

Der G-Score – ein Screeninginstrument zur Erfassung der subjektiven körperlichen Gesundheit

The G-Score – a Short Questionnaire for the Measurement of Subjective Physical Health

Autoren

Juliana Schneider¹, Markus Zenger^{2, 3}, Elmar Brähler⁴, Yve Stöbel-Richter^{3, 5}, Hendrik Berth¹

Institute

- 1 Psychosoziale Medizin und Entwicklungsneurowissenschaften, Forschungsgruppe Angewandte Medizinische Psychologie und Medizinische Soziologie, Universitätsklinikum Carl Gustav Carus, Dresden
- 2 Fachbereich Angewandte Humanwissenschaften, Hochschule Magdeburg-Stendal
- 3 Integriertes Forschungs- und Behandlungszentrum (IFB) AdipositasErkrankungen – Verhaltensmedizin, Medizinische Psychologie und Medizinische Soziologie, Universitätsmedizin Leipzig
- 4 Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Mainz
- 5 Fakultät Management- und Kulturwissenschaften, Hochschule Zittau/Görlitz

Schlüsselwörter

Screeninginstrument, körperliche Gesundheit, Sächsische Längsschnittstudie

Key words

physical health, Saxony Longitudinal Study, screening

eingereicht 23.03.2017

akzeptiert 23.05.2017

Bibliografie

DOI <https://doi.org/10.1055/s-0043-113558>

Online-Publikation: 11.9.2017

Psychother Psych Med 2018; 68: 361–368

© Georg Thieme Verlag KG Stuttgart · New York

ISSN 0937-2032

Korrespondenzadresse

Prof. Dr. Hendrik Berth

Psychosoziale Medizin und Entwicklungsneurowissenschaften
Forschungsgruppe Angewandte Medizinische Psychologie
und Medizinische Soziologie

Universitätsklinikum Carl Gustav Carus

Fetscherstraße 74

01307 Dresden

berth@inhaltsanalyse.de

ZUSAMMENFASSUNG

Einleitung Die Erfassung der subjektiv eingeschätzten körperlichen Gesundheit ist wichtig im Kontext vieler klinischer und Forschungsfragestellungen. In der vorliegenden Arbeit wird eine Untersuchung der psychometrischen Kriterien des G-Scores, einem 4 Items umfassenden Screeninginstrument zur Selbstbeurteilung der körperlichen Gesundheit, vorgenommen.

Material und Methoden Es erfolgten Einschätzungen der Objektivität, der Validität und der Reliabilität, eine Itemanalyse und die Bestimmung eines geeigneten Cut-Off-Wertes zur Unterscheidung gesunder von körperlich auffälligen Personen anhand von Daten der Sächsischen Längsschnittstudie 1998–2013 (N = 324–417 Personen der Normalbevölkerung).

Ergebnisse Die Objektivität des G-Scores kann aufgrund seiner Standardisierung als gegeben angesehen werden. Als Cut-Off-Wert zur Unterscheidung körperlich auffälliger von gesunden Personen wird ein G-Score von 4 vorgeschlagen. Die Inhaltsvalidität ist anhand des a posteriori definierten Konstruktes der körperlichen Gesundheit als gut einzuschätzen. Die Korrelationen mit assoziierten und nicht-assozierten Konstrukten sprechen für einen Zusammenhang körperlicher mit psychischen Beschwerden. Die prädiktive Validität des Instruments ist gut. Die Reliabilitätsanalysen ergeben Werte im zufriedenstellenden Bereich. Dies spricht für eine hohe Änderungssensitivität des erfassten Konstruktes.

Diskussion und Schlussfolgerungen Aufgrund der Ergebnisse empfiehlt sich der G-Score für den Einsatz als Screeninginstrument im Forschungsbereich. Es sollten weitere psychometrische Untersuchungen des G-Scores an einer bevölkerungsrepräsentativen Stichprobe und mit objektiven Vergleichsdaten durchgeführt werden.

ABSTRACT

Aim The measurement of subjective physical health is important in clinical settings as well as for research purposes. In the present paper, the psychometric properties of the G-Score, a 4 item screening questionnaire for the self-assessment of one's physical health, is explored.

Methods The Objectivity, validity and reliability were estimated. Moreover, an item analysis and a suitable cut-off-value for the differentiation between healthy and presumably ill subjects

were conducted. Data from the Saxony Longitudinal Study (“Sächsische Längsschnittstudie”) 1998–2013 were analysed (N = 324–417 healthy individuals).

Results The objectivity of the G-Score is estimated as very good. As a cut off for the differentiation between healthy and presumably ill subjects, a G-Score of 4 is suggested. The content validity is slightly lacking. Correlations with associated and non-associated constructs hint an interaction of physical with psychological discomfort. A good

predictive validity of the instrument is assessed. Reliability estimates show acceptably good results. This indicates a high sensitivity for changes in the measured construct.

Conclusion Putting all results into consideration, a use of the G-Score as a screening questionnaire in research is recommended. However, more psychometric investigations with representative samples and objective comparative data should be carried out.

Einleitung

Im Zuge der steigenden Prävalenz chronischer Erkrankungen rückt neben der Behandlung auch die gesundheitsbezogene Lebensqualität ins Zentrum medizinischer, psychologischer und gesundheitswissenschaftlicher Forschung [1]. Hier sind v. a. Forschungsarbeiten relevant, in denen die medizinischen und demografischen Veränderungen zusammengeführt werden, um gesundheitspolitische Aspekte langfristig zu berücksichtigen [2]. Gleichzeitig ist das Verständnis von Morbidität und Gesundheit im Zusammenhang mit Lebensqualität von Bedeutung, um Angebote und Maßnahmen wirksam gestalten zu können [3]. Gesundheitspolitisch soll v. a. die sogenannte „Über-, Unter- und Fehlversorgung“ verringert werden [4].

Es gibt sehr viele Definitionen des Gesundheitsbegriffs. Tendenziell sind in den letzten Jahren viele Dimensionen integriert worden, wie auch die Definition von Hurrelmann [5] zeigt: „Gesundheit bezeichnet den Zustand des objektiven und subjektiven Befindens einer Person, der dann gegeben ist, wenn sie sich in den physischen, psychischen und sozialen Bereichen ihrer Entwicklung im Einklang mit den eigenen Möglichkeiten und Zielvorstellungen und den jeweils gegebenen äußeren Lebensbedingungen befindet.“

Brähler und Schumacher [6] stellen 4 Möglichkeiten hinsichtlich des Zusammenspiels von objektivem Befund und subjektiver Befindlichkeit dar: Personen mit körperlichem Befund und mit subjektiven Beschwerden; Personen ohne körperlichen Befund und ohne subjektive Beschwerden; Personen mit körperlichem Befund, aber ohne subjektive Beschwerden; Personen ohne körperlichen Befund, aber mit subjektiven Beschwerden.

Körperliche Gesundheit im engeren Sinne lässt sich daher zunächst als die Abwesenheit objektiver Krankheitsfaktoren und subjektiver Beschwerden verstehen. Aber auch wenn keine subjektiven Beschwerden vorliegen, kann, objektiven Befunden zufolge, eine Erkrankung vorliegen. Umgekehrt ist es möglich, dass objektiv kein Befund vorliegt, die betroffene Person jedoch Beschwerden hat.

In der Forschung werden standardisierte Verfahren mit möglichst guten psychometrischen Eigenschaften zur Messung der subjektiv wahrgenommenen körperlichen Gesundheit benötigt, die gleichzeitig im Einsatz ökonomisch sein sollten. Es liegen einige Instrumente vor: Das Item „Gegenwärtiger Gesundheitszustand“ erwies sich im Bundesgesundheitsurvey bereits als zuverlässiger Indikator für die Gesundheit [7]. Die Somatic Symptom Scale-8 [8], ein 8 Items umfassender Kurzfragebogen zur Selbstbeurteilung, misst die Belastung der Befragten durch somatische Symptome in den letzten 7 Tagen anhand einer 5-stufigen Ratingskala. Die aus 8 Items bestehende Kurzform des Gießener Beschwerdeboogens [9] bildet aktuelle subjektive körperliche Beschwerden ab.

Für die wachsende Anzahl bevölkerungsbasierter, biomedizinisch-sozialwissenschaftlicher Langzeitstudien, die mit einem breiten Methodenspektrum von ärztlicher Untersuchung über verschiedene Laborparameter, bildgebender Verfahren usw. bis hin zu Lebensgewohnheiten und Persönlichkeitsmerkmalen, eine Vielzahl unterschiedlichster Parameter erfassen (vgl. [10]), ist es wichtig, auf validierte, im Längsschnitt verwendbare Kurzskaalen zurückgreifen zu können, von denen bereits einige verfügbar sind [11].

Ziel der vorliegenden Arbeit ist daher die Prüfung der psychometrischen Gütekriterien eines mehrfach eingesetzten Instruments zur Messung von körperlicher Gesundheit und Beschwerden, des G-Scores.

Material und Methoden

Der G-Score

Der G-Score („Gesundheits-Score“) wurde für die Sächsische Längsschnittstudie konstruiert [12, 13]. Die Instruktion lautet: „Hatten Sie in den letzten 12 Monaten folgende Beschwerden? Bitte geben Sie an, wie häufig diese auftraten“.

Die 4 Items lauten:

- 1) Nervosität
- 2) Magenbeschwerden
- 3) Schlaflosigkeit
- 4) Herzbeschwerden.

Den Befragten stehen folgende Antwortmöglichkeiten zur Verfügung: „ja, häufig“ (3), „ja, ab und zu“ (2), „ja, selten“ (1) und „nein, nie“ (0). Die Antworten der 4 Items werden aufsummiert. Der mögliche Wertebereich liegt zwischen 0 und 12. Höhere Werte stehen für eine stärkere Belastung mit Körperbeschwerden. Der G-Score erfragt im Unterschied zu den einleitend genannten Instrumenten den Gesundheitszustand der letzten 12 Monate. Er ist kürzer als bisherige Instrumente, die ähnliche Konstrukte erfassen. Zudem existiert durch den langjährigen Einsatz in der Sächsischen Längsschnittstudie eine umfangreiche Datenbasis [12].

Stichprobe

In der Sächsischen Längsschnittstudie wird seit 1987 eine für die ehemalige DDR repräsentative Stichprobe damals 14-jähriger Schüler in den DDR-Bezirken Karl-Marx-Stadt (heute Chemnitz) und Leipzig (ursprünglich N = 1281) untersucht [14]. Nach dem Abschluss der 3. Welle im Frühjahr 1989 erklärten sich N = 587 Befragte bereit, auch weiterhin an der Studie teilzunehmen [15]. Die Stu-

die konnte seitdem mit nahezu jährlich stattfindenden Erhebungen fortgesetzt werden. Im wissenschaftlichen Fokus der Studie stehen das Erleben der ostdeutschen Transformation seit 1989/1990 aus sozialwissenschaftlicher Perspektive, gesundheitspsychologische Fragestellungen sowie Aspekte der Familiengründung und des Erwerbslebens.

Der G-Score wird bereits seit 1996 (Welle 12) fortlaufend in der Sächsischen Längsschnittstudie eingesetzt. Bis einschließlich Welle 27 nahmen an den Erhebungen mit Einsatz des G-Scores mind. N = 326 Personen kontinuierlich teil. ▶ **Tab. 1** gibt einen Überblick über soziodemografische Charakteristika der Stichprobe und die verwendeten Instrumente in den untersuchten Erhebungswellen. An den im Folgenden ausgewerteten Erhebungswellen der Studie nahmen stets über 300 Personen (mehr als 55 % Rücklauf der ursprünglichen Stichprobe) teil. Es waren etwas mehr Männer als Frauen. Die Stichprobe ist altershomogen (Geburtsjahrgang 1973).

Itemanalyse

Anhand der Histogramme der Wellen 13, 17, 19, 24 und 27 wird eingeschätzt, ob eine Normalverteilung vorliegt. Zusätzlich werden Minimum und Maximum betrachtet, um sicherzugehen, dass von allen Antwortmöglichkeiten der Skala Gebrauch gemacht wurde. Die Schätzung der Itemschwierigkeit erfolgt über den Mittelwert anstelle des Medians, da dieser bei Fragebögen mit sehr wenigen Items differenzierungsfähiger ist [16]. Die part-whole-korrigierte Trennschärfe wird mittels Inter-Item-Korrelation ermittelt. Um die Itemhomogenität zu überprüfen, werden die mittleren Inter-Item-Korrelationen sowie das Minimum der Inter-Item-Korrelation herangezogen. Das Minimum wird betrachtet, um festzustellen, ob es negative Korrelationen gibt.

Validität

Für die Berechnungen der Validität wird der Rangkorrelationskoeffizient nach Spearman verwendet, da der G-Score keine Normalverteilung vorweist. Es werden Korrelationen mit der Angstska-

larskala der Depressionsskala der HADS-D [17], dem D-Score [15, 20] und der SCL-K-9 [18] berechnet. Des Weiteren werden Korrelationen mit dem GBB-24 [9], der SSS-8 [8], dem Vorkommen und der Häufigkeit von Arztbesuchen und der Einschätzung des gegenwärtigen Gesundheitszustands [7] berechnet.

Die Hospital Anxiety and Depression Scale–Deutsche Version (HADS-D) ist ein 14 Items umfassender Selbstbeurteilungsfragebogen mit 2 Subskalen zu Angst und Depressivität. Die Befragten werden gebeten, für jedes Item eine Einschätzung auf einer 4-stufigen Antwortskala (Kodierung: 0–3) anzugeben. Durch Addition der Antwortwerte werden dann je ein Angst- und ein Depressivitätsscore gebildet. Validität und Reliabilität konnten in zahlreichen Studien belegt werden. Es liegen repräsentative Normwerte vor [19].

Der D-Score [20–15], ebenfalls eine Eigenentwicklung der Sächsischen Längsschnittstudie, ist eine Kurzskala mit 4 Items. Die Items lauten: „Ich fühle mich oft niedergeschlagen und mutlos“, „Manchmal weiß ich nicht mehr, was der Sinn meines Lebens ist“, „Ich bin oft ratlos, verstehe die Welt nicht mehr“, und „Haben Sie Angst vor der Zukunft?“. Die Befragten sollen anhand einer 3-stufigen Antwortskala („Ja, und zwar seit Jahren“ - „Ja, aber erst seit einigen Monaten“ - „Nein“, kodiert mit 2 – 1 – 0 Punkten) ihre psychische Befindlichkeit einschätzen. Anschließend wird der D-Score durch Addition der jeweiligen Antwortangaben berechnet. Er kann zwischen 0 und 8 liegen, wobei 8 die schlechteste psychische Gesundheit bedeutet. Der D-Score zeichnet sich durch zufriedenstellende Gütekriterien aus [15, 20].

Die SCL-K-9 [18] ist ein Fragebogen zur Selbstbeurteilung der psychischen Belastung. Die Befragten sollen anhand einer 5-stufigen Ratingskala einschätzen, wie sehr sie in den letzten 7 Tagen unter den Beschwerden Somatisierung, Zwanghaftigkeit, Unsicherheit im Sozialkontakt, Depressivität, Ängstlichkeit, Aggressivität/Feindseligkeit, Phobische Angst, Paranoides Denken und Psychozismus litten. Die Antwortoptionen „überhaupt nicht“, „ein wenig“, „ziemlich“, „stark und sehr stark werden“ respektive mit

▶ **Tab. 1** Stichprobenmerkmale der Sächsischen Längsschnittstudie und verwendete Instrumente.

Welle	Jahr	Teilnehmer	Alter (M, SD)	Geschlecht (n, %)	Verwendete Instrumente (neben G-Score)
13	1998	367	25,09 (0,45)	M: 166 (45,1) W: 202 (54,9)	Inanspruchnahme von Ärzten und Beratung
17	2003	417	30,05 (0,40)	M: 192 (45,8) W: 227 (54,2)	Inanspruchnahme von Ärzten und Beratung, gegenwärtiger Gesundheitszustand
18	2004	414	31,14 (0,40)	M: 192 (46,4) W: 222 (53,6)	Inanspruchnahme von Ärzten und Beratung
19	2005	377	32,11 (0,40)	M: 179 (46,6) W: 205 (53,4)	gegenwärtiger Gesundheitszustand
20	2006	393	33,17 (0,42)	M: 176 (45,5) W: 211 (54,5)	Inanspruchnahme von Ärzten und Beratung
24	2010	324	37,29 (0,47)	M: 148 (45,4) W: 178 (54,6)	Inanspruchnahme von Ärzten und Beratung, gegenwärtiger Gesundheitszustand
26	2012	350	39,23 (0,56)	M: 158 (45,1) W: 192 (54,9)	gegenwärtiger Gesundheitszustand, GBB-24, SCL-K-9, D-Score
27	2013	326	40,20 (0,45)	M: 145 (43,9) W: 185 (56,1)	Inanspruchnahme von Ärzten und Beratung, gegenwärtiger Gesundheitszustand, D-Score, SSS-8

0–4 kodiert. Die SCL-9-K ist eine Kurzform der Symptomcheckliste SCL-90 [21]. Validität und Reliabilität sind gegeben [18].

Der Gießener Beschwerdebogen (GBB-24) [9] ist ein valides und reliables Instrument zur Selbstbeurteilung mit 24 Items zur Erfassung körperlicher Beschwerden. Der Fragebogen umfasst die 4 Skalen Erschöpfung, Magenbeschwerden, Gliederschmerzen und Herzbeschwerden. Die Antwortmöglichkeiten lauten „nie“, „kaum“, „einigermaßen“, „erheblich“ sowie „stark“ (kodiert mit 0–4). Es wird weiterhin ein Gesamtwert (Summe) des Beschwerdedrucks bestimmt.

Die Somatic Symptom Scale (SSS-8) [8] ist ein Kurzfragebogen zur Selbstbeurteilung der Belastung durch somatische Symptome, z. B. „Bauchschmerzen oder Verdauungsbeschwerden“, „Schwindel“ oder „Schlafstörungen“. Die Befragten werden gebeten, ihre somatischen Symptome in den letzten 7 Tagen anhand einer 5-stufigen Ratingskala („gar nicht“ – „wenig“ – „mittel“ – „stark“ – „sehr stark“, kodiert mit 0–4) einzuschätzen. Der Summenscore kann zwischen 0 und 32 liegen.

In den Wellen 18 (2004), 20 (2006) und 24 (2010) wurde außerdem erfragt, ob (1) ein Hausarzt oder Allgemeinmediziner, (2) ein Facharzt (Chirurg, Augenarzt, Orthopäde), (3) ein Neurologe/Psychiater, (4) ein Psychotherapeut/Psychologe, (5) Beratung/Selbsthilfegruppe/Pfarrer oder (6) Familienangehörige/Freunde/Bekanntete in Anspruch genommen wurden. In der Welle 27 (2013) wurde gefragt, ob und wenn ja, wie oft (1) ein Hausarzt oder Allgemeinmediziner, (2) ein Facharzt oder (3) ein Psychotherapeut in Anspruch genommen wurden.

Um Sensitivität und Spezifität des G-Scores zu beurteilen und einen geeigneten Cut-Off-Wert zu bestimmen, wird für die Daten der Welle 27 ein Vergleich mit der SSS-8 [8] einschließlich einer Betrachtung der ROC-Kurven und des Youden-Index vorgenommen. Es wird eine logistische Regression mit dem G-Score der Welle 19 (2005) als Prädiktor und der SSS-8 [8] der Welle 27 (2013) als Indikatorvariable berechnet.

Reliabilität

Die Reliabilität wird mittels interner Konsistenz (Cronbachs Alpha) und Retest-Reliabilität im Einjahresabstand mittels Guttman-Split-Half-Koeffizienten anstelle der Spearman-Brown-Korrektur berechnet, da in sämtlichen Erhebungswellen die Testhälften starke Unterschiede zwischen den Varianzen aufweisen. So zeigt Testhälfte 1, bestehend aus Item 1 und 2, durchweg größere Varianzen als Testhälfte 2, bestehend aus Item 3 und 4 [16].

Ergebnisse

Itemanalyse

Die Histogramme aus den Wellen 13, 17, 19, 24 und 27 zeigen vornehmlich rechtsschiefe, eingipflige Verteilungen. Am häufigsten wurde die Antwort „nein, nie“ (0) gegeben; die Antwort „ja, häufig“ (3) am seltensten. Die Minima und Maxima in ▶ **Tab. 2** zeigen, dass von der Skala vollständig Gebrauch gemacht wurde. Die part-whole-korrigierten Trennschärfen betragen im Mittel $r = 0,47$ (Item 1), $r = 0,52$ (Item 2), $r = 0,49$ (Item 3) und $r = 0,36$ (Item 4). Die arithmetischen Mittel der Inter-Item-Korrelationen betragen $r = 0,318$ (Welle 13), $r = 0,268$ (Welle 17), $r = 0,350$ (Welle 19), $r = 0,382$ (Welle 24), $r = 0,394$ (Welle 27) und im Mittel aller Wellen $r = 0,342$ ($SD = 0,002$).

Objektivität

Durch das standardisierte Format des Fragebogens mit einer durchgehend verbal verankerten Rating-Skala, die festgelegte Zuordnung von numerischen Werten zu den Antworten auf der Rating-Skala und einer definierten Verrechnungsvorschrift sind die Durchführungs- und die Auswertungsobjektivität sehr gut gegeben. Eine gute Interpretationsobjektivität ist durch die Bestimmung des Cut-Off-Wertes gegeben. Insgesamt ist die Objektivität damit als sehr gut zu bewerten.

Reliabilität

Die Interne Konsistenz erreicht Werte zwischen $\alpha = 0,589$ und $\alpha = 0,769$, im Mittel $\alpha = 0,662$. Die Werte für die Einjahres-Retest-Reliabilität liegen zwischen $r_{tt} = 0,516$ (Welle 13/17) und $r_{tt} = 0,761$ (Welle 24/25) (▶ **Tab. 3**). Im Mittel liegt die Einjahres-Retest-Reliabilität bei $r_{tt} = 0,656$.

Konstruktvalidität

Die Korrelationen mit anderen Konstrukten liegen zwischen $\rho = 0,378$ und $\rho = 0,686$. Die Korrelationen mit dem Aufsuchen verschiedener Ärzte und anderer Personengruppen fallen mit höchstens $\rho = 0,412$ geringer aus. Alle Korrelationen sind bis auf 2 Ausnahmen signifikant (▶ **Tab. 4**).

Die ROC-Kurve hat eine Fläche von $AUC = 0,824$. Für einen Cut-Off-Wert des G-Scores bei 3,5 (4) ergibt sich eine Sensitivität von 0,775 und eine Falsch-Positiv-Rate ($1 - \text{Spezifität}$) von 0,279 sowie den höchsten Youden-Index von 0,496.

In allen Erhebungswellen, mit Ausnahme in Welle 19 (2005), berichteten Frauen im Mittel signifikant stärkere Beschwerden als Männer (▶ **Tab. 5**).

Prädiktive Validität

Es ergibt sich für die logistische Regression im Omnibus-Test der Modellkoeffizienten zur Überprüfung der Modellgüte eine signifikante Veränderung im $-2LL$, verglichen mit dem Nullmodell ($\Delta -2LL = 54,15$; $p = < 0,01$). Nach Cox und Snell kann der G-Score 17,0% und nach Nagelkerke 23,5% der Gesamtvarianz der körperlichen Befindlichkeit aufklären. Sowohl das Bestimmtheitsmaß nach Cox und Snell als auch nach Nagelkerke – welches einer normierten Version des Bestimmtheitsmaßes nach Cox und Snell entspricht – stellen ein dem R^2 in linearen Regressionen ähnliches Instrument zur Einschätzung der Anpassungsgüte dar. Es ist zu beachten, dass das Bestimmtheitsmaß nach Cox und Snell den Wert 1 nicht annehmen kann, was in der Interpretation zu berücksichtigen ist. In der Klassifizierungstabelle zeigt sich für nach SSS-8 als unauffällig eingestufte Personen eine Sensitivität von 84,4%, für als auffällig eingestufte eine Spezifität von 41,4%. Der G-Score sagt also 84,4% der nach SSS-8 unauffälligen und 41,4% der nach SSS-8 auffälligen Personen korrekt vorher. Die Gesamttrefferwahrscheinlichkeit liegt bei 69,8%. Das Odds-Ratio für den G-Score betrug 1,529 [95% KI (1,345; 1,738)]. Es unterscheidet sich signifikant von Null (▶ **Tab. 6**). Mit jeder Steigerung des Wertes im G-Score steigt also die Wahrscheinlichkeit, in der SSS-8 als auffällig klassifiziert zu werden, um das 1,529-fache an.

Diskussion

Mit dem G-Score wird ein sehr kurzes und somit ökonomisches Instrument zur Erfassung von körperlichen Beschwerden vorgestellt.

► **Tab. 2** Item-Kennwerte des G-Scores.

Welle/Item	M	SD	Min	Max	Schiefe	Exzess	Schwierigkeits-index	korrigierte part-whole-Trennschärfen
13,1	0,723	0,892	0	3	0,968	-0,110	0,723	0,475
13,2	1,190	0,989	0	3	0,257	-1,051	1,190	0,548
13,3	0,870	1,004	0	3	0,768	-0,683	0,870	0,541
13,4	0,229	0,629	0	3	3,028	8,876	0,229	0,240
17,1	0,549	0,822	0	3	1,445	1,298	0,549	0,379
17,2	1,065	0,965	0	3	0,419	-0,920	1,065	0,434
17,3	0,969	1,006	0	3	0,604	-0,860	0,969	0,375
17,4	0,257	0,623	0	3	2,691	7,113	0,257	0,320
19,1	0,692	0,885	0	3	1,060	0,123	0,692	0,497
19,2	1,117	0,909	0	3	0,215	-0,991	1,117	0,522
19,3	0,960	0,903	0	3	0,580	-0,560	0,960	0,462
19,4	0,257	0,574	0	3	2,467	6,266	0,257	0,394
24,1	0,699	0,872	0	3	1,048	0,172	0,699	0,537
24,2	1,132	0,958	0	3	0,389	-0,842	1,132	0,526
24,3	1,012	0,947	0	3	0,613	-0,564	1,012	0,557
24,4	0,314	0,643	0	3	2,123	4,021	0,314	0,398
27,1	0,881	0,924	0	3	0,662	-0,645	0,881	0,477
27,2	1,257	0,950	0	3	0,308	-0,813	1,257	0,591
27,3	1,171	1,010	0	3	0,388	-0,965	1,171	0,533
27,4	0,413	0,721	0	3	1,620	1,633	0,413	0,453
								Cronbachs Alpha
G-Score 13	3,020	2,52	0	3	0,889	0,254	-	0,666
G-Score 17	2,840	2,32	0	12	0,905	0,238	-	0,589
G-Score 19	3,030	2,37	0	12	0,883	0,251	-	0,769
G-Score 24	3,150	2,53	0	12	0,915	0,270	-	0,712
G-Score 27	3,720	2,67	0	12	0,760	0,269	-	0,720

► **Tab. 3** Retest-Reliabilitäten des G-Scores (Intervall 1 Jahr).

Wellen	rtt	Wellen	rtt	Wellen	rtt
12/13	0,574 **	19/20	0,621 **	23/24	0,722 **
13/17	0,516 **	20/21	0,657 **	24/25	0,761 **
17/18	0,589 **	21/22	0,665 **	25/26	0,742 **
18/19	0,601 **	22/23	0,725 **	26/27	0,694 **

* $p < 0,05$, ** $p < 0,01$

Im Vergleich zu den Kurzfragenbogen GBB [9] und SSS-8 [8] ist dieser Fragebogen mit 4 Items nur halb so lang. Die Durchführung und die Auswertung benötigen weniger als eine Minute. Für Forschungsfragestellungen hat der Fragebogen daher eine hohe Nützlichkeit. So kann der G-Score erste Hinweise auf Beschwerdebilder bei den Befragten liefern. Für Studien, in denen das körperliche Beschwerdeerleben im Mittelpunkt des Interesses steht, empfehlen sich ausführlichere Instrumente, wie etwas die Langversion des GBB [9]. Bei

Untersuchungen mit zahlreichen verschiedenen Fragestellungen, bei denen das subjektive Gesundheitsbefinden nur eines unter vielen interessierenden Konstrukten darstellt, ist dieses Screeninginstrument eine ökonomische und flexibel einsetzbare Alternative. In Online-Surveys könnte nach Erreichen des Cut-Offs (Sprungregel) eine detailliertere Evaluation der körperlichen Gesundheit mit einem ausführlicheren Instrument erfolgen.

Die Objektivität des G-Scores ist durch die Standardisierung von Instruktion, Fragenwortlaut, Antwortkategorien und Auswertung gegeben.

Im Mittel liegt die Inter-Item-Korrelation bei $r = 0,342$ und damit im empfohlenen Bereich [22]. Die Trennschärfen erreichen den hohen Bereich [22]. Insgesamt zeigt die Itemanalyse, dass die Items in ihrer Item-Schwierigkeit, Trennschärfe und Homogenität durchaus im erwartungskonformen Bereich liegen und Abweichungen inhaltlich und statistisch plausibel zu erklären sind. So weisen Items, die eine hohe Item-Schwierigkeit haben und gleichzeitig eine geringe Standardabweichung, eher niedrigere Trennschärfen auf [16].

Die Inhaltsvalidität ist als gut einzuschätzen, da die 4 Items eindeutig Aspekte des a posteriori definierten Konstrukts „körperliches Gesundheitsbefinden“ abbilden. Die Korrelationen des G-Scores mit

► Tab. 4 Korrelationen des G-Scores mit anderen Verfahren (Spearman-Rang-Korrelationen).

G-Score Welle	HADS-D Depressivität	HADS-D Angst	GBB-24	D-Score	SCL-K-9	SSS-8	gegenwärtiger Gesundheitszustand	Inanspruchnahme: Hausarzt/ Psychotherapeut
18	-	-	-	-	-	-	0,402 **	0,170 ** / 0,126 / 0,127 *
20	-	-	-	-	-	-	0,389 **	0,140 ** / 0,078 / 0,116 *
24	-	-	-	-	-	-	0,500 **	0,308 ** / 0,250 ** / 0,230 **
26	0,421 **	0,611 **	0,685 **	0,449 **	0,568 **	-	0,441 **	-
27	-	-	-	0,517 **	-	0,686 **	0,392 **	0,227 ** / 0,234 ** / 0,267 **

* p < 0,05, ** p < 0,01

dem Gießener Beschwerdebogen, der Somatic-Symptom-Scale-8, dem allgemeinen Gesundheitszustand und mit der Häufigkeit der Arztbesuche liegen erwartungsgemäß im mittleren und hohen positiven Bereich. Frauen berichteten mehr Symptome als Männer.

Die Korrelationen mit der Angstskala sowie der Depressionsskala der HADS-D, mit der SCL-K-9, dem D-Score und mit der Häufigkeit der Inanspruchnahme von Psychotherapeuten und von Beratungsangeboten, Familie und Freunden sind ebenfalls signifikant positiv. Dies ist inhaltlich allerdings begründbar, da körperliche Beschwerden durch psychische Beschwerden verursacht sein können [23]. Ebenfalls können körperliche Erkrankungen zu psychischen Komorbiditäten führen [24].

Die berichteten Beschwerden, die Inanspruchnahme medizinischer Leistungen sowie deren Korrelationen sind in dieser Studie jedoch relativ gering ausgeprägt. Dies kann u. a. auf das Alter der Stichprobe (M 2013 = 40,2) und dem daher zu vermutenden relativ guten Gesundheitszustand zurückgeführt werden. Die gestellte Frage nach der Inanspruchnahme medizinischer Leistungen schließt Vorsorgeuntersuchungen verschiedenster Art nicht aus, welche nicht mit Beschwerden in Zusammenhang stehen müssen. Weiterhin könnten Inanspruchnahmen ärztlicher Dienstleistungen aufgrund von Beschwerden erfolgt sein, welche nicht im G-Score abgefragt werden. Die retrospektive Erfassung von Arztbesuchen in der Vergangenheit kann insbesondere bei befundlosen Untersuchungen deutlichen Vergessensprozessen unterliegen. Die geringen Zusammenhänge zwischen Inanspruchnahme/körperlichem Befund und subjektivem Beschwerdeerleben sind aus vielen Studien bekannt [6].

Die Konstruktdefinition der körperlichen Gesundheit legt keinen eindeutigen Punkt fest, an dem die Grenze zwischen „gesund“ und „krank“ überschritten ist. Schlussendlich bleibt es offen zu bestimmen, ab wann eine Person als körperlich auffällig gilt. Dies könnte festgelegt sein bei Beantwortung eines einzigen Items mit „ja, häufig“ oder jedoch auch bei Beantwortung aller 4 Items mit „ja, selten“. Zudem verfügen die Antwortoptionen möglicherweise nicht über eine ausreichende inhaltliche Trennschärfe. Aus diesem Grund sollten zukünftig weitere Cut-Off-Wert-Bestimmungen mit anderen Maßen vorgenommen werden.

Auf der Basis der Werte des G-Scores im Jahr 1999 können für 2013 in der SSS-8 zu 84,4% die Einstufung als unauffällig vorherzusagen und die Einstufung als auffällig zu 41,4%. In der Konsequenz ist eher nicht zu empfehlen, den G-Score als Verfahren zu nutzen, um mögliche körperliche Gesundheitsbeschwerden über lange Zeiträume vorherzusagen. Der G-Score ist unter der Annahme, dass körperliche Beschwerden meist passagerer Natur sind, ein änderungssensitives Instrument. Zusammenfassend sind die Ergebnisse der logistischen Regressionsanalyse als Indikator für eine gute prädiktive Validität des G-Scores anzusehen. Es ist zu empfehlen, in späteren Arbeiten die prädiktive Validität anhand weiterer diagnostischer Instrumente erneut zu überprüfen.

Die mittlere interne Konsistenz (Cronbachs Alpha) ist mit $\alpha = 0,662$ als niedrig einzuschätzen. Diese ist durch die Kürze des Instruments mit nur 4 Items erklärbar. Niedrige Werte der internen Konsistenzen können auch durch geringe Varianzen bedingt sein [25]. Im Mittel liegt die Einjahres-Retest-Reliabilität bei $r_{tt} = 0,656$ und ist damit als niedrig einzuschätzen. Die Einjahres-Retest-Reliabilität im niedrigen Bereich ist inhaltlich plausibel begründbar

► **Tab. 5** Geschlechtsunterschiede im G-Score (M, SD, t-Test).

Welle	Frauen		Männer		t-Test		
	M	SD	M	SD	t	df	p
13	3,39	2,69	2,57	2,22	-3,125	365	0,002
17	3,28	2,26	2,31	2,29	-4,370	414	0,000
18	3,72	2,32	2,65	2,14	-4,791	402	0,000
19	3,24	2,42	2,79	2,29	-1,842	374	0,066
20	3,54	2,43	2,67	2,25	-3,587	377	0,000
24	3,63	2,71	2,59	2,18	-3,745	322	0,000
26	3,79	2,73	3,00	2,28	-2,095	348	0,004
27	4,29	2,78	2,97	2,35	-4,567	324	0,00

► **Tab. 6** Prädiktive Validität G-Score (1999)→SSS-8 (2013).

	B	SE	Wald Z	df.	Sig.	OR	OR unt.	OR ob.
G-Score	0,425	0,650	42,170	1,000	0,000	1,529	1,345	1,738
Konst.	-2,041	0,258	62,507	1,000	0,000	0,130		

durch die hohe Änderungssensitivität des erfassten Konstrukts körperliches Gesundheitsbefinden.

Beim G-Score handelt es sich um eine Eigenkonstruktion, entwickelt zum Einsatz als sehr kurzes Screening körperlicher Beschwerden in der Sächsischen Längsschnittstudie. Es könnte zukünftig eine konkretere verbale Verankerung als Antwortvorgabe gewählt werden, die sowohl die konkrete Häufigkeit des Auftretens als auch die Dauer der Beschwerden berücksichtigt. Außerdem kritisch zu betrachten ist in diesem Zusammenhang, dass der G-Score bislang nur bei einer altershomogenen Stichprobe von aus Ostdeutschland stammenden Personen eingesetzt wurde. Die vier abgefragten Bereiche decken inhaltlich nicht das gesamte Spektrum möglicher Körperbeschwerden ab.

Zur Prüfung der hier vorgestellten Ergebnisse ist es zukünftig geplant, weitere psychometrische Untersuchungen des G-Scores an einer bevölkerungsrepräsentativen Stichprobe und mit weiteren objektiven Vergleichsdaten durchzuführen. Größere Stichproben ermöglichten dann auch die Prüfung soziodemografischer Einflüsse (wie Alter, Bildungsstand usw.) auf den G-Score.

FAZIT FÜR DIE PRAXIS

Die psychometrischen Eigenschaften des Kurzfragebogens G-Score zur Erfassung subjektiv wahrgenommener körperlicher Beschwerden zeigen, gemessen an seiner Kürze, ein zufriedenstellendes Potenzial für den Einsatz als Screeninginstrument im Forschungsbereich. Ein Einsatz in der Individualdiagnostik und zur Veränderungsmessung sollte aufgrund der bislang vorhandenen Daten unter Vorbehalt erfolgen.

Interessenkonflikt

Die Autoren geben an, dass kein Interessenkonflikt besteht.

Literatur

- [1] Statistisches Bundesamt., (Hrsg.). Gesundheitsbericht für Deutschland. Stuttgart: Verlag Metzler-Poeschel; 1998
- [2] Rattay P, Butschalowsky H, Rommel A et al. Inanspruchnahme der ambulanten und stationären medizinischen Versorgung in Deutschland. Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013; 56: 832–844
- [3] Radoschewski M. Gesundheitsbezogene Lebensqualität – Konzepte und Maße. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2000; 43: 165–189
- [4] Verbraucherzentrale Bundesverband., (Hrsg.). Abbau von Über-, Unter- und Fehlversorgung. Den Blick auf den Bedarf der Patienten richten. Positionspapier des Verbraucherzentrale Bundesverbandes zum Eckpunktepapier zum Versorgungsgesetz (05.05.2011). Im Internet: http://www.vzbv.de/sites/default/files/downloads/versorgungsgesetz_positionspapier_05_05_2011.pdf (17.05.2017)
- [5] Hurrelmann K. Gesundheitssoziologie. Weinheim: Juventa; 2000
- [6] Brähler E, Schumacher J. Befund und Befinden: Psychologische Aspekte körperlicher Beschwerden. In: Brähler E, Strauß B, (Hrsg.). Handlungsfelder in der Psychosozialen Medizin. Göttingen: Hogrefe; 2002: 208–241
- [7] Helmert U. Subjektive Einschätzung der Gesundheit und Mortalitätsentwicklung. Gesundheitswesen 2003; 65: 47–54
- [8] Gierk B, Kohlmann S, Kroenke K et al. The somatic symptom scale-8 (SSS-8): A brief measure of somatic symptom burden. JAMA Intern Med 2014; 174: 399–407

- [9] Brähler E, Hinz A, Scheer JW. GBB-24. Der Gießener Beschwerdebogen. Manual (3. Auflage)Göttingen: Hogrefe; 2008
- [10] Nationale Akademie der Wissenschaften Leopoldina, acatech – Deutsche Akademie der Technikwissenschaften, Union der deutschen Akademien der Wissenschaften. Wissenschaftliche und gesellschaftspolitische Bedeutung bevölkerungsweiter Längsschnittstudien. Halle (Saale): Deutsche Akademie der Naturforscher Leopoldina; 2016 https://www.leopoldina.org/uploads/tx_leopublication/2016_Stellungnahme_Laengsschnittstudien.pdf (abgerufen am 17.05.2017)
- [11] Kemper CJ, Brähler E, Zenger M (Hrsg.). Psychologische und sozialwissenschaftliche Kurzskalen. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2014
- [12] Berth H, Förster P, Brähler E et al. Arbeitslosigkeit und körperliche Gesundheit. Ausgewählte Ergebnisse einer Längsschnittstudie. In Fangerau H, Kessler S, (Hrsg.). Achtung und Missachtung in der Medizin. Anerkennung und Selbstkonstitution als Schlüsselkategorien zur Deutung von Krankheit und Armut. Freiburg: Alber; 2012: 215–232
- [13] Berth H, Brähler E, Zenger M et al. (Hrsg.). Gesichter der ostdeutschen Transformation. Die Teilnehmerinnen und Teilnehmer der Sächsischen Längsschnittstudie im Porträt. Gießen: Psychosozial-Verlag; 2015
- [14] Förster P. Junge Ostdeutsche auf der Suche nach der Freiheit. Eine systemübergreifende Längsschnittstudie zum politischen Mentalitätswandel vor und nach der Wende. Opladen: Leske + Budrich; 2002
- [15] Berth H, Förster P, Stöbel-Richter Y et al. Arbeitslosigkeit und psychische Belastung. Ergebnisse einer Längsschnittstudie 1991 bis 2004. *Z Med Psychol* 2006; 15: 111–116
- [16] Bühner M. Einführung in die Test- und Fragebogenkonstruktion. 3. Aufl.München: Pearson Studium; 2011
- [17] Herrmann C, Buss U, Snaith RP. Hospital Anxiety and Depression Scale – Deutsche Version. Ein Fragebogen zur Erfassung von Angst und Depressivität in der somatischen Medizin. Bern: Huber; 1995
- [18] Klaghofer R, Brähler E. Konstruktion und teststatistische Prüfung einer Kurzform der SCL-90-R. *Z Klin Psychol Psychiatr Psychother* 2001; 49: 115–124
- [19] Hinz A, Brähler E. Normative values for the hospital anxiety and depression scale (HADS) in the general German population. *J Psychosom Res* 2011; 71: 74–78
- [20] Feige T. D-Score – Bewertung eines Screeninginstrumentes zur Messung der Psychischen Befindlichkeit. Psychologische Diplomarbeit, Technische Universität Dresden; 2014
- [21] Franke GH. SCL-90-R. Symptom Checkliste von L.R. Derogatis – Deutsches Manual. 2. Aufl Göttingen: Beltz; 2002
- [22] Moosbrugger H, Kelava A. Testtheorie und Fragebogenkonstruktion. 2., aktual. u. überarb. Aufl. Berlin: Springer; 2012
- [23] Rief W. Somatisierungsstörung. In Linden M, Hautzinger M, (Hrsg.). Verhaltenstherapiemanual. Heidelberg: Springer Medizin; 2015: 591–595
- [24] Jacobi F. Psychische Störungen bei Patienten mit körperlichen Erkrankungen in der Allgemeinbevölkerung. In Harter M, Baumeister H, Bengel J, (Hrsg.). Psychische Störungen bei körperlichen Erkrankungen. Heidelberg: Springer Medizin; 2007: 45–53
- [25] Leonhart R. Datenanalyse mit SPSS. Göttingen: Hogrefe; 2010