. Those outcomes are often experiences in comparison to others such as having been superior or inferior to others in certain competitive situations. In this process, for instance, the questions that are designed to measure autonomous self-esteem are easily changed into questions to assess heteronomous (contingent) self-esteem. Thus, it is highly possible that adaptive self-esteem such as autonomous self-esteem cannot be assessed using self-reports, while contingent or heteronomous self-esteem that is expressed as concrete behaviors can be. Yamasaki et al. (2017) also considered that the same logic is applied to all self-report questionnaires to measure adaptive and healthy self-esteem such as true and secure self-esteem.

In line with this consideration, certain methods to assess adaptive self-esteem nonconsciously are needed. Projection methods such as TAT and Rorschach tests represent such candidates. In fact, TAT has been used to measure self-esteem (Ackerman, Hisenroth, Clemence, Weatherill, & Flowler, 2001). However, the administration and scoring in projection methods are both time- and labor-consuming, in addition to suffering from a lack of objectivity. Semi-projection methods include similar faults. Another promising candidate is the Implicit Association Test (IAT; e.g., Greenwald & Banaji, 1995). In the IAT, the implicit association between two kinds of stimuli (category and attribute stimuli) is measured. For instance, when assessing racial prejudice against back people, the implicit association between a white (or black) person and a pleasant (or unpleasant) word is measured. In the case of using computers, if reactions in the task in which a participant presses a left (or right) key when a white person or a pleasant word appears on the computer display and a right (or left) key when a black person or an unpleasant word appears are faster and more accurate than those in the task in which a participant presses a left (or right key) when a white person or an unpleasant word appears and a right (or left) key when a black person or a pleasant word appears, racial prejudice against black people is predicted to exist. In general, implicit characteristics are considered as lying on the preconscious domain (Quirin, Kazén, Rohrmann, & Kuhl, 2009). However, to put it more accurately, they should be considered as lying on the nonconscious domain, since the preconscious domain is largely influenced by the unconscious domain.

The IAT attempts to assess various psychological characteristics including self-esteem (cf., Bosson, Brown, Zeigler-Hill, & Swann, 2003). However, the previous IATs to assess self-esteem include several limitations. First, results are different, depending on the opposite category stimuli against oneself. Karpinski (2004) revealed that negative category words against oneself falsely make the scores of self-esteem in the IAT higher compared to the positive category words. Regarding this problem, several studies (e.g., Jordan et al., 2003; Zeigler- Hill, 2006) recommended using neutral category words such as "that" and "it." Second, although attribute stimuli are usually pleasant or unpleasant words in the IAT, the words directly representing the core characteristics of self-esteem such as "confident" and "satisfied" need to be utilized.

It is possible that words such as "joy" and "beautiful" would show positive association with high self-esteem, but the association is indirect and weaker than when using words representing core self-esteem characteristics. From these limitations, it is understood that extant IAT measures of self-esteem do not show any high positive correlations to each other (Bosson, Swann, & Pennebaker, 2000).

5. Development of New IAT Measures

Considering the limitations of previous IATs for self-esteem, Yokoshima et al. (2017) developed a new IAT measure for children (the paper and pencil version of Self-Esteem Implicit Association Test for Children (SE-IAT-C). The reason why they developed the children's version was to examine the effectiveness of a program to cultivate adaptive autonomous self-esteem for children. In developing it, they considered the following two points: (1) it can be simultaneously applied to a group of children using a paper-and-pencil type method, (2) the category words against self are neutral such as "that" and "it", and (3) the attribute words are positive and negative words representing the core characteristics of self-esteem. Nearly according to Jordan et al. (2003), this test consists of seven trial blocks as shown in Table 2. Depending on the instruction, the task is to divide the words into the right or left space by marking with a pencil. The final scores are calculated by the differences between the numbers of right answers in Blocks 3 and 4 versus 6 and 7 (or 6 and 7 versus 3 and 4). Blocks <3 and 4> and <6 and 7> are counterbalanced.

Table 2. Procedure of the Paper and Pencil Version of Self-Esteem Implicit Association Test for Children (SE-IAT-C) (adapted from Yokoshima et al., 2017)

Blocks	Task labels	Task details	Time
1	practice for attribute words	pleasant vs. unpleasant	20 sec.
2	practice for category words	self vs. not-self	20 sec.
3	Scored trial 1	self -pleasant vs. not-self -unpleasant	20 sec.
4	Scored trial 2	self -pleasant vs. not-self -unpleasant	20 sec.
5	practice for category words	self vs. not-self	20 sec.
6	Scored trial 3	not-self -pleasant vs. self -unpleasant	20 sec.
7	Scored trial 4	not-self -pleasant vs. self -unpleasant	20 sec.

After developing the original version of this test, Yokoshima et al. (2017) and Yokoshima, Ogami, Kaya, and Yamasaki (2019) examined test reliability and validity. No previous IAT measures tested the validity directly in terms of the construct itself. In most cases, the validity was examined in concurrent and divergent validity using other measures that often lack validity. Yokoshima, Ogami, et al. (2019) examined the test-retest reliability over an interval of four weeks. The correlations were similar to or higher than previous IAT measures (Bosson et al., 2000; Greenwald & Farnham, 2000). Yokoshima et al. (2017) examined validity using teachers' assessments of children with high and low scores on this test. The teachers were the participants' homeroom teachers who had known the children well. The teachers assessed the children in terms of autonomy, aggression, and anxiety. It was predicted that children with high scores would be assessed as high in autonomy, but low in aggression and anxiety thatarise with low autonomous self-esteem. As predicted, results showed that the teachers rated the children with high autonomous self-esteem as having higher autonomy, being less anxious, and being less aggressive than those with a low autonomous trait.

Most recently, Yokoshima, Ogami, et al. (2019) developed the tablet PC version of their paper-and-pencil version. Compared with the paper-and-pencil version, the tablet PC version needs less time and labor to prepare and implement, automatically providing the scores as soon as the session is over. This version was developed with the android OS using JAVA through the Android Studio.

The category and attribute words and the structure of the trial blocks are the same as in the paper-and-pencil version. Children are highly motivated and proficient at operating tablet PCs. Moreover, Yokoshima, Ogami, et al. (2019) confirmed the reliability and validity of the tablet PC version. Validity was confirmed via homeroom teachers' assessment of the children with high and low scores on the test in terms of autonomous self-esteem using similar methods to Yokoshima et al. (2017).

6. Modification of Autonomous and Heteronomous Self-Esteem by a Universal Program

Once the new concepts and measuring tools with respect to self-esteem are developed, one of the next concerns is education in schools. In schools, it is a crucial matter to cultivate self-esteem. When children show certain behavioral or psychological problems, school teachers often attribute the causes to lower self-esteem. However, in the programs to enhance self-esteem at schools, their successful results seem to enhance both desirable and undesirable self-esteem, because most of the good results were obtained using self-report questionnaires such as Rosenberg's scale that could represent both types of self-esteem. In addition, the contents of the programs do not include any consideration for the nonconscious functions of children's psychological characteristics. It is clear that past programs were not efficient in increasing autonomous self-esteem and decreasing heteronomous self-esteem.

Given these present conditions, Yamasaki and colleagues (Yamasaki, Uchida, Yokoshima, Kaya, & Michishita, 2018; Yokoshima, Kaya, Michishita, & Yamasaki, 2018) developed a new universal prevention program for children to cultivate autonomous self-esteem and decrease heteronomous self-esteem, based on a group of programs termed "TOP SELF (Trial Of Prevention School Education for Life and Friendship)" (e.g., Uchida, Yamasaki, & Sasaki, 2014; Yamasaki, Murakami, Yokoshima, & Uchida, 2015). Mainly according to the somatic-marker hypothesis (Damasio, 1994), the programs of the TOP SELF include a theoretical background that children learn desirable psychological characteristics while a great amount of emotions arise. Emotions represent somatic responses that are nonconsciously evoked (Damasio, 1994, 2003). The methods of the programs are enjoyable for children using games, music, and animated stories to evoke emotions. The new program for autonomous self-esteem adopted the theories and methods of TOP SELF and constructed the revised structure and contents of the program purposes representing the process and components to cultivate autonomous self-esteem (see Yamasaki, Uchida, Yokoshima, Kaya, & Michishita, 2018, for details). After completing the program, Yamasaki, Michishita, Yokoshima, Kaya, and Uchida (2018) examined the effectiveness of the program targeting children in elementary schools. The program was implemented in 45-minute regular classes weekly for four successive weeks. Results showed that autonomous self-esteem by the IAT significantly increased from the pre- to post-intervention periods in the intervention group, while it did not change in the control group. Regarding heteronomous self-esteem measured by a reliable and valid questionnaire (Kaya, Yamaguchi, Yokoshima, Uchida, & Yamasaki, 2018), it significantly decreased from the pre- to post-intervention periods in the intervention group, but did not change in the control group. These results clarified that the program led to enhancing healthy and adaptive self-esteem, and to decreasing unhealthy and nonadaptive self-esteem.

7. Future AvenuesRegarding Research and Education for Self-Esteem

The current article discussed the characteristics of healthy and adaptive self-esteem. It also introduced new concepts such as autonomous and heteronomous self-esteem, underscoring the importance of distinguishing between desirable and undesirable self-esteem. In line with this consideration, the following two points were proposed: (1) Rosenberg's (1965) scale should not be used because it includes elements of both desirable and undesirable self-esteem that cannot be distinguished, and (2) desirable self-esteem such as autonomous types needs to be nonconsciously measured using certain implicit tests, although undesirable self-esteem such as heteronomous types can be consciously measured using self-report questionnaires. Moreover, because this modification of concepts and measurement methods of self-esteem suggests a new type of program to cultivate desirable self-esteem in schools, a new school-based program for children was introduced, along with the demonstrated effectiveness in enhancing autonomous self-esteem and decreasing heteronomous self-esteem. Research on self-esteem has thus far been severely criticized in terms of concepts and measures. Although the criticism pointed out serious problems that cannot be overlooked, many researchers have been continuing to conduct studies on self-esteem using the usual concepts and measures, which resulted in inconsistent confounded findings in this domain.

To resolve this confounded situation, first, the concept of self-esteem needs to be unified or revised. However, so long as many researchers insist on their own definitions of self-esteem, it seems difficult to reach total agreement, because any decision cannot be empirically made regarding the psychological construct per se. It might be necessary for key researchers to meet and discuss to determine a unified concept.

In the current article, we recommended the concept of autonomous and heteronomous self-esteem with the new IATs and programs to measure and cultivate autonomous self-esteem. In Japan, in addition to writing scientific papers and books, and presenting papers at academic meetings, we have been opening symposiums, lectures, and workshops to both researchers and school teachers to explain about these new self-esteem principles. Nevertheless, the dissemination is now far from progressing well. Psychology is somewhat of an ambiguous science, unlike natural sciences. It seems no so easy to conclude what is right and what is wrong using empirical methods, when many different notions continue to coexist. This paper emphasized various necessities to accelerate scientific activities regarding desirable true self-esteem by drastically improving concepts, measurement, and programs.

Acknowledgments

The author thanks Kanako Uchida, Takayuki Yokoshima, and Ikuko Kaya for their contributions.

Funding

This paper was supported by JSPS KAKENHI Grant Number 18K03140.

References

Ackerman, S. J., Hilsenroth, M. J., Clemence, A. J., Weatherill, R., & Fowler, J. C. (2001). Convergent validity of Rorschach and TAT scales of object relations. *Journal of Personality Assessment*, *77*, 295-306.doi:10.1207/S15327752JPA7702_11

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightman (Eds.), *Measures of personality and social psychological attitudes* (3rd ed.) (pp. 115-160). Ann Arbor, MI: Institute of Social Research.

Bosson, J. K., Brown, R. P., Zeigler-Hill, V., & Swann, Jr., W. B. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity*, *2*, 169-187. doi:10.1080/15298860309029

Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631-643. doi:10.1037/0022-3514.79.4.631

Coopersmith, S. (1967). *The antecedents of self-esteem.* San Francisco: W. H. Friedman and Company.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Putnam.

Damasio, A. R. (2003). *Looking for Spinoza: Joy, sorrow and the feeling brain.* New York: Harcourt.

Deci, E. L., & Ryan, R. M. (1995). Human autonomy: The basis for true self-esteem. In Kernis, M. H. (Ed.), *Efficacy, agency, and self-esteem* (pp. 31-49). New York: Plenum Press.

Demo, D. H. (1985). The measurement of self-esteem: Refining our methods. *Journal of Personality and Social Psychology*, *48*, 1940-1502. doi:10.1037/0022-3514.48.6.1490

Donnellan, M. B., Trzesniewski, K. H., & Robins, R. W. (2015). Measures of self-esteem. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measure of personality and social psychological constructs* (pp. 131-157). New York: Elsevier.

Falk, C. F., Heine, S. J., Takemura, K., Zhang, C. X. J., & Hsu, C.-W. (2015). Are implicit self-esteem measures valid for assessing individual and cultural differences? *Journal of Personality*, *83*, 56-68. doi:10.1111/jopy.12082

Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem. II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, *46*, 404-421. doi:10.1037/0022-3514.46.2.404

Garvie, C. T. (1979). Focus of motivation in sport: A principle factor. Unpublished dissertation, University of Alberta, Edmonton. Cited from Robinson, T. T., & Carron, A. V. (1982). Personal and situational factors associated with dropping out versus maintaining participation in competitive sport. *Journal of Sport Psychology*, *4*, 364-378.

Gnambs, T., Scharl, A., & Schroeders, U. (2018). The structure of the Rosenberg Self-Esteem Scale: A cross-cultural meta-analysis. *Zeitschrift für Psychologie*, *226*, 14-29. doi:10.1027/2151-2604/a000317

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27. doi:10.1037/0033-295X.102.1.4

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022-1038. doi:10.1037/0022-3514.79.6.1022

Griffiths, R. A., Beumont, P. J. V., Giannakopoulos, E., Russell, J., Schotte, D., Thornton, C., ...Varano, P. (1999). Measuring self-esteem in dieting disordered patients: The validity of the Rosenberg and Coopersmith contrasted. *International Journal of Eating Disorders*, *25*, 227-231. doi:10.1002/(SICI)1098-108X(199903)25:2<227::AID-EAT13>3.0.CO;2-4

Hagborg, W. J. (1993). The Rosenberg Self-Esteem Scale and Harter's Self-Perception Profile for Adolescents: A concurrent validity study. *Psychology in the Schools*, *30*, 132-136. doi:10.1002/1520-6807(199304)30:2<132::AID-PITS2310300205>3.0.CO;2-Z

Harter, S. (1985). *The Self-Perception Profile for Children: Revision of the Perceived Competence Scale for Children*. Unpublished manuscript, University of Denver, Denver, Co. Cited from Harter, S. (2012). *Self-Perception Profile for Adolescents: Manual and questionnaires*. University of Denver, Denver, Co.

Izuma, K., Kennedy, K., Fitzjohn, A., Sedikides, C., & Shibata, K. (2018). Neural activity in the reward-related brain regions predicts implicit self-esteem: A novel validity test of psychological measures using neuroimaging. *Journal of Personality and Social Psychology*, *114*, 343-357. doi:10.1037/pspa0000114

Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, *85*, 969-978. doi:10.1037/0022-3514.85.5.969

Kaplan, H. B., & Pokorny, A. D. (1969). Self-derogation and psychosocial adjustment. *Journal of Nervous and Mental Disease*, *149*, 421-434. doi:10.1097/00005053-196911000-00006

Karpinski, A. (2004). Measuring self-esteem using the implicit association test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22-34. doi:10.1177/0146167203258835

Kaya, I., Yamaguchi, S., Yokoshima, T., Uchida, K., & Yamasaki, K. (2018). Development of the heteronomous (contingent) self-esteem scale for children (HSES-C): Examination of the reliability and validity of the scale considering the applicability to school education. *Journal for the Science of Schooling*, *19*, 1-12.

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*, 1-26. doi:10.1207/S15327965PLI1401_01

Quirin, M., Kazén, M., Rohrmann, S., & Kuhl, J. (2009). Implicit but not explicit affectivity predicts circadian and reactive cortisol: Using the Implicit Positive and Negative Affect Test. *Journal of Personality*, *77*, 401-425. doi:10.1111/j.1467-6494.2008.00552.x

Richardson, C. G., Ratner, P. A., & Zumbo, B. D. (2009). Further support for multidimensionality within the Rosenberg Self-Esteem Scale. *Current Psychology*, *28*, 98-114. doi:10.1007/s12144-009-9052-3

Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality correlates of self-esteem. *Journal of Research in Personality*, *35*, 463-482. doi:10.1006/jrpe.2001.2324

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68-78. doi:10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Ed.), Handbook of self-determination research. Rochester, NY: University of Rochester Press.

Sinclair, S. J., Blais, M. A., Gansler, D. A., Sandberg, E., Bistis, K., & LoCicero, A. (2010). Psychometric properties of the Rosenberg Self-Esteem Scale: Overall and across demographic groups living within the United States. *Evaluation & the Health Professions*, *33*, 56-80. doi:10.1177/0163278709356187

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences*, *31*, 653-673. doi:10.1016/S0191-8869(00)00169-0

Uchida, K., Yamasaki, K., & Sasaki, M. (2014). Attractive, regularly-implementable universal prevention education program for health and adjustment in schools: An innovation from Japan. *Procedia – Social and Behavioral Sciences*, *116*, 754-764.

Watson, D., Suls, J., & Haig, J. (2002). Global self-esteem in relation to structural models of personality and affectivity. *Journal of Personality and Social Psychology*, *83*, 185-197. doi:10.1037/0022-3514.83.1.185

Wegener, I., Geiser, F., Alfter, S., Mierke, J., Imbierowicz, K., Kleiman, A., ...Conrad, R. (2015). Changes of explicitly and implicitly measured self-esteem in the treatment of major depression: Evidence for implicit self-esteem compensation. *Comprehensive Psychiatry*, *58*, 57-67. doi:10.1016/j.comppsych.2014.12.001

Yamasaki, K., Michishita, N., Yokoshima, T., Kaya, I., & Uchida, K. (2018). *Effectiveness of a school-based universal prevention program for enhancing autonomous self-esteem at elementary schools: Utilizing a newly developed implicit association test and questionnaire.* Paper presented at the 5th annual meeting of the International Academic Conference on Social Sciences, Sydney, Australia.

Yamasaki, K., Murakami, Y., Yokoshima, T., & Uchida, K. (2015). Effectiveness of a school-based universal prevention program for enhancing self-confidence: Considering the extended effects associated with achievement of the direct purposes of the program.*International Journal of Applied Psychology,5*, 152-159. doi:10.5923/j.ijap.20150506.02

Yamasaki, K., Uchida, K., Yokoshima, T., Kaya, I., & Michishita, N. (2018). Development of the purposes of a universal prevention program for autonomous self-esteem and a plan for developing educational methods. *Bulletin of Center for Collaboration in Community (Naruto University of Education)*, *32*, 91-100.

Yamasaki, K., Yokoshima, T., & Kanko, U. (2017). Reconstruction of the concept and measuring method of self-esteem: Toward innovatively reborn research of self-esteem. *Research Bulletin of Naruto University of Education*, *32*, 1-19.

Yokoshima, T., Kaya, I., Uchida, K., & Yamasaki, K. (2018). Effects of a school-based universal prevention program "Development of Self-Confidence": Examining the effectiveness of the program Utilizing the paper and pencil version of Self-Esteem Implicit Association Test for Children. *Japanese Journal of School Health*, *60*, 5-17.

Yokoshima, T., Ogami, Y., Kaya, I., & Yamasaki, K. (2019). *Development of the tablet PC version of Self-Esteem Implicit Association Test (T-SE-IAT) for children.* Manuscript submitted for publication.

Yokoshima, T., Uchiyama, Y., Uchida, K., & Yamasaki, K. (2017). Development of the paper and pencil version of Self-Esteem Implicit Association Test for Children (SE-IAT-C): Investigation of the reliability and validity utilizing Rosenberg's Self-EsteemScale and assessment of children by teachers. *Journal for the Science of Schooling*, *18*, 1-13.

Yokoshima T., Uchiyama., Y., Uchida., K., & Yamasaki, K. (2019). *What does the Rosenberg Self-Esteem Scale for Children (RSES-C) measure? Focusing on the adaptive and nonadaptive components of self-esteem*. Manuscript submitted for publication.

Zeigler-Hill, V. (2006). Discrepancies between implicit and explicit self-esteem: Implications for narcissism and self-esteem instability. *Journal of Personality*, *74*, 119-143. doi:10.1080/00224545.2012.688778