Kellner’s Symptom Questionnaire, a Highly Sensitive Patient-Reported Outcome Measure: Systematic Review of Clinimetric Properties

Giada Benasi\textsuperscript{a} Giovanni A. Fava\textsuperscript{b} Chiara Rafanelli\textsuperscript{a}

\textsuperscript{a}Department of Psychology, University of Bologna, Bologna, Italy; \textsuperscript{b}Department of Psychiatry, University at Buffalo, State University of New York, Buffalo, NY, USA

Keywords
Symptom questionnaire · Patient-reported outcome · Self-rating scale · Clinimetrics · Clinical pharmacopsychology · Anxiety · Depression · Somatization · Hostility · Psychological well-being

Abstract
Introduction: Patient-reported outcomes (PROs) are of increasing importance in clinical medicine. However, their evaluation by classic psychometric methods carries considerable limitations. The clinimetric approach provides a viable framework for their assessment. Objective: The aim of this paper was to provide a systematic review of clinimetric properties of the Symptom Questionnaire (SQ), a simple, self-rated instrument for the assessment of psychological symptoms (depression, anxiety, hostility, and somatization) and well-being (contentment, relaxation, friendliness, and physical well-being). Methods: The PRISMA guidelines were used. Electronic databases were searched from inception up to March 2019. Only original research articles, published in English, reporting data about the clinimetric properties of the SQ, were included. Results: A total of 284 studies was selected. The SQ has been used in populations of adults, adolescents, and older individuals. The scale significantly discriminated between subgroups of subjects in both clinical and nonclinical settings, and differentiated medical and psychiatric patients from healthy controls. In longitudinal studies and in controlled pharmacological and psychotherapy trials, it was highly sensitive to symptoms and well-being changes and discriminated between the effects of psychotropic drugs and placebo. Conclusions: The SQ is a highly sensitive clinimetric index. It may yield clinical information that similar scales would fail to provide and has a unique position among the PROs that are available. Its use in clinical trials is strongly recommended.

Introduction

Patient-reported outcomes (PROs), any report coming directly from patients about how they function or feel in relation to a health condition or its therapy, are of increasing importance in clinical medicine and psychology [1]. Some PROs focus on self-rated evaluation of specific disease-related conditions, such as cancer, pain, or depression [1, 2]. Other indices are focused on more general perceptions, such as quality of life and psychological well-being [3–5].

In clinical psychology, the use of self-rating scales for evaluating the psychological status in conjunction or al-
ternative to observer-rated methods has paved the ground for the developments of PROs [3, 6]. In psychiatry, self-rating scales, long before the appearance of PROs, have been part of assessment tools for clinical trials [7]. Guidelines for inclusion of PROs in clinical trial protocols have recently been developed [8]. However, the application of the classical psychometric model to the clinical challenges appears to be inadequate [9, 10]. The homogeneity of items, as measured by statistical tests such as Cronbach’s alpha, has often been the most important requirement for a rating scale [10]. However, the same properties that give a scale a high score for internal consistency may obscure its ability to detect change. The redundant nature of the items of a scale may increase Cronbach’s alpha but decrease its sensitivity [10]. The ability of a rating scale to discriminate between different groups of patients suffering from the same illness (e.g., depressed inpatients and outpatients) and to reflect changes in experiments in therapeutics such as drug trials has been defined by Kellner [11] as sensitivity. Scales may be valid and reliable but may lack sensitivity. This is particularly important when treatment effects are small and with limited sample sizes. The concept of sensitivity refers to both the detection of psychological states (whether symptoms or well-being) and their changes with treatment.

In 1982, Alvan Feinstein [12] introduced the term clinimetrics, to indicate a domain concerned with the measurement of clinical issues that do not find room in customary clinical taxonomy. Clinimetrics has a set of rules which govern the structure of indices, the choice of component variables, the evaluation of consistency and validity, and that differ from classical psychometrics, which developed outside the clinical field, mainly in the educational and social areas [9, 10, 13–15].

The Symptom Questionnaire (SQ) is a simple, self-rated questionnaire that was developed by Robert Kellner in 1976 [16] and can be used for the assessment of both symptoms and well-being. Its psychometric properties have been outlined by the author in a paper published more than 30 years ago [17]. The aim of this paper is to update this work and provide a comprehensive review of clinimetric properties of the SQ.

Development and Characteristics of the SQ

The SQ was developed by Robert Kellner with the aim of providing a scale to be used in clinical research that could be more sensitive than other instruments [17].

The SQ was originated from the Symptom Rating Test (SRT) [18], a self-rating scale specifically designed to measure changes in distress among neurotic patients participating in efficacy trials, such as drug trials. The design principles of the SQ were therefore those of a distress scale, and its items were derived from the same list of neurotic symptoms used to create the SRT.

Nevertheless, the SQ has had a long evolution [16] and each stage in its development was based on empirical findings [17]. The SQ anxiety, depression, and somatization scales were created from a review of the literature on factor analyses of symptoms of psychiatric patients and normal controls. Items were included in the final version of these scales based on their ability to discriminate between depressed and anxious nonpsychotic psychiatric patients and normal subjects, and between psychotropic drug and placebo in three drug trials [17]. The hostility scale was constructed according to a clinimetric approach. Statements that two investigators agreed to consider as expression of anger or hostility were selected from interviews with patients with neurotic or personality disorders. Items were retained in the final version of the scale if they were reported significantly more frequently by hostile patients compared with normal subjects or patients judged to be not hostile in two studies. Furthermore, the literature on factor analyses of symptoms in psychiatric patients and normal subjects was searched, and items were retained if they were part of a factor of anger, hostility, or irritability [17].

The SQ differs from the SRT in that it has brief items instead of questions, and yes/no or true/false responses instead of scales of severity or frequency of symptoms. The total number of items was increased, and statements of well-being were included in order to improve the sensitivity of the scale [19, 20].

Description

The final version of the SQ consists of 92 items and yields 4 main scales: depression, anxiety, hostility, and somatization (see online suppl. Appendix 1; see www.karger.com/doi/10.115900506110 for all online suppl. material). Each scale can be divided into 2 subscales, one concerned with symptoms and the other with well-being, for a total of 8 subscales. Therefore, each of the main scales includes items from both the symptoms and the well-being subscales (Table 1).

Answers are dichotomous, and the respondent is asked to check YES/NO or TRUE/FALSE for each item. Scales and subscales can be scored separately, and the sum of the 4 main scale scores yields a total distress score. Instructions
for scoring are reported in the online supplementary Appendix 2. Norms for the interpretation of results are available in Kellner [17]. Moderate distress is indicated by a scale score between 1 and 2 standard deviations above the mean for normal subjects, while severe distress or psychopathology are suggested by a scale score of 2 standard deviations above the mean. A single high score in one or more of the SQ scales is not enough to make a diagnosis of psychopathology, and further clinical assessment is required. In particular, a high score on the somatization scale needs to be interpreted with caution when a medical disease is present.

Two forms of the SQ are available, which differ only for a different time focus. The week form is concerned with feelings experienced by the respondent during the past week, while the day form with feelings experienced on the day of the test. These two forms may serve different purposes in research. The week form is the most commonly used [17].

Clinical Applications

The SQ has been translated into several languages, such as Italian, French, German, Spanish, Portuguese, Dutch, Urdu, Punjabi, Cantonese, Mandarin, Arabic dialects, Russian, Swedish, and Hindi. Most of the studies in which the SQ was administered were performed in Italy, the USA, Canada, and the UK. In these studies, the SQ has been used to assess levels of distress and well-being across clinical and nonclinical populations, and as an outcome variable to test the efficacy of pharmacological and psychological interventions. In longitudinal studies, the scale has been used to detect significant changes in symptomatology. Moreover, several investigations were concerned with the relationship with other rating scales or constructs.

Methods

Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [21] were used to perform the present systematic review.

The search was carried out from inception until March 2019 in the following databases: PubMed, PsycINFO, and MEDLINE, KCI-Korean Journal Database, Russian Science Citation Index, and SciELO citation index via Web of Science. In each database the key words “Symptom Questionnaire,” “SQ,” and “Kellner” were used.

The databases SCOPUS and Web of Science were consulted for articles citing the work “A Symptom Questionnaire” of Kellner [17], that is the main reference for the SQ. The reference lists of the retrieved articles were also examined for additional studies.

Study Selection

Only articles published in English and reporting data about the clinimetric properties of the SQ were selected. A first screening of titles and abstracts was performed to exclude articles published in a language other than English, nonoriginal research articles (e.g., books, meeting abstracts, letters, commentaries, reviews, etc.), and papers that were clearly irrelevant. The full texts of the remaining papers were analyzed.

The search, selection, and analysis of the selected studies were performed independently by two reviewers (G.B. and C.R.); disagreements were resolved by consensus among these primary raters and a senior investigator (G.A.F.).

Data Extraction

Data were independently extracted by both reviewers with the use of a precoded form. The following data were extracted from studies meeting criteria for inclusion in the systematic review: country and field in which the study was performed; sample characteristics and size; measures and statistical analyses; clinimetric data.

Table 1. SQ scales and subscales

<table>
<thead>
<tr>
<th>Scales (92 items)</th>
<th>Symptom subscales (68 items)</th>
<th>Well-being subscales (24 items)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>Depression</td>
<td>Contentment</td>
</tr>
<tr>
<td></td>
<td>(e.g., item 6: sad, blue;</td>
<td>(e.g., Item 4: cheerful;</td>
</tr>
<tr>
<td></td>
<td>item 61: not interested in</td>
<td>item 7: happy)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Anxiety</td>
<td>Relaxation</td>
</tr>
<tr>
<td></td>
<td>(e.g., item 1: nervous;</td>
<td>(e.g., item 9: feeling calm;</td>
</tr>
<tr>
<td></td>
<td>item 36: scared)</td>
<td>item 29: relaxed)</td>
</tr>
<tr>
<td>Hostility</td>
<td>Hostility</td>
<td>Friendliness</td>
</tr>
<tr>
<td></td>
<td>(e.g., item 3: irritable;</td>
<td>(e.g., item 13: feeling kind</td>
</tr>
<tr>
<td></td>
<td>item 20: angry)</td>
<td>toward people; item 35:</td>
</tr>
<tr>
<td>Somatization</td>
<td>Somatization</td>
<td>patient)</td>
</tr>
<tr>
<td></td>
<td>(e.g., item 12: feeling of</td>
<td>(e.g., item 10: feeling</td>
</tr>
<tr>
<td></td>
<td>not enough air; item 77:</td>
<td>healthy; item 19: no pains</td>
</tr>
<tr>
<td></td>
<td>muscle pains)</td>
<td>anywhere)</td>
</tr>
</tbody>
</table>
Results

A total of 284 studies were included in the review (online suppl. Fig. 1). Of these, 232 research articles were found to display the clinimetric properties of the SQ and are included in the present review. Other papers in which the SQ was used will not be discussed in detail (online suppl. Additional References).

Specifically, we focused on data about discriminant validity, sensitivity to change, concurrent and divergent validity, relations to other dimensions (including brain correlates and biomarkers), and predictive validity. The differential sensitivity of the SQ compared to other scales, and differences between symptoms and well-being subscales, have been highlighted and discussed.

**Discriminant Validity**

**Patients versus Controls**

Most of the SQ scales and subscales were able to significantly discriminate in the expected direction between psychiatric patients and controls [22–41]. In some of these studies, the SQ showed particularly high discriminant validity, being able to sensitively differentiate remitted patients with residual affective symptoms from healthy subjects [24, 29–31, 40, 41].

In several investigations, the SQ scores were also able to differentiate medical patients from healthy controls. In all of these studies, levels of distress and well-being were significantly worse among family practice patients [25, 35], breast cancer survivors [42], and patients affected by endocrine [43–50] and gastrointestinal [51, 52] conditions than among healthy subjects.

In two studies, the SQ was administered to a sample of adolescents and significantly discriminated between those with or without epilepsy [53] or endocrine disorders [54].

**Different Groups of Patients and Healthy Subjects**

The SQ has been used in many clinical investigations to discriminate between subgroups of patients across different medical and psychiatric settings.

In psychiatry, the SQ significantly discriminated between subgroups of patients with affective disorders [22, 25, 27, 28, 37, 55–63] or post-traumatic stress disorder [64, 65].

Most of the SQ scales and subscales were able to discriminate between different subgroups of medical patients in cardiology [66–87], endocrinology [43–47, 50, 88], oncology [42, 89–93], dermatology [94], gynecology [95–100], pneumology [101], and general medical settings [102–104], based on the presence of specific medical or psychological features and comorbidities.

In two studies comparing psychiatric and medical patients with each other, the SQ scores were significantly worse among psychiatric patients than among patients in family practice [25, 35].

The SQ has been used in a variety of other studies to discriminate between subgroups of healthy subjects. In some of these investigations, the scale allowed the identification of significant differences in levels of distress and well-being among subjects facing different sources of distress, such as illness of a family member [105–108], the occurrence of a major life event [109, 110], and exposure to childhood maltreatment [111–118] or intimate partner violence [119, 120]. In two other studies, the SQ discriminated between personality clusters among parents of patients with eating disorders [121], and between law and medical students [122].

There were no clear indications for gender differences in levels of distress and well-being assessed by the SQ [28, 84, 91, 105, 122–128].

**Symptoms versus Well-Being Subscales**

Except for two studies, in which the well-being subscales appeared to discriminate more sensitively between patients and controls [38, 47], most of the symptoms subscales were more sensitive than their well-being counterparts in differentiating psychiatric and medical patients from healthy controls [22, 23, 30, 34, 35, 37, 40, 42, 50].

An opposite pattern was observed for the well-being subscales when discriminating among different subgroups of psychiatric and medical patients, or healthy subjects [22, 37, 42, 50, 55, 62, 80, 111, 122].

**Cutoff Scores**

Various cutoff points have been suggested for the SQ scales and subscales [17], yet there has not been consensus about their positioning [22, 53, 69, 70, 72, 78, 79, 81–86, 90, 100, 103, 104, 129–137].

The SQ can be used as a screening tool for the identification of cases on the basis of normative parameters and cutoff points. However, it is not intended to be a diagnostic instrument. Rather, in the diagnostic process, a thorough clinical interview and assessment could include both self-report and observer-rated measures.

**Sensitivity to Change/Responsiveness**

Sensitivity to change has been assessed in numerous studies in which the SQ was used as a repeated rating outcome measure.
Pharmacological Trials

In a variety of open studies, the SQ was able to show significant improvements in levels of psychological distress and well-being among patients with different psychiatric [27, 32, 37–39, 59, 61, 63, 100, 138–153] and medical conditions [154, 155], receiving psychotropic medications or other types of pharmacological treatments. In three of these studies [145, 147, 148], a pain subscale, based on the pain-related items of the SQ somatization scale, was used and showed to be sensitive to treatment changes.

In four studies [139, 151, 156, 157], the SQ scales were sensitive to changes which occurred at treatment interruption among patients with affective disorders who were taking psychotropic medications. In two investigations [139, 157], the SQ was also able to discriminate withdrawal symptoms that occurred after the interruption of selective serotonin reuptake inhibitors.

The ability of the SQ to discriminate drug effects has been evaluated in a number of placebo-controlled trials. Most of the SQ scales and subscales were highly sensitive in discriminating between effects of drugs and placebo among both psychiatric and medical patients [19, 20, 158–163]. Compared to placebo, drugs yielded significantly greater improvements in all studies.

In four negative placebo-controlled trials [100, 137, 164, 165], significant improvements from baseline to the end of treatment were detected in SQ scores, without significant differences between drugs and placebo. The same results were obtained with observer-rated instruments [137, 165].

Finally, the SQ was found to detect significant changes not only in treatment studies, but also after pharmacological challenges in patients with depression after tryptophan depletion [166] and fenfluramine challenge [167].

Psychotherapy Studies

In a number of open trials, the SQ scores changed with psychological interventions in both clinical [26, 29, 64, 168–171] and nonclinical populations [172–175].

The ability of the SQ to discriminate treatment effects has been evaluated in several controlled trials. In these trials, at least one of the SQ scales and subscales significantly discriminated between the effects of different psychological interventions [31, 176–183]. Three of these studies were performed among middle and high school students receiving either a school-based protocol derived from well-being therapy or attention placebo [181], anxiety management [182], and cognitive behavioral therapy [180].

In other controlled studies, the SQ scales showed significant improvements from baseline to the end of psychological interventions, without significant differences between treatment groups [57, 58, 184–188].

Rehabilitation Studies

In several investigations, levels of anxiety, depression, somatization, and hostility significantly decreased, according to the SQ scales and subscales, in different subgroups of patients undergoing cardiac or pulmonary rehabilitation [67–70, 72, 78, 80–87, 189–196], and among employees receiving a worksite health intervention [197]. In some of these studies, the SQ significantly discriminated between different degrees of improvement among subgroups of coronary artery disease patients [72, 78, 82, 84, 87, 191, 195].

Medical Course and Procedures

The SQ scores changed significantly among women undergoing amniocentesis [34, 95, 96, 98], fetoscopy [97], or ultrasound examination [198]. In all of these studies, levels of distress and well-being significantly improved after the performance of the tests, with a further improvement when normal results were communicated.

In other longitudinal studies, the SQ scales and subscales were sensitive to changes in symptomatology observed among primipara women during the first 15 days postpartum [199], mothers of premature infants during the first 24 days of hospitalization in a neonatal intensive care unit [200], family caregivers of an elderly person at 2 weeks and 2 months after discharge from the hospital [201], and among medical patients at different evaluations over a period of time between 6 months and 5 years [66, 91, 92, 103, 202].

In four studies, the SQ scores significantly changed after the performance of a specific medical procedure, such as surgery or diagnostic testing [203–206].

Nonmedical Contexts

The SQ has also been administered in healthy populations. In the elderly a significant increase in levels of distress was detected after a major life crisis [110].

In two other studies [207, 208], the SQ was administered to medical and dental students at different points during their programs. The SQ symptoms scores significantly increased over time starting from the beginning of the program. In one of these studies [207], medical students involved in a new, problem-based, and student-centered curriculum experienced significantly lower levels of psychological distress overall as compared to students in the traditional curriculum.
Symptoms versus Well-Being Subscales

With the exception of three studies, in which at least one of the well-being subscales was more sensitive in discriminating between treatment effects [20, 177, 180], some or all of the symptom subscales appeared to be more sensitive to treatment changes and differential effects than the corresponding well-being subscales, in both pharmacological and psychotherapy trials [29, 37, 61, 80, 142, 143, 155, 159, 179, 181].

However, only the friendliness subscale was able to show significant improvements after treatment with amitriptyline in a subgroup of patients with major depressive disorder reporting losses [37].

Comparison with Other Scales

In eight studies [20, 158–160, 177, 180–182], the SQ discriminated between the effects of drugs and placebo and those of different psychological interventions more sensitively than other self-rating scales, such as the SRT [18], the Beck Depression Inventory [209], the visual analogue scale for pain [210], the Psychological Well-Being scales [211], and the Revised Children’s Manifest Anxiety Scale [212]. In two of these studies [20, 158], the SQ scales discriminated between the effects of psychotropic drugs and placebo more sensitively than observer-rated scales, including the Hamilton Anxiety Rating Scale (HARS) [213]. When a reduction in sample size was made by random methods to examine the sensitivity of the scales in a smaller sample [20], only the SQ and the SRT [18], but not the HARS [213], were still able to discriminate between chlordiazepoxide and placebo effects.

Only in two studies [214, 215] the SQ was not able to discriminate between the effect of drug and placebo, showing lower sensitivity than the visual analogue scale for pain [210] and the Hamilton Rating Scale for Depression (HRSD) [216].

In four longitudinal studies [91, 92, 199, 202], the SQ was more sensitive to symptom changes than the HRSD [216] and the Clinical Interview for Depression (CID) [217, 218], while in one investigation [66], the CID scales of anxiety and depression [217, 218] showed greater sensitivity than the corresponding scales on the SQ.

Concurrent Validity

The concurrent validity of the SQ has been examined with other self-rating scales measuring psychological distress. At least one of the SQ symptom scales significantly and positively correlated with scales measuring similar constructs on the Profile of Mood States [219], the Trauma Symptom Inventory [220], the SRT [18], the Hopkins Symptom Checklist [221], and the Center for Epidemiologic Studies Depression Scale (CES-D) [222], in both clinical and nonclinical populations [17, 34, 46, 199, 223]. In these studies, correlation coefficients ranged from 0.39 to 0.93. Moreover, significant correlations were observed between the SQ hostility scale and the Perceived Stress Scale [224] \( r = 0.46 \) [27], the SQ contentment subscale and the CES-D [222] \( r = 0.48–0.53 \) [34], and the SQ depression subscale and the SRT total neuroticism scale [18] \( r = 0.65–0.78 \) [23], in depressed patients and controls.

Correlations between the SQ and self-rating scales of psychological well-being have also been evaluated in subjects from the general Italian population [126], and among psychiatric or medical patients and healthy controls [29, 76]. Most of the Psychological Well-Being scales [211] significantly and negatively correlated with the SQ symptom scores, and significantly and positively correlated with the well-being scores [29, 76, 126]. However, the degree of these correlations was highly variable for both symptoms \( r = 0.15–0.91 \) and well-being \( r = 0.11–0.84 \) subscales.

The relationship between the SQ and several observer-rated measures of distress has been investigated. Significant and positive correlations were observed between the HRSD [216], the HARS [213], the Montgomery-Åsberg Depression Rating Scale [225], the Brief Depression Rating Scale [226], and the CID [217, 218], and at least one of the SQ symptom scales among psychiatric or medical patients [17, 22, 27, 34, 38, 63, 124, 129, 136, 202], and healthy subjects [34, 37, 38, 128, 129]. Specifically, correlation coefficients between the SQ depression score and the HRSD [216] ranged from 0.36 to 0.72 [34, 124, 128, 129], and were slightly higher than those observed with the Brief Depression Rating Scale [226] \( r = 0.45–0.57 \) [38] and the CID [217, 218] \( r = 0.28–0.68 \) [38, 202]. Correlation coefficients between the SQ anxiety score and the HARS [213] ranged from 0.56 to 0.69 [17, 128] and were higher than those observed with the CID anxiety scale [217, 218] \( r = 0.35–0.54 \) [202]. Moreover, in one investigation [34], the SQ contentment subscale significantly correlated with the HRSD [216] in both depressed patients \( r = 0.61 \) and controls \( r = 0.54 \). In three studies [38, 136, 202], the correlation between the SQ and observer-rated measures increased with improvements of patients’ clinical state and appeared to be higher when ratings did not reflect the severity of symptoms, but simply the presence or absence of symptoms.
**Associations with Other Dimensions**

### Illness Severity and Quality of Sleep

Significant and positive associations were observed between the SQ measures of distress and severity of disease in patients with medical [91, 92, 101, 155, 227] and psychiatric disorders [143].

In patients with chronic nightmare disorder receiving psychological interventions, the SQ scores of distress were significantly and inversely associated with nightmare decrease and quality of sleep [64, 168, 187].

### Distress Risk Factors

In patients with cancer, SQ scores of depression were significantly and positively correlated with age [91, 92] and days spent in isolation with fever during hospitalization [205].

In a sample of depressed patients treated with fluoxetine, posttreatment SQ scores of distress were significantly and inversely predicted by an earlier time to onset of clinical improvement (i.e., the first time point at which the HRSD score [216] decreased by at least 30% from baseline) [228] and baseline caffeine consumption [229], and significantly and positively predicted by baseline alcohol consumption and a greater number of medical comorbidities [230].

Finally, in a sample of patients with coronary artery disease, long-term statin use was significantly associated with a lower risk of developing symptoms of depression, anxiety, and hostility, according to the SQ scores [132].

### Biomarkers and Brain Correlates

Only a few studies investigated the relationship between SQ scores of distress and brain correlates or biomarkers. Specifically, significant associations were found between SQ ratings of distress and changes in brain bioenergetic function among subjects with major depression [231], alterations in the white matter tract integrity among young adults exposed to childhood adversities [114, 232], and the activation of specific areas in an extended neural network involved in facial expression response, among healthy subjects with a genetic variation near the CREB1 [233].

In a sample of depressed outpatients [234], significant and positive correlations were found between the SQ score of anxiety and cardiovascular risk factors, such as levels of cholesterol and QTc interval.

In one study [110], changes in the SQ score of distress were significantly associated with physiological indicators of distress, such as blood levels of cortisol and the absorption of calories.

In two other studies [88, 161], the SQ depression and somatization scores were significantly associated with changes in levels of growth hormone among patients with prior acromegaly, with or without current growth hormone deficiency.

### Traumatic Experiences

Among young adults, exposure to various forms of maltreatment during childhood had a significant effect on SQ ratings of distress [128, 235–238].

Moreover, intimate partner violence and coercion were significantly associated with higher levels of anxiety, depression, somatization, and hostility in samples of university students of different ethnicity [119, 120, 239–242].

### Other Psychological Constructs

Several studies have shown the existence of a significant association between at least one of the SQ ratings of distress and well-being and a variety of other constructs, in both clinical and nonclinical settings. These constructs include: measures of illness perception and hypochondriacal concerns [40, 109, 243–246]; alexithymia and emotional intelligence and regulation [183, 203, 247, 248]; positive and negative cognitions [27, 108]; anxiety sensitivity [156, 249]; socioaffective vigilance [250]; perceptions of learning environment [207]; and different aspects of well-being, such as spiritual well-being, personal and social resourcefulness, coping abilities, perception of social support, and gratitude [93, 108, 251, 252].

Among high school and undergraduate students (i.e., 13–23 years old), the SQ scores of distress were significantly associated with measures of illness behavior [253], perception of emotional intelligence [254], and hyperactivity/inattention [123].

### Predictive Validity

The predictive validity of the SQ has been examined in a number of studies with reference to its ability to predict a criterion measure at a later time.

### Health Outcomes

In patients affected by different medical conditions, higher SQ scores of distress significantly predicted worse health outcomes, such as greater disability, symptomatology, hospitalization, and use of medications [101, 130, 255, 256].

In cardiology, patients who reported higher scores on the SQ distress scales had a significantly higher risk for adverse cardiac outcomes and mortality at 12- to 161-month follow-ups [69, 70, 72, 79, 80, 86, 133, 257, 258].
Moreover, worse SQ scores during pregnancy were significant risk factors for postpartum depression [135, 259], early cessation of breastfeeding [134], and developmental or health problems in children at 3 years of follow-up [260, 261].

Treatment Outcomes

Among patients with affective disorders receiving psychotropic medication, higher SQ scores of anxiety and somatization at baseline significantly predicted worse treatment outcomes, such as reports of side effects [262], relapse [263], and poorer or delayed onset of clinical response (i.e., ≥50% decrease in HRSD-17 [216] scores from baseline to end point) [125, 264–266]. Similarly, early improvements in the anxiety, depression, and hostility scores significantly predicted response and remission (i.e., HRSD-17 [216] score <8 at end point) after 8 weeks of treatment with fluoxetine in patients with major depressive disorder [149].

Among women with growth hormone deficiency [162] and patients with coronary artery disease [78, 82, 84, 87], higher SQ scores of distress at baseline significantly predicted greater improvements in SQ scores of distress after growth hormone therapy or cardiac rehabilitation. A different trend was observed among oncological patients, where higher baseline SQ scores of depression and anxiety were significant predictors of the development of depression after autologous bone marrow transplantation [90].

Discussion

The SQ has been used in a variety of studies, in which it was administered to different populations of adolescents, adults, and elderly. Most of the SQ scales and subscales were able to discriminate between subgroups of subjects in both clinical and nonclinical settings, and to differentiate medical or psychiatric patients from healthy controls. In all studies in which the SQ was administered to psychiatric patients and healthy controls, most scales and subscales significantly discriminated between groups. The same scales showed to be highly sensitive to changes with pharmacological and psychotherapy interventions, and to significantly discriminate between treatment effects. In longitudinal studies, both symptom and well-being scores changed over time in the expected direction. The present results have confirmed those summarized by Kellner in 1987 [17].

When compared with other self- and observer-rated scales, a greater sensitivity of the SQ has emerged in its ability to discriminate between the effects of drugs and placebo, and the effects of different psychological interventions [20, 158–160, 177, 180–182].

Compared to the well-being subscales, the SQ symptom subscales showed greater discriminant validity in differentiating between patients and controls [22, 23, 30, 34, 35, 37, 40, 42, 50], and greater sensitivity to treatment changes and differential effects [29, 37, 61, 80, 142, 143, 155, 159, 179, 181]. However, the well-being subscales were more sensitive to other differences between groups, such as those between subgroups of healthy subjects or between subgroups of patients affected by the same medical or psychiatric condition [22, 37, 42, 50, 55, 62, 80, 111, 122], and were able to show treatment changes in a specific subgroup of depressed patients, when other scales failed to do so [37]. Nevertheless, both subscales showed high sensitivity, being able to differentiate between remitted patients and healthy controls [24, 29, 30, 40, 41].

In line with data reported in Kellner [17], significant correlations were found between the SQ and other self- and observer-rating scales. In psychometrics, a high correlation is often regarded as evidence that two scales measure the same factor. However, high correlations do not necessarily indicate similar sensitivity [10, 20], as illustrated in several of the studies examined in the present review [20, 37, 38, 46, 108, 199, 202]. High, statistically significant correlations are due to the presence of common contents among scales measuring similar constructs, but the items and properties that these scales do not share determine their sensitivity [10, 20, 267, 268]. The use of both observer- and self-rated scales has been recommended to yield information that might not be revealed if only one scale is used [129].

As outlined in several studies in the present review, the SQ scores have shown a significant prognostic value, being associated with a variety of medical outcomes and biological or psychological variables. Therefore, the assessment of both levels of well-being and distress contributes not only to a more complete evaluation of the patient’s health status, but also to the development of more effective and personalized interventions. Moreover, due to its extreme sensitivity, the SQ appears to be useful in studies with small or moderate sample sizes, in which the sensitivity of the scale is important, and in psychiatric and psychosomatic investigations, where specific changes in the patient’s psychological condition are investigated.

The SQ, compared to other similar PROs, carries a number of advantages. First, it has brief items instead of questions and yes/no and true/false responses instead of scales of severity or frequency of symptoms. For its brev-
ity and simplicity, it is thus particularly suitable for populations of subjects with limited verbal skills and can be used in busy clinical practice or as an epidemiological screening procedure to differentiate moderate and severe distress or for other clinical routines, such as the assessment of treatment effects and the psychological reactions to medical procedures. Even though it is generally assumed that graded response scales provide maximum sensitivity, these may be more confusing for many patients than simple judgments of items as present or absent [269]. The sensitivity of the SQ may be related to the fact that patients are required to make a simple yes/no judgment about the presence of each symptom, with fewer opportunities for subjects to amplify or minimize symptoms to make a distinction about qualities, degrees, and patterns of distress [269].

Second, there are many self-report inventories that have been developed that are geared to the assessment of one specific dimension, often subsumed under the rubric of depression or anxiety [268, 270]. The SQ provides simultaneous assessment of both symptom and well-being dimensions. The four symptom dimensions (anxiety, depression, hostility, and somatization) are key elements of a subject’s clinical state. The subscale of hostility captures an affective component that is often neglected in diagnostic interviewing, and this is subsumed under the rubric of irritable mood, i.e., a feeling state characterized by irritability which requires an increased effort of control over temper or results in irascible verbal and behavioral outbursts [6]. There are other multidimensional scales that are available and have been found to be sensitive in clinical trials, such as the Hopkins Symptom Checklist [221, 269]. However, they lack the psychological well-being dimensions that may provide important information in specific clinical settings, such as with subclinical symptomatology and/or impaired psychological well-being.

The findings of this review thus indicate that the SQ fulfills the criteria for a comprehensive and highly sensitive clinimetric index and represents one of the most sensitive PRO measure available, if not the most sensitive. It may supply clinical information that other similar scales fail to provide and that may supplement the data derived from interview methods. It can be used with adult subjects, as well as with adolescents and in older individuals. Its use is recommended in clinical investigations concerned with psychiatric patients, with particular reference to drug trials [7], psychotherapy studies [271], and network analyses of psychopathology [272]. Because of its sensitivity, the SQ may be particularly suitable for detecting psychological distress or impairment in well-being in populations characterized by subclinical or mild symptoms [38]. In medical settings, it may disclose distress and impaired quality of life associated with disorders [6], iatrogenic psychopathology [273], and be a screening method to assess psychosocial problems [103, 274, 275]. The joint assessment of well-being and distress is in line with current emphasis on psychological well-being and the evolving science of euthymia [5]. Robert Kellner should be credited for developing in the seventies a clinimetric tool that was far ahead of its time and that is more timely than ever.

Disclosure Statement

All authors have no conflicts of interest to declare.

Author Contributions

All authors conceived the project. G. Benasi and C. Rafanelli performed the searches and collected the data. All authors analyzed the data. All authors drafted and revised the paper.

References


Symptom Questionnaire


Symptom Questionnaire


