



# Investigating Stable and Dynamic Aspects of the Vietnamese Self-Compassion Scale using Generalisability Theory

Quoc Cuong Truong<sup>1</sup> · Thanh Minh Nguyen<sup>2</sup> · An Hong Nguyen<sup>2</sup> · Angela Medvedeva<sup>3</sup> · Oleg N. Medvedev<sup>1</sup>

Accepted: 12 July 2022 / Published online: 20 July 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

**Objectives** Evaluating comprehensive reliability of the Vietnamese Self-Compassion Scale (VSCS) and its ability to distinguish between trait (stable) vs state (dynamic) aspects of self-compassion using Generalisability Theory (G-Theory) is necessary. This investigation contributes to both reliability and validity of research that uses the VSCS to measure self-compassion in Vietnamese adults.

**Methods** In a sample of 155 Vietnamese adults who completed the VSCS at three occasions that were each 2 weeks apart, a G-study was conducted to measure reliability and trait vs state aspects of each VSCS subscale and the short-form VSCS, and a D-study was conducted to examine the effects of removing subscales on overall scale reliability as well as evaluate trait vs state aspects of each item.

**Results** With G-coefficients of 0.93–0.98, both the complete and short-form VSCS (VSCS-SF) demonstrated excellent reliability in measuring trait self-compassion. Three of the six subscales—self-judgement, mindfulness, and kindness—also demonstrated excellent reliability, with G-coefficients of 0.82–0.85. Eighteen of the 26 items measured trait more than state. The remaining eight items reflected a mixture of trait and state, but this did not affect overall reliability.

**Conclusions** This study indicated that the VSCS, VSCS-SF, and three VSCS subscales reliably measured trait self-compassion, with scores generalisable across the Vietnamese population and occasions. Thus, overall self-compassion levels remained stable over time, which is useful for evaluating the effectiveness of an intervention because significant changes of self-compassion are likely to be long-lasting.

**Keywords** Self-Compassion · State and trait · Measurement · Vietnamese Self-Compassion Scale · Generalisability Theory

The recent development of a Vietnamese version of the widely used Self-Compassion Scale (SCS; Neff, 2003a) has greatly enhanced the study of self-compassion in Vietnam. The Vietnamese Self-Compassion Scale (VSCS; Tran & Tran, 2017) is the most widely used measure of self-compassion in Vietnamese adults and has showed strong psychometric properties according to Classical Test Theory (CTT) metrics such as Cronbach's alpha (Nguyen & Le, 2021). There is increasing research attention to the benefits

of mindfulness-based practice such as exercising self-compassion, including improving human health and well-being. Accordingly, the SCS has been cross-culturally validated and widely applied, including in Vietnam. A thorough evaluation of the Vietnamese SCS is particularly important given that Buddhism, of which mindfulness-based practice is a central part, has a historically strong influence and prevalence in Vietnam. However, the VSCS has not been examined in terms of comprehensive generalisability and reliability properties using Generalisability Theory (G-Theory) developed by Cronbach et al. (1963), which can precisely distinguish between dynamic and stable aspects of a measure, i.e., if self-compassion is a trait and in turn, is stable over time. Thus, using G-theory, it is important to evaluate the applicability of the VSCS to basic and intervention-based research and identify limitations of the VSCS that should be addressed.

✉ Thanh Minh Nguyen  
nguyenminhthanhpsy@gmail.com

<sup>1</sup> School of Psychology, University of Waikato, Hamilton, New Zealand

<sup>2</sup> Department of Psychology, Hoa Sen University, District 1, Ho Chi Minh City 71000, Vietnam

<sup>3</sup> Department of Chemistry, Rice University, Houston, TX, USA

In the construction of the SCS, self-compassion was defined as being open-minded; connecting to one's own suffering; experiencing feelings of kindness to oneself; and taking an understanding and non-judgmental perspective confronting one's shortcomings and failures, with the awareness that these experiences are part of the common human experience (Neff, 2003a). The SCS consists of 26 items divided into six subscales: Self-Kindness, Self-Judgment, Common Humanity, Isolation, Mindfulness, and Over-Identification (Neff, 2003a). In the initial validation with 391 university students, the SCS demonstrated excellent overall consistency, Cronbach's  $\alpha=0.92$ , with fair internal consistency for individual subscales, Cronbach's  $\alpha=0.75$  to 0.81, lowest for Mindfulness and highest for Over-Identification (Neff, 2003a). The original six-factor structure of SCS has been supported by subsequent validations in China (Chen et al., 2011), Korea (Lee & Lee, 2010), Italy (Petrocchi et al., 2013), and Brazil (de Souza & Hutz, 2016). A 12-item short-form of the SCS (SCS-SF) was validated in Dutch and American undergraduates, and the English version demonstrated a significantly high correlation with the full form in scores ( $r=0.98$ ;  $r=0.89$  to 0.93 for subscale score correlations; Raes et al., 2011). However, while the results of the validation supported the use of total score, Cronbach's  $\alpha=0.80$ , Cronbach's alpha ranged from 0.54 to 0.75 for subscale scores, suggesting that subscale scores should not be used (Raes et al., 2011). The SCS was translated and validated into Vietnamese (Nguyen & Le, 2021; Tran & Tran, 2017). It was translated and back-translated by English-Vietnamese bilingual researchers who were proficient in English and Vietnamese (Tran & Tran, 2017). Following that, pilot testing of wording was performed with five university students (Tran & Tran, 2017). The Vietnamese SCS demonstrated satisfactory reliability, Cronbach's  $\alpha=0.76$  to 0.92. The VSCS was significantly associated with higher psychological well-being (Nguyen & Le, 2021) and gratitude (Nguyen et al., 2020) and lower COVID-19-related distress (Nguyen & Le, 2021), narcissism, and anxiety (Quang et al., 2021).

Self-compassion may positively affect individuals' mental health (Wilson et al., 2019). Higher levels of self-compassion were associated with psychological well-being (Nguyen & Le, 2021) and may reduce mental health symptoms such as depression, stress, and anxiety in both adults and adolescents (MacBeth & Gumley, 2012; Marsh et al., 2018). Compassion-based therapies were implemented to further investigate the relationship between compassion and mental health (Leaviss & Uttley, 2015). Self-compassion-based psychotherapy contributed to improving self-compassion and other mental health outcomes (see Kirby et al., 2017 for a meta-analysis). Self-compassion may also improve interpersonal relationships (Yarnell & Neff, 2013). Specifically, higher levels of compassion are associated with a higher

ability to compromise with others, lower levels of emotional disturbance, and higher levels of authenticity and relationship happiness (Yarnell & Neff, 2013).

Self-compassion is positively associated with secure attachment, positive parenting behaviours; healthier relationships with a romantic partner, family, and friends; and more constructive conflict resolutions (see Lathren et al., 2021 for a review). In Vietnamese adults, parental self-compassion is associated with positive parenting practices such as an authoritative parenting style or mindful parenting (Nguyen et al., 2020). However, further research is needed to investigate self-compassion in Vietnamese adults. Vietnam, like other countries with strong Buddhist influences, provides an important socio-cultural context for studying mindfulness because many non-Buddhists may already practice mindfulness or be aware of mindfulness practices. In the third study of the initial validation of the SCS, there were significantly higher self-compassion scores for Buddhist practitioners ( $n=43$ ) of the mindfulness-based practice Vipassana than undergraduates ( $n=232$ ), and specifically scores were higher for Self-Kindness, Common Humanity, and Mindfulness and lower for Self-Judgement, Over-Identification, and Isolation, and longer time of practice in years was associated with higher self-compassion scores (Neff, 2003a).

However, these results have not been replicated across all Buddhist samples, indicating that further research is needed to understand the relationship between Buddhist practices and self-compassion (Zeng et al., 2016). The concept of self-compassion as defined by the SCS may differ from that of Buddhism, or relationships between SCS-measured self-compassion and Buddhism may vary by practice (e.g., mindfulness-based Vipassana was practiced in the Neff's (2003a) sample) and be stronger in the case of mindfulness-based Buddhist practice. Measures like the VSCS may help to indirectly measure the continued influence of mindfulness-based Buddhist practices in Vietnam in the midst of a decline in Buddhist followers, which is important for designing mindfulness-based interventions to secular populations with existing knowledge of Buddhism. Thus, it is important to conduct longitudinal psychometric studies of the VSCS with robust methods such as G-theory that evaluate how the scale measures state vs trait self-compassion. In turn, the reliability of the VSCS can be strengthened, and the VSCS can be better applied to basic and intervention-based research contexts.

Individual self-compassion was widely referred to as a stable aspect or trait (Breines & Chen, 2013; Medvedev et al., 2021b; Neff, 2003b; Neff et al., 2017). A trait was defined as a stable individual reaction pattern spanning different periods and varied locations, while a state was an individual's response to an occasion that happened at a particular time (Epstein, 1984; Spielberger et al., 1970). The precision of clinical evaluations and the validity of studies

were highly dependent on whether the measurements could distinguish between state and trait aspects: states could be subject to variations in conditions while traits remained stable over time, so it was important to have valid measures of each and distinguish between them (Medvedev et al., 2017; Truong et al., 2020). A treatment or intervention could aim to change trait aspects because impacting trait was associated with a long-term change, and relapse may be less likely to occur (Medvedev et al., 2017). Changes in state aspects indicate that a specific condition such as training or practice may play a role, although these changes may be short-lasting without continued practice and susceptible to other influences such as an individual's mood or physical state.

Recently, self-compassion was also proposed as a state with the same six components as a trait (Neff et al., 2021) suggesting that self-compassion includes both trait and state aspects but the degree to which the overall construct and its facets reflect a trait or a state remains unknown and more research is needed to elucidate this quantitatively. Studies have suggested that when an individual is in difficult situations or being threatened, their state aspects of self-compassion may be more present than the trait self-compassion (Flentje, 2021; Thøgersen-Ntoumani et al., 2017). It means that self-compassion may rapidly change and hence increasing common humanity in such situations (Waring & Kelly, 2019). This may help individuals to become more compassionate towards themselves and motivate their engagement in more adaptive behaviours (Leary et al., 2007; Waring & Kelly, 2019). However, research found that trait self-compassion also increased more adaptive responses to failure (e.g., reduce shame), and positively predicted state self-compassion (Waring & Kelly, 2019). Therefore, measures of both state and trait aspects of self-compassion are necessary, and it is important to identify how well the measure assesses and distinguishes each, in basic and intervention-based research settings (Cahn & Polich, 2006; Chiesa and Serretti, 2009). Even though the SCS was well-validated and increasingly applied for evaluating an individual's self-compassion levels, research conducted to evaluate and distinguish between dynamic and stable aspects of self-compassion using comprehensive methodology (e.g., G-theory) is still limited (Muris et al., 2018; Neff et al., 2021).

The Classical Test Theory (CTT) method traditionally obtains Cronbach alpha and intraclass correlation coefficients (ICC) for the internal consistency and temporal reliability of a measure, respectively. In addition, CTT also computes test–retest coefficients to distinguish between state and trait aspects of a measure; coefficients of 0.70 and above reflect a measure of trait, and coefficients below 0.60 reflect a measure of state (Ramanaiah et al., 1983; Spielberger, 1999). However, the accuracy of CTT has been questioned because measurement error is considered a single component, while in complex natural environments there are multiple sources

of error that can impact the precision of measurement, for example time and respondent state and properties of the environment (Bloch & Norman, 2012; Medvedev et al., 2017). Particularly, both the test–retest coefficient and the ICC only represent the correlations of scale scores at different occasions (e.g., occasion 1 and occasion 2) and do not account for measurement error due to individual items of the scale and/or interactions between person (participant), scale items, and occasions (Bloch & Norman, 2012).

G-Theory is a comprehensive statistical method for evaluating the reliability of psychometric instruments and investigating their ability to distinguish state from trait (Arterberry et al., 2014; Cronbach et al., 1963; Medvedev et al., 2017; Truong et al., 2020, 2022). G-Theory is an extension of CTT and can evaluate possible sources of measurement error that may influence the reliability of an instrument (Allen & Yen, 1979; Cronbach et al., 1963). Specifically, G-theory can account for effects of all relevant factors including person, methodology (individual items) and situation (different occasions) that may contribute to measurement error independently or via interactions between factors. In summary, while CTT only accounts for one aspect of reliability at a time (e.g., test–retest coefficient, ICC, Cronbach's alpha), G-theory closely evaluates all possible aspects influencing reliability simultaneously, including interactional effects between factors, which is useful for improving the accuracy of a psychometric instrument. As such, previous studies demonstrated that G-theory was the most suitable application to investigate the overall reliability and distinguish between state and trait aspects of a measure (Arterberry et al., 2014; Medvedev et al., 2017; Paterson et al., 2018; Truong et al., 2021).

To date, only the English SCS-SF was examined using G-theory, and the results indicated that the scale predominantly measured the trait aspect of self-compassion (Medvedev et al., 2021b). However, the full 26-item English and Vietnamese SCS scales and the Vietnamese short-form scale were not subjected to robust G-theory examination. This study aimed to apply G-theory to the VSCS, VSCS-SF, and subscales to (1) investigate trait and state aspects of self-compassion, (2) evaluate the temporal reliability, and (3) identify potential sources of error that may affect the measurement. Specifically, a generalisability study (G-study) evaluated reliability and generalisability of the total scores of the VSCS, the VSCS-SF, and the scores of individual VSCS subscales and identified possible sources of measurement error affecting these measures (Cardinet et al., 2010). Next, a decision study (D-study) was carried out to assess the psychometric properties of individual VSCS items and examine how removing each subscale/item can affect the overall reliability of the VSCS. Results from D-study could be used to identify items that measure state or trait or both aspects of self-compassion to a comparable degree as well as

to optimise the measurement design of the scale to enhance reliability by experimenting with removing items/subscales levels (Truong et al., 2020).

## Method

### Participants

Four hundred and eighteen participants were included at the baseline sample (occasion 1). However, the sample decreased to 237 and 155 participants at occasions 2 and 3 respectively. We excluded participants who did not complete the VSCS at followed-up occasions. Therefore, the sample selected for the generalisability analyses included 155 Vietnamese participants aged 18–46 ( $M$  age = 27.05,  $SD$  = 6.46;  $n$  = 20 or 12.90% male). According to a power analysis for a repeated-measures ANOVA over three occasions to detect an effect size of 0.25 with  $\beta$  = 0.95 and  $p$  < 0.05, a sample of 43 was required, so the sample size far exceeded the minimum requirement (Faul et al., 2007; Forrest et al., 2021; Medvedev et al., 2021a).

### Procedures

To recruit the participants, the advertisement of the study was publicised online through many social networks such as Facebook, Twitter and university websites. There was a hyperlink and QR code in the advertisement led potential participants to the information about the study. Participants completed the VSCS online at three occasions with 2-week intervals. Respondents were also required to provide demographic information such as sex, age, educational level, and living area. To ensure the completed forms matching for each participant, they were asked to provide their email address at the first occasion. Every 2 weeks after, participants received an email to remind them to complete the instrument again.

### Measures

The VSCS (Tran & Tran, 2017) consisted of 26 items measuring six subscales including (1) Self-Kindness (e.g., “I try to be loving towards myself when I’m feeling emotional pain”), (2) Self-Judgment (e.g., “I’m disapproving and judgmental about my own flaws and inadequacies”), (3) Common Humanity (e.g., “When things are going badly for me, I see the difficulties as part of life that everyone goes through”), (4) Isolation (e.g., “When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world”), (5) Mindfulness (e.g., “When something upsets me I try to keep my emotions in balance”), and (6) Over-Identification (e.g., “When I’m feeling down I

tend to obsess and fixate on everything that’s wrong”) (Tran & Tran, 2017). A 5-point Likert scale from 1 (almost never) to 5 (almost always) was used for respondents to rate each individual item. There were 9 items which located in the subscales of Self-Judgment (i.e., items: 1, 8, 11, 16, and 21) and Over-Identification (i.e., items: 2, 6, 20, and 24) that required reverse coding before conducting data analysis. The total scores of the scale and individual subscale scores were represented by the mean of responses to the relevant items together after conducting reverse coding. The VSCS-SF had 12 items and was grouped according to the same format as the English SCS-SF (Raes et al., 2011). In the current study, the VSCS-SF scores were extracted from the full VSCS scale for each participant.

### Data Analyses

Descriptive statistics including means ( $M$ ), standard deviation ( $SD$ ), Cronbach’s alpha, McDonald’s Omega coefficients, test–retest coefficients, and intraclass correlation coefficients (ICC) for the VSCS, the VSCS-SF, and individual subscales of the VSCS were computed by using IBM SPSS Statistics 27 software. Swiss Society for Research in Education’s EduG 6.1-e software (Swiss Society for Research in Education Working Group, 2006) was employed to conduct Generalisability analyses (G-analyses). All G-analyses involved in both the G- and D-studies were performed following the guidelines described elsewhere (Cardinet et al., 2010; Truong et al., 2020). This study utilised a two-facet design incorporating a random effects model: person (P) by item (I) by occasion (O), express as  $P \times I \times O$ . The object of measurement in this G-theory study was the facet P (person) which also called the differentiation facet and was not a source of measurement error (Brennan, 2011; Cardinet et al., 2010). The other two facets (i.e.,  $I$  = item and  $O$  = occasion) were instrumentation facets. The facets P and O were set infinite and the I facet was fixed because the same VSCS questionnaires were used across all occasions. Supplementary Table 1 presents formulae using to compute G-Theory estimates and component variances for the two-facet design of person by item by occasion, express as  $P \times I \times O$  for this study design (Shavelson et al., 1989). There are two obtained reliability/generalisability coefficients (G-coefficients): the relative G-coefficient ( $G_r$ ) and the absolute G-coefficient ( $G_a$ ), which are reflecting the reliability of a trait measure.  $G_r$  of 0.80 or higher and  $G_a$  of 0.70 or higher indicate a measure reliably assessing trait if person is the object of measurement (Arterberry et al., 2014; Truong et al., 2020). Additionally, a G-coefficient is computed by a proportion of person (P) variance and all error variance that consist of person variance and total error variance due to other facets and their interactions (i.e., I,

O,  $I \times O$ ,  $P \times I$ ,  $P \times O$ , and  $P \times I \times O$ ) (Forrest et al., 2021), expressed as:

$$G = \frac{\text{Person variance}}{\text{Person variance} + \text{Total error variance}}$$

Thus, it can be interpreted, for instance, if a G-coefficient equals 0.80, it reflects that person variance explains 80% of all error variance and the total error variance accounts for the remaining 20% of all error variance in the data.

Trait component indices (TCI) and state component indices (SCI), which represent the proportion of measurement error attributed to trait and state aspects of a measure were also computed in both G- and D-study by following the guidelines described in Medvedev et al., (2017). In G-study, these coefficients were used to indicate whether a measure was measuring trait or state aspects. Specifically, an SCI of less than 0.40 ( $TCI \geq 0.60$ ) reflects a measure of trait, and an SCI of greater than or equal to 0.60 ( $TCI < 0.40$ ) indicates a reliable state measure. TCI and SCI ranging between these benchmarks (i.e.,  $0.40 \leq SCI < 0.60$ ) suggest that comparable amounts of variance are attributed to both state and trait aspects of a measure (Truong et al., 2020). In this case, this means that an instrument is measuring both state and trait to a comparable degree. In D-study, TCI and SCI were used to identify variation attributable to dynamic or stable aspects. Individual items with an SCI of less than 0.40 ( $TCI \geq 0.60$ ) reflect stable/trait characteristics, while items with a SCI greater than or equal to 0.60 are indicative of psychological states (Medvedev et al., 2017).

## Results

### Descriptive Statistics

Descriptive statistics for the total scores and six subscale scores of the VSCS across three occasions are displayed in Table 1. The VSCS achieved excellent internal consistency with both Cronbach's alpha and McDonald's Omega of 0.94 at occasions 1 and 2, and 0.95 at occasion 3 (Cronbach, 1951; McDonald, 2013). The VSCS also achieved excellent ICC of 0.97 (95%CI [0.96, 0.98]), and excellent test–retest reliability coefficients of 0.96 between occasion 1 and occasion 2, and 0.95 between occasion 1 and occasion 3. Slightly lower Cronbach's alpha and McDonald's Omega obtained for the VSCS-SF across occasions but their estimates are still in the range of good to excellent internal consistency ( $\geq 0.88$ ). The VSCS-SF also achieved excellent ICC of 0.97 (95%CI [0.95, 0.97]), and excellent test–retest reliability coefficients of 0.95 and 0.94 for occasion 2 and occasion 3, respectively (with reference to occasion 1).

The subscales of the VSCS had fair to good internal consistency with Cronbach's alpha ranging from

0.75 to 0.90 and with McDonald's Omega ranging from 0.75 to 0.89. The ICC values of all VSCS subscales are above 0.88. Overall, traditional statistical approaches indicate that the VSCS and its six subscales showed high temporal reliability and fair to excellent internal consistency, which reflected trait measures.

### G-Study

Table 2 displays error variance due to person (P), item (I), and occasion (O), and their interactions ( $P \times I$ ,  $P \times O$ ,  $I \times O$ ,  $P \times I \times O$ ) together with G-coefficients (i.e.,  $G_r$  and  $G_a$ ), as well as trait and state component indices (TCI, SCI) for the VSCS and its individual subscales. Both  $G_r$  and  $G_a$  of the VSCS were 0.98 which indicated excellent reliability and generalisability of VSCS scores across persons and occasions and with the remaining 2% of error variance entirely (100%) explained by the interaction between person and occasion ( $P \times O$ ). Interestingly, the TCI value was also 0.98, reflecting that the VSCS was reliably measuring stable aspects of self-compassion. In addition, reliability of the VSCS was not influenced by other sources of error variance. Moreover, slightly lower but still excellent  $G_r$  and  $G_a$  values of 0.94 and 0.93, respectively, were observed for the VSCS-SF, with only 6% of error variance mainly explained by  $P \times O$  and  $P \times O \times I$  interactions. The TCI value of 0.83 was reflecting the ability of the VSCS-SF to reliably assess self-compassion trait.

The score of the full VSCS has the highest reliability/generalisability following by the short version VSCS-SF, the subscales of self-judgement (0.85), mindfulness (0.83), and kindness (0.82) while reliability of the subscales of humanity (0.77), isolation (0.79), and over-identification (0.75) subscales falls slightly below the cut-off point of 0.80 (Arterberry et al., 2014). Most individual VSCS subscales' errors were explained by interactional effects between study facets (i.e.,  $P \times O$ ,  $P \times I$ ,  $I \times O$ , and  $P \times I \times O$ ). Over-identification subscale was the only subscale affected by occasion error (6.2%) suggesting that the overall sample scores on this subscale were influenced by occasion increasing the overall error variance. The TCI values were also computed for all six subscales of the VSCS, and they all were above 0.94, clearly indicating that all VSCS subscales were measuring self-compassion trait. Overall, these estimates (i.e., TCI and G-coefficients) demonstrated that the VSCS and its short form VSCS-SF are valid and reliable measures of trait self-compassion, while only three subscales (i.e., self-judgement, mindfulness and kindness) met the requirements for reliable trait measures.

**Table 1** Means, standard deviation (*SD*), Cronbach's alpha, McDonald's Omega coefficients, test–retest coefficients, and intraclass correlation coefficient (ICC) for the VSCS total and its six subscales ( $n = 155 \times 3$  occasions)

Scale/assessment	Occasion 1	Occasion 2	Occasion 3	ICC (95%CI)
<b>Self-Kindness</b>				0.91 (0.88–0.93)
Mean ( <i>SD</i> )	3.61 (0.78)	3.52 (0.82)	3.53 (0.85)	
Cronbach's alpha/McDonald's Omega	0.82/0.81	0.87/0.87	0.90/0.89	
Test–retest ( $r^a$ )	–	0.85	0.85	
<b>Self-Judgment</b>				0.93 (0.91–0.95)
Mean ( <i>SD</i> )	2.85 (0.82)	2.84 (0.79)	2.77 (0.86)	
Cronbach's alpha/McDonald's Omega	0.80/0.80	0.80/0.82	0.84/0.85	
Test–retest ( $r^a$ )	–	0.88	0.91	
<b>Common Humanity</b>				0.89 (0.86–0.92)
Mean ( <i>SD</i> )	3.62 (0.76)	3.47 (0.81)	3.48 (0.78)	
Cronbach's alpha/McDonald's Omega	0.77/0.78	0.81/0.81	0.79/0.79	
Test–retest ( $r^a$ )	–	0.84	0.83	
<b>Isolation</b>				0.92 (0.90–0.94)
Mean ( <i>SD</i> )	2.88 (0.95)	2.84 (0.92)	2.84 (0.88)	
Cronbach's alpha/McDonald's Omega	0.81/0.82	0.82/0.79	0.80/0.75	
Test–retest ( $r^a$ )	–	0.88	0.87	
<b>Mindfulness</b>				0.89 (0.86–0.92)
Mean ( <i>SD</i> )	3.71 (0.73)	3.60 (0.76)	3.64 (0.80)	
Cronbach's alpha/McDonald's Omega	0.75/0.75	0.83/0.84	0.83/0.87	
Test–retest ( $r^a$ )	–	0.80	0.85	
<b>Over-Identification</b>				0.94 (0.91–0.95)
Mean ( <i>SD</i> )	3.31 (0.91)	3.24 (0.94)	3.14 (0.91)	
Cronbach's alpha/McDonald's Omega	0.82/0.83	0.87/0.87	0.85/0.85	
Test–retest ( $r^a$ )	–	0.89	0.91	
<b>VSCS-SF</b>				0.97 (0.95–0.97)
Mean ( <i>SD</i> )	3.33 (0.70)	3.30 (0.73)	3.33 (0.74)	
Cronbach's alpha/McDonald's Omega	0.88/0.88	0.90/0.89	0.91/0.91	
Test–retest ( $r^a$ )	–	0.95	0.94	
<b>VSCS Total</b>				0.97 (0.96–0.98)
Mean ( <i>SD</i> )	3.32 (0.67)	3.28 (0.67)	3.32 (0.70)	
Cronbach's alpha/McDonald's Omega	0.94/0.94	0.94/0.94	0.95/0.95	
Test–retest ( $r^a$ )	–	0.96	0.95	

<sup>a</sup>Test-retest bivariate correlations between occasion 1 and subsequent occasions 2 and 3

## D-Study

Table 3 shows generalisability estimates for individual items of the VSCS including error variance attributed to the effects of person, occasion, and person-occasion interaction together with SCI. There were 18 VSCS individual items with SCIs below 0.40, reflecting measurement of *trait* self-compassion and indicating that these items were least sensitive to transitory changes. For the remaining 8 items, SCIs ranged from 0.41 to 0.58, indicating that the items measured both aspects to a comparable degree and thus could not be considered as measuring state more than trait or vice versa. Table 3 also includes a comparison of the VSCS SCIs with those from the English SCS-SF. Items 6, 2, 11, 1, and 26 were a mixture of state and trait for the SCS-SF (SCIs 0.43–0.55), but the corresponding items except for item 11

(0.43) were strongly trait (SCI < 0.40, SCIs 0.27–0.34) for the VSCS. Of the items that were strongly trait for the SCS-SF (12, 15, 13, 25, 9, and 14), only 13, 14, and 25 were also clearly trait for the VSCS. There were no items in the VSCS that met the cut-off criteria (SCI ≥ 0.60) for clearly state, but item 10 was clearly state (SCI = 0.60) for the SCS-SF. The item in the SCS-SF with the highest trait score (lowest SCI, SCI = 0.23) was a mixture of state and trait (SCI = 0.48) for the VSCS.

A series of generalisability analyses were carried out to examine whether removing one subscale at a time would influence reliability of the VSCS (Table 4). The first analysis involved with removing the first subscale of the VSCS (i.e., self-kindness). Five analyses were then subsequently conducted with removing one another subscale at a time. These analyses found that reliability of the VSCS remained

**Table 2** G-study estimates for the VSCS, the VSCS-SF, and six VSCS subscales including grand mean (GM), standard errors of the grand mean (SE), coefficient G relative (Gr), coefficient G absolute (Ga), trait component index (TCI), state component index (SCI), and error variance (in %) attributed for the two-facet design express as person (P) × occasion (O) × item (I) (n = 155)

Facets	VSCS total		VSCS-SF		Self-Kindness		Self-Judgment		Common Humanity		Isolation		Mindfulness		Over-Identification	
	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%
P	0.324	98.0	0.350	93.0	0.296	79.0	0.320	82.0	0.274	72.0	0.323	77.0	0.379	77.0	0.291	69.0
I	0.000	0.0	0.000	0.0	0.002	0.7	0.000	0.0	0.000	0.0	0.003	0.8	0.015	3.0	0.001	0.2
O	0.000	0.0	0.003	0.8	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.013	2.6	0.027	6.2
P × I	0.000	0.0	0.003	0.9	0.016	4.3	0.013	3.3	0.020	5.4	0.022	5.1	0.017	3.5	0.027	6.4
P × O	0.006	2.0	0.007	1.8	0.013	3.5	0.000	0.0	0.005	1.2	0.015	3.4	0.018	3.6	0.016	3.7
I × O	0.000	0.0	0.003	0.7	0.012	3.1	0.016	4.1	0.025	6.6	0.008	1.9	0.008	1.6	0.006	4.3
P × I × O	0.000	0.0	0.011	2.8	0.036	9.5	0.042	10.6	0.056	14.8	0.051	11.8	0.044	8.8	0.056	13.2
GM	3.309		3.295		3.392		3.296		3.315		3.387		3.302		3.143	
SE	0.001		0.079		0.119		0.128		0.158		0.107		0.188		0.182	
Gr	0.98		0.94		0.82		0.85		0.77		0.79		0.83		0.75	
Ga	0.98		0.93		0.79		0.82		0.72		0.77		0.77		0.69	
TCI	0.98		0.83		0.96		> 0.99		0.98		0.96		0.95		0.95	
SCI	0.02		0.17		0.04		< 0.01		0.02		0.04		0.05		0.05	

consistently high with all G-coefficients of 0.97–0.98, and all TCIs of 0.98 compared to the total VSCS scale. In addition, Table 4 also presents the results of three further generalisability analyses which involved removing items with different SCI benchmarks to test whether removing these items would impact reliability of the VSCS. These analyses resulted in G-coefficients and TCIs of the VSCS either remained unchanged with removing items with  $SCI \geq 0.50$  or slightly lower G-coefficients (0.96–0.97) with removing items with  $SCI \geq 0.70$  and  $\geq 0.60$  compared to the total VSCS scale. Together, these findings suggest that the VSCS is measuring trait self-compassion and its reliability is optimal with the current measurement design.

### Discussion

This research aimed to investigate trait and state aspects of self-compassion measured by the VSCS and its subscales and evaluate the temporal reliability and generalisability of VSCS assessment scores using G-Theory. The findings of this research showed that the total VSCS was reliably measuring trait self-compassion with G-coefficients and TCI of 0.98. These estimates indicated that only 2% of variance in the total VSCS scores were explained by error variance attributed to state aspects of self-compassion and thus the impact of measurement error was negligible (Shavelson et al., 1989). The findings also indicated that the VSCS-SF was reliable in measuring trait self-compassion with both G relative and absolute coefficients of 0.94 and 0.93, respectively, and TCI of 0.83. Additionally, three subscales of the

VSCS (self-judgement, mindfulness and kindness) achieved temporal reliability in measuring stable aspects of self-compassion with G-coefficients above 0.80 and TCIs above 0.94 (Arterberry et al., 2014; Medvedev et al., 2017). The scores of individual VSCS subscales were predominantly influenced by measurement error due to the interactions between the differentiation facet (P) with item (i.e., P × I), and with item by occasion (i.e., P × I × O), indicating that the total VSCS scores were more reliable in comparison to individual subscales. Overall, the findings of this research demonstrated that the scores of VSCS and VSCS-SF are generalisable across sample population and occasions, and the VSCS and VSCS-SF are adequate measures of trait self-compassion in the current measurement design. Three VSCS subscales naming self-judgement, mindfulness and kindness can be used for reliable assessment of trait self-compassion levels. Moreover, the results reflected that self-compassion levels are less sensitive to change over time, which is consistent to the previous study (Medvedev et al., 2021b). Accordingly, if significant changes in self-compassion levels take place due to an intervention, they are more likely to last for a long term.

A D-study was carried out to evaluate psychometric properties of individual VSCS items in measuring stable versus dynamic aspects of self-compassion. It is interesting that there were no individual VSCS items reflecting state more than trait self-compassion, but 18 out of 26 VSCS items measured self-compassion as a trait. The remaining 8 items were reflecting both state and trait aspects of self-compassion to a comparable degree. This could be explained by the fact that these items using the semantics that were relevant

**Table 3** Error variance attributed to person (P), occasion (O), and P×O interaction together with trait component index (TCI) for individual VSCS items

Items/subscales	P	O	P×O	SCI
<i>Self-Kindness</i>				
5. I try to be loving towards myself when I'm feeling emotional pain	0.452	0.013	0.198	0.30
<b>12. When I'm going through a very hard time, I give myself the caring and tenderness I need. (6)</b>	0.192	0.140	0.266	0.58 ( <b>0.38</b> )
19. I'm kind to myself when I'm experiencing suffering	0.374	0.053	0.220	0.37
23. I'm tolerant of my own flaws and inadequacies	0.395	0.046	0.186	0.32
<b>26. I try to be understanding and patient towards those aspects of my personality I don't like. (2)</b>	0.545	0.057	0.217	0.28 ( <b>0.55</b> )
<i>Self-Judgment</i>				
<b>1. I'm disapproving and judgmental about my own flaws and inadequacies. (11)</b>	0.477	0.130	0.233	0.33 ( <b>0.47</b> )
8. When times are really difficult, I tend to be tough on myself	0.379	0.028	0.235	0.38
<b>11. I'm intolerant and impatient towards those aspects of my personality I don't like. (12)</b>	0.347	0.062	0.273	0.44 ( <b>0.43</b> )
16. When I see aspects of myself that I don't like, I get down on myself	0.388	0.153	0.227	0.37
21. I can be a bit cold-hearted towards myself when I'm experiencing suffering	0.391	0.047	0.223	0.36
<i>Common Humanity</i>				
3. When things are going badly for me, I see the difficulties as part of life that everyone goes through	0.238	0.131	0.278	0.54
7. When I'm down, I remind myself that there are lots of other people in the world feeling like I am	0.509	0.160	0.213	0.30
<b>10. When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people. (10)</b>	0.431	0.117	0.235	0.35 ( <b>0.60</b> )
<b>15. I try to see my failings as part of the human condition. (5)</b>	0.286	0.005	0.266	0.48 ( <b>0.23</b> )
<i>Isolation</i>				
4. When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world	0.344	0.035	0.279	0.45
<b>13. When I'm feeling down, I tend to feel like most other people are probably happier than I am. (4)</b>	0.470	0.005	0.256	0.35 ( <b>0.38</b> )
18. When I'm really struggling, I tend to feel like other people must be having an easier time of it	0.448	0.058	0.229	0.34
<b>25. When I fail at something that's important to me, I tend to feel alone in my failure. (8)</b>	0.429	0.018	0.184	0.30 ( <b>0.37</b> )
<i>Mindfulness</i>				
<b>9. When something upsets me I try to keep my emotions in balance. (7)</b>	0.334	0.064	0.230	0.41 ( <b>0.36</b> )
<b>14. When something painful happens I try to take a balanced view of the situation. (3)</b>	0.509	0.070	0.164	0.24 ( <b>0.39</b> )
17. When I fail at something important to me I try to keep things in perspective	0.519	0.051	0.225	0.30
22. When I'm feeling down I try to approach my feelings with curiosity and openness	0.471	0.000	0.216	0.31
<i>Over-Identification</i>				
<b>2. When I'm feeling down I tend to obsess and fixate on everything that's wrong. (9)</b>	0.527	0.056	0.274	0.34 ( <b>0.53</b> )
<b>6. When I fail at something important to me I become consumed by feelings of inadequacy. (1)</b>	0.553	0.022	0.206	0.27 ( <b>0.43</b> )
20. When something upsets me I get carried away with my feelings	0.241	0.017	0.334	0.58
24. When something painful happens I tend to blow the incident out of proportion	0.339	0.112	0.236	0.41

Highlighted in bold are the items included in the English SCS-SF with the item number in brackets after the item, and the corresponding SCI for the English SCS-SF item is in brackets and bolded in the SCI column, reproduced with permission from Medvedev et al., 2021b Table 6

to specific transient conditions that may differ significantly across individuals and over time (e.g., going through a very hard time, intolerant and impatient towards those aspects of my personality, failings as part of the human condition), which may lead to higher variability of the state aspect of self-compassion. Even so, these findings indicated that individual VSCS items are measuring predominantly trait self-compassion. Moreover, the D-study also involved analyses with removing one subscale of the VSCS at a time, and removing items with different TCI benchmarks to examine how this may impact on reliability of the VSCS. However, these analyses resulted in merely negligible changes of the

VSCS reliability (G-coefficients ranged from 0.96 to 0.98) compared to the analysis with the original VSCS items.

It is noteworthy that the results for the VSCS-SF are consistent to the previous study conducted by Medvedev et al., (2021b) which validated the SCS-SF using G-Theory methodology. The G-study outcomes from Medvedev et al.'s (2021b) study also indicated that the SCS-SF was reliably measuring trait self-compassion with G-coefficients of 0.84. More interestingly, the D-studies from both investigations revealed that none of individual items in the short-forms of the VSCS was sensitive to change over time because all obtained TCIs for individual items were



**Table 4** D-study reliability estimates and variance components for the person (P)×occasion (O)×item (I) design, including interactions when removing one subscale or items with different SCI benchmarks at a time

Facets	Self-Kindness		Self-Judgment		Common Humanity		Isolation		Mindfulness		Over-Identification		Items with SCI > 0.50		Items with SCI > 0.40		Items with SCI < 0.30	
	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	$\sigma^2$	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%	$\sigma^2$	%
P	0.333	97.0	0.323	97.0	0.330	97.0	0.318	97.0	0.312	97.0	0.326	97.0	0.342	98.0	0.348	96.0	0.311	96.0
I	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0	0.000	0.0
O	0.000	0.0	0.000	0.0	0.000	0.1	0.000	0.0	0.000	0.1	0.001	0.2	0.000	0.0	0.000	0.0	0.000	0.0
PI	0.001	0.2	0.001	0.2	0.001	0.2	0.001	0.2	0.001	0.2	0.001	0.2	0.000	0.1	0.002	0.5	0.001	0.3
PO	0.006	1.9	0.006	1.9	0.006	2.0	0.005	2.0	0.006	2.0	0.008	2.1	0.006	1.5	0.007	2.1	0.006	2.1
IO	0.001	0.2	0.001	0.2	0.000	0.1	0.001	0.2	0.001	0.2	0.000	0.1	0.000	0.1	0.001	0.3	0.001	0.3
PIO	0.002	0.7	0.002	0.7	0.002	0.6	0.002	0.6	0.002	0.6	0.002	0.5	0.001	0.3	0.004	1.1	0.004	1.2
SE	0.025		0.024		0.024		0.023		0.027		0.033		0.018		0.034		0.032	
Gr	0.97		0.97		0.97		0.98		0.97		0.97		0.98		0.96		0.97	
Ga	0.97		0.97		0.97		0.97		0.97		0.97		0.98		0.96		0.96	
GM	3.29		3.31		3.31		3.30		3.31		3.34		3.33		3.35		3.29	
TCI	0.98		0.98		0.98		0.98		0.98		0.98		0.98		0.98		0.98	
SCI	0.02		0.02		0.02		0.02		0.02		0.02		0.02		0.02		0.02	

below 0.40 (i.e., no  $SCI \geq 0.60$ ), while only one item in the English SCS-SF (i.e., “When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people”) had an SCI score of 0.60 which met the cut-off score for a state item. Besides that, another item from the Common Humanity subscale, “I try to see my failings as part of the human condition” which showed notable differences in SCI/TCI scores between the English SCS-SF and the VSCS-SF. This item had an SCI score of 0.23 in English SCS-SF but it was only 0.48 in the VSCS-SF. The minor differences between two studies, such as the obtained G-coefficients, may be due to heterogeneity in study designs, such as different sample sizes, cultural differences, or different languages used (English versus Vietnamese version). Therefore, this study added cross-cultural evidence supporting the psychometric properties of individual items in the short-forms of the Self-Compassion Scale.

Our findings demonstrated that the construct of self-compassion operationalised by (Neff, 2003a) and translated into Vietnamese represents a personality trait. This is an important contribution to the theoretical understanding of individual differences and the nature of this construct, which is consistent across English and Vietnamese samples. This means that by measuring self-compassion traits using both the SCS and VSCS in individuals, we can reliably predict relevant health-related outcomes (e.g., distress, psychological well-being) as already shown by several studies (MacBeth & Gumley, 2012; Marsh et al., 2018; Nguyen & Le, 2021). However, we also found specific aspects of self-compassion that are more dynamic compared to others, which suggests that such aspects can be easily amended by means

of an intervention because reinforcing a state leads to developing a trait (Miller et al., 2021).

### Limitations and Directions of Future Research

The sample in this research contained large population of females (80%) and this gender imbalance may impact on the results. Previous research indicated that women tended to report slightly lower self-compassion scores on the SCS compared to men (Yarnell et al., 2015, 2019). Therefore, it would be beneficial to replicate the analyses in future studies with a more gender balanced Vietnamese sample. In addition, the VSCS-SF was analysed using data from the full scale which is a potential limitation that may influence the results because participants may respond to items differently when completing the VSCS-SF separately. Unlike the English SCS-SF, which was validated by Raes et al. (2011), the VSCS-SF has not validated before this study. Therefore, the findings of this study support psychometric properties of the Vietnamese SCS-SF. Furthermore, even though the evidence showed that the scores of self-compassion scales were valid and reliable, these scores constitute an ordinal measure, which still have common limitations of an ordinal scale (Hobart & Cano, 2009). Therefore, further studies can be conducted to enhance the precision of these measures up to interval-level data by applying Rasch analysis (Rasch, 1961). The VSCS could be adapted to use with children and adolescents and subsequently validated with G-theory and Rasch analysis. Subsequently, longitudinal studies of self-compassion and other mindfulness-related measures can be conducted in Vietnam, from childhood or adolescence to adulthood,

to evaluate how exposure to culturally or socially prevalent Buddhism can influence trait mindfulness and self-compassion over time. Experiments and interventions can then be designed and conducted based on the results. For example, mindfulness practices may be indirectly transferred through social interaction with Buddhist followers or non-Buddhist mindfulness practitioners and thus confer benefits in health and well-being to those exposed to them. Finally, a comparison of SCIs between the VSCS and the English SCS-SF suggested that there could be important cultural differences affecting the degree of state vs trait for an item, but future studies should quantify these differences in SCIs between cultures and investigate why there may be differences between state and trait aspects. Moreover, one could compare SCIs between natives and residents of Vietnam vs Vietnamese immigrants in other countries to start to disentangle country-specific vs culture-specific influences. Similarly, one could compare SCIs and self-compassion scores over time for immigrants who move to Vietnam and indirectly observe cultural effects, including exposure to Buddhism.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12671-022-01950-3>.

**Author Contribution** QCT: directed and designed the study, analysed the data, and authored the paper. TMN: coordinated data collection, contributed to research design, and edited the manuscript. AHN: contributed to data collection and manuscript revision. AM: contributed to writing and editing the manuscript. ONM: contributed to designing the study and edited the manuscript. All co-authors have made substantial contributions that meet the stated requirements for authorship.

**Data Availability** All data are available at the Open Science Framework (<https://osf.io/uh2wn>).

## Declarations

**Ethical Statement** This study was approved by The Ethical Committee of Psychological Research, Department of Psychology, Hoa Sen University, Vietnam (approval code: HSUPSY/21/03.01).

**Informed Consent Statement** All participants provided informed consent prior to their participation in the study.

**Conflict of Interest** The authors declare no competing interests.

## References

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Brooks/Cole.
- Arterberry, B. J., Martens, M. P., Cadigan, J. M., & Rohrer, D. (2014). Application of generalizability theory to the big five inventory. *Personality and Individual Differences*, 69, 98–103. <https://doi.org/10.1016/j.paid.2014.05.015>
- Bloch, R., & Norman, G. (2012). Generalizability theory for the perplexed: A practical introduction and guide: AMEE guide no. 68. *Medical Teacher*, 34, 960–992. <https://doi.org/10.3109/0142159X.2012.703791>
- Breines, J. G., & Chen, S. (2013). Activating the inner caregiver: The role of support giving schemas in increasing state self-compassion. *Journal of Experimental Social Psychology*, 49(1), 58–64. <https://doi.org/10.1016/j.jesp.2012.07.015>
- Brennan, R. (2011). Generalizability theory and classical test theory. *Applied Measurement in Education*, 24, 1–21. <https://doi.org/10.1080/08957347.2011.532417>
- Cahn, B. R., & Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin*, 132, 180–211. <https://doi.org/10.1037/0033-2909.132.2.180>
- Cardinet, J., Pini, G., & Johnson, S. (2010). *Applying generalizability theory using eduG*. Routledge Academic.
- Chen, J., Yan, L.-S., & Zhou, L.-H. (2011). Reliability and validity of Chinese version of self-compassion scale. *Chinese Journal of Clinical Psychology*, 19(6), 734–736.
- Chiesa, A., & Serretti, A. (2009). Mindfulness-based stress reduction for stress management in healthy people: A review and meta-analysis. *Journal of Alternative and Complementary Medicine*, 15(5), 593–600. <https://doi.org/10.1089/acm.2008.0495>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, 16(2), 137–163. <https://doi.org/10.1111/j.2044-8317.1963.tb00206.x>
- de Souza, L. K., & Hutz, C. S. (2016). Adaptation of the self-compassion scale for use in Brazil: Evidences of construct validity. *Temas Em Psicologia*, 24(1), 159–172. <https://doi.org/10.9788/tp2016.1-11>
- Epstein, S. (1994). Trait theory as personality theory: Can a part be as great as the whole? *Psychological Inquiry*, 5(2), 120–122. [https://doi.org/10.1207/s15327965pli0502\\_4](https://doi.org/10.1207/s15327965pli0502_4)
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.1007/s12671-017-0850-7>
- Flentje, L. (2021). *The relationship between self-efficacy and self-compassion: An experience sampling study* (Master's thesis, University of Twente, Netherlands).
- Forrest, S. J., Siegert, R. J., Krägeloh, C. U., Landon, J., & Medvedev, O. N. (2021). Generalizability theory distinguishes between state and trait anxiety. *Psychological Assessment*, 33(11), 1080–1088. <https://doi.org/10.1037/pas0001060>
- Hobart, J., & Cano, S. (2009). *Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods*. Prepress Projects.
- Kirby, J. N., Tellegen, C. L., & Steindl, S. R. (2017). A meta-analysis of compassion-based interventions: Current state of knowledge and future directions. *Behavior Therapy*, 48, 778–792. <https://doi.org/10.1016/j.beth.2017.06.003>
- Lathren, C. R., Rao, S. S., Park, J., & Bluth, K. (2021). Self-compassion and current close interpersonal relationships: A scoping literature review. *Mindfulness*, 12, 1078–1093. <https://doi.org/10.1007/s12671-020-01566-5>
- Leary, M. R., Tate, E. B., Adams, C. E., Batts Allen, A., & Hancock, J. (2007). Self-compassion and reactions to unpleasant self-relevant events: The implications of treating oneself kindly. *Journal of Personality and Social Psychology*, 92(5), 887–904. <https://doi.org/10.1037/0022-3514.92.5.887>
- Leaviss, J., & Uttley, L. (2015). Psychotherapeutic benefits of compassion-focused therapy: An early systematic review. *Psychological*

- Medicine*, 45(5), 927–945. <https://doi.org/10.1017/S0033291714002141>
- Lee, W. K., & Lee, K. (2010). The validation study of the Korean version of Self-Compassion Scale with adult women in community. *Journal of Korean Neuropsychiatric Association*, 49(2), 193–200.
- MacBeth, A., & Gumley, A. (2012). Exploring compassion: A meta-analysis of the association between self-compassion and psychopathology. *Clinical Psychology Review*, 32, 545–552. <https://doi.org/10.1016/j.cpr.2012.06.003>
- Marsh, I. C., Chan, S. W. Y., & MacBeth, A. (2018). Self-compassion and psychological distress in adolescents—A meta-analysis. *Mindfulness*, 9, 1011–1027. <https://doi.org/10.1007/s12671-017-0850-7>
- McDonald, R. P. (2013). *Test theory: A unified treatment*. Psychology press.
- Medvedev, O. N., Berk, M., Dean, O. M., Brown, E., Sandham, M. H., Dipnall, J. F., McNamara, R. K., Sumich, A., Krägeloh, C. U., Narayanan, A., & Siegert, R. J. (2021a). A novel way to quantify schizophrenia symptoms in clinical trials. *European Journal of Clinical Investigation*, 51(3), e13398. <https://doi.org/10.1111/eci.13398>
- Medvedev, O. N., Dailianis, A. T., Hwang, Y. S., Krägeloh, C. U., & Singh, N. N. (2021b). Applying generalizability theory to the Self-Compassion Scale to examine state and trait aspects and generalizability of assessment scores. *Mindfulness*, 12(3), 636–645. <https://doi.org/10.1007/s12671-020-01522-3>
- Medvedev, O. N., Krägeloh, C. U., Narayanan, A., & Siegert, R. J. (2017). Measuring mindfulness: Applying generalizability theory to distinguish between state and trait. *Mindfulness*, 8(4), 1036–1046. <https://doi.org/10.1007/s12671-017-0679-0>
- Miller, Y. R., Medvedev, O. N., Hwang, Y.-S., & Singh, N. N. (2021). Applying generalizability theory to the Perceived Stress Scale to evaluate stable and dynamic aspects of educators' stress. *International Journal of Stress Management*, 28(2), 147–153. <https://doi.org/10.1037/str0000207>
- Muris, P., van den Broek, M., Otgaar, H., Oudenhoven, I., & Lennartz, J. (2018). Good and bad sides of self-compassion: A face validity check of the Self-Compassion Scale and an investigation of its relations to coping and emotional symptoms in non-clinical adolescents. *Journal of Child and Family Studies*, 27(8), 2411–2421. <https://doi.org/10.1007/s10826-018-1099-z>
- Neff, K. D. (2003a). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. <https://doi.org/10.1080/15298860309027>
- Neff, K. D. (2003b). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2, 85–102. <https://doi.org/10.1080/15298860309032>
- Neff, K. D., Tóth-Király, I., Knox, M. C., Kuchar, A., & Davidson, O. (2021). The development and validation of the state Self-Compassion Scale (long- and short form). *Mindfulness*, 12(1), 121–140. <https://doi.org/10.1007/s12671-020-01505-4>
- Neff, K. D., Whittaker, T., & Karl, A. (2017). Evaluating the factor structure of the Self-Compassion Scale in four distinct populations: Is the use of a total self-compassion score justified? *Journal of Personality Assessment*, 99(6), 596–607. <https://doi.org/10.1080/00223891.2016.1269334>
- Nguyen, T. M., & Le, G. N. H. (2021). The influence of COVID-19 stress on psychological well-being among Vietnamese adults: The role of self-compassion and gratitude. *Traumatology*, 27(1), 86–97. <https://doi.org/10.1037/trm0000295>
- Nguyen, T. M., Bui, T. T. H., Xiao, X., & Le, V. H. (2020). The influence of self-compassion on mindful parenting: A mediation model of gratitude. *The Family Journal*, 28(4), 455–462. <https://doi.org/10.1177/1066480720950421>
- Paterson, J., Medvedev, O. N., Sumich, A., Tautolo, E. S., Krägeloh, C. U., Sisk, R., McNamara, R. K., Berk, M., Narayanan, A., & Siegert, R. J. (2018). Distinguishing transient versus stable aspects of depression in New Zealand Pacific Island children using generalizability theory. *Journal of Affective Disorders*, 227, 698–704. <https://doi.org/10.1016/j.jad.2017.11.075>
- Petrocchi, N., Ottaviani, C., & Couyoumdjian, A. (2013). Dimensionality of self-compassion: Translation and construct validation of the self-compassion scale in an Italian sample. *Journal of Mental Health*, 23(2), 72–77. <https://doi.org/10.3109/09638237.2013.8418>
- Quang, A. M. T., Van Pham, M., Mai, T. T., Le, G. N. H., & Song, G. A. N. (2021). Self-compassion and Students' Well-Being Among Vietnamese Students: Chain Mediation Effect of Narcissism and Anxiety. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, 1–16. <https://doi.org/10.1007/s10942-021-00431-1>
- Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*, 18, 250–255. <https://doi.org/10.1002/cpp.702>
- Ramanaiah, N. V., Franzen, M., & Schill, T. (1983). A psychometric study of the State-Trait Anxiety Inventory. *Journal of Personality Assessment*, 47(5), 531–535. [https://doi.org/10.1207/s15327752jpa4705\\_14](https://doi.org/10.1207/s15327752jpa4705_14)
- Rasch, G. (1961). On general laws and the meaning of measurement. *Psychology: Proceedings of the Fourth Berkeley Symposium Mathematics Statistics and Probability*, 5, 321–333.
- Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist*, 44(6), 922–932. <https://doi.org/10.1037/0003-66X.44.6.922>
- Spielberger, C. D. (1999). *Manual for the State-Trait Anger Expression Inventory-2*. Psychological Assessment Resources Inc.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Test manual for the State Trait Anxiety Inventory*. Consulting Psychologists Press.
- Swiss Society for Research in Education Working Group. (2006). *EDUG user guide*. IRDP.
- Thøgersen-Ntoumani, C., Dodos, L., Chatzisarantis, N., & Ntoumanis, N. (2017). A diary study of self-compassion, upward social comparisons, and body image related outcomes. *Applied Psychology: Health and Well-Being*, 9(2), 242–258. <https://doi.org/10.1111/aphw.12089>
- Tran, T. H., & Tran, M. D. (2017). Evaluation of self-compassion: A quantitative study in Vietnamese students [in Vietnamese]. *The Journal of Psychology*, 10(223), 13–23.
- Truong, Q. C., Choo, C., Numbers, K., Merkin, A. G., Brodaty, H., Kochan, N. A., et al. (2021). Applying generalizability theory to examine assessments of subjective cognitive complaints: whose reports should we rely on – participant versus informant? *International Psychogeriatrics*, 1–11. <https://doi.org/10.1017/S1041610221000363>
- Truong, Q. C., Choo, C., Numbers, K., Merkin, A. G., Brodaty, H., Kochan, N. A., Sachdev, P. S., Feigin, V. L., & Medvedev, O. N. (2022). Clinical investigation of dynamic and enduring aspects of global cognition in aged population. *European Journal of Clinical Investigation*, 52, e13681. <https://doi.org/10.1111/eci.13681>
- Truong, Q. C., Krägeloh, C. U., Siegert, R. J., Landon, J., & Medvedev, O. N. (2020). Applying generalizability theory to differentiate between trait and state in the Five facet mindfulness questionnaire (FFMQ). *Mindfulness*, 11(4), 953–963. <https://doi.org/10.1007/s12671-020-01324-7>

- Waring, S. V., & Kelly, A. C. (2019). Trait self-compassion predicts different responses to failure depending on the interpersonal context. *Personality and Individual Differences, 143*, 47–54. <https://doi.org/10.1016/j.paid.2019.01.043>
- Wilson, A. C., Mackintosh, K., Power, K., & Chan, S. W. (2019). Effectiveness of self-compassion related therapies: A systematic review and meta-analysis. *Mindfulness, 10*, 979–995. <https://doi.org/10.1007/s12671-018-1037-6>
- Yarnell, L. M., & Neff, K. D. (2013). Self-compassion, interpersonal conflict resolutions, and well-being. *Self and Identity, 12*(2), 146–159. <https://doi.org/10.1080/15298868.2011.649545>
- Yarnell, L. M., Neff, K. D., Davidson, O. A., & Mullarkey, M. (2019). Gender differences in self-compassion: Examining the role of gender role orientation. *Mindfulness, 10*(6), 1136–1152.
- Yarnell, L. M., Stafford, R., Neff, K. D., Reilly, E., Knox, M. C., & Mullarkey, M. (2015). Meta-analysis of gender differences in self-compassion. *Self and Identity, 14*, 499–520.
- Zeng, X., Wei, J., Oei, T. P., & Liu, X. (2016). The self-compassion scale is not validated in a Buddhist sample. *Journal of Religion and Health, 55*(6), 1996–2009. <https://doi.org/10.1007/s10943-016-0205z>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.