

Psychologische Rundschau

Herausgeber

Karl Christoph Klauer
Jens Bölte
Oliver Dickhäuser
Barbara Krahe
Klaus Moser

Offizielles Organ der Deutschen Gesellschaft für Psychologie (DGPs)

Themenheft

Gewissheit und Skepsis:
Wissenschaftskommunikation als
Forschungsthema der Psychologie

Gastherausgeber

Rainer Bromme und Dorothe Kienhues

TBS-TK-Rezension

Sorge- und Umgangsrechtliche Testbatterie (SURT)

Joseph Salzgeber, Gesellschaft für Wissenschaftliche Gerichts- und Rechtspsychologie (GWG), München

Johannes Bach, GWG

Michael Wiedemann, Sachverständiger in Berlin/Brandenburg

Allgemeine Informationen

Zur Diagnostik hinsichtlich der Thematik des Sorgerechts und Umgangsrechts existieren nur sehr wenige standardisierte Verfahren, welche den Anforderungen an einen psychometrischen Test genügen, dies gilt insbesondere für jüngere Kinder. Die Sorge- und Umgangsrechtliche Testbatterie (SURT) (Hommer, 2009) schließt diese Lücke. Sie erfasst im Moment einer familienrechtspsychologischen Begutachtung emotionale Beziehungen von Kindern zwischen vier und acht Jahren zu ihren Eltern über drei Untertests, die auch einzeln vorgegeben werden können: den projektiven Familienszenen-Test (PFST), dem semiprojektiven Entscheidungsfragebogen (SPEF) und dem nicht projektiv ansetzenden Test Elternwahrnehmungsunterschiede (EWU). Testergebnisse werden über Normen bewertet. Wieso das Verfahren projektiv bzw. semiprojektiv benannt wird, bleibt unklar.

Theoretische Grundlagen

Der Testautor geht davon aus, dass die Zuschreibungen von Eigenschaften und die Wahl zwischen den Eltern in speziellen Situationen im SPEF (ein Kind oder ein Tierkind entscheidet über die Nähe zu einem Elternteil) die emotionale Beziehung des Kindes zu den Eltern quantifizierbar abbildet. Der Test berücksichtigt die spezielle Situation in der familienrechtspsychologischen Begutachtung, nämlich die Ausgangssituation der Nullhypothese (die Bindung des Kindes zu den Eltern weicht nicht gesichert vom Mittelwert der Normstichprobe ab) und das zivilrechtliche Beweisprinzip des Überwiegens von Indizien für eine Seite der streitenden Parteien. Das generelle Prinzip der Auswertung und Interpretation ist der Vergleich von Differenzwerten zwischen den Eltern im jeweiligen Untertest, nämlich ob sich im Vergleich mit einer Normgruppe von Familien in Sorgerechtsverfahren diese Differenzen signifikant von den empirischen Mittelwerten der drei Untertests unterscheiden. Im PFST ist noch eine zweite Bewertung isoliert für einen Elternteil möglich.

Objektivität

Die Testinstruktion für die Durchführung des Verfahrens ist im Handbuch klar und nachvollziehbar dargestellt. Es werden wörtliche Instruktionen für den Testleiter vorgegeben sowie genaue Angaben bezüglich des Settings gemacht. Ausführungen zu besonderen Kompetenzen des Testleiters aufgrund des speziellen Einsatzgebietes wären angemessen gewesen. Wünschenswert wären auch Hinweise sowohl auf den Umgang mit Rückfragen der Kinder als auch auf mögliche Reaktionen darauf. Auch die Problematik der Verweigerungshaltung des Kindes in der Testsituation wird nicht thematisiert. Bezüglich der Auswertung der Ergebnisse liegen klare und nachvollziehbare Instruktionen vor, insofern ist die Auswertungsobjektivität gegeben. Hilfreich für die Interpretation der Ergebnisse sind die drei dargestellten Fallvignetten, anhand derer sich (proto)typische Fallkonstellationen nachvollziehen lassen. Detaillierte Informationen zur Interpretation der Ergebnisse werden jedoch nicht gegeben.

Normierung

Zur Gewinnung der Normstichprobe wurden einige Stichproben, welche unter anderem mit Hilfe von Diplomarbeiten gewonnen wurden, zusammengelegt. In diesem Zusammenhang wird darauf verwiesen, dass sich die Einzelstichproben hinsichtlich der Mittelwerte nicht signifikant unterscheiden. Es fehlt jedoch eine detaillierte Übersicht über die zentralen Kennwerte der unterschiedlichen Stichproben, die zur Normierung herangezogen wurden. Die Repräsentativität der Stichprobe wird im Handbuch angenommen, ihr Nachweis jedoch nicht geführt. Der Hinweis auf die Schwierigkeit der Datengewinnung im Kontext der gerichtlichen Begutachtung ist berechtigt, dennoch können die vorliegenden Daten keinen Anspruch auf Repräsentativität erheben bzw. der Nachweis fehlt. Des Weiteren fehlt eine Darstellung, ob es bedeutsame Unterschiede hinsichtlich der Testwerte einerseits von Kindern, die alle Untertests bearbeitet haben, und andererseits von denen, die nur einen oder zwei Untertests bearbeitet haben, gibt. Je nach Untertest lag die Größe der Stichprobe zwischen

$N = 392$ und $N = 896$, dabei in etwa gleich viele Jungen wie Mädchen und Kinder jünger als sieben bzw. älter als sechs Jahre. Für einzelne Werte liegen getrennte Normen für die Geschlechter vor. Wieso keine getrennten Altersnormen vorliegen, selbst wenn einzelne Gütekriterien wie die zur konvergenten Validität für zwei Altersgruppen berechnet wurden, wird nicht erläutert.

Zuverlässigkeit

Reliabilitäten wurden immer zu den Differenzen zwischen den Eltern berechnet. Zum Untertest PFST werden drei Werte, nämlich zu den drei Qualitäten „Abstoßend“ (PFST-A), „Gefährlich“ (PFST-G) und „Schön“ (PFST-S) angegeben. Zum Untertest EWU wurden die Reliabilitäten zur Gesamtpunktzahl und zum Präferenzwert (Häufigkeit der Bevorzugung von Vater oder Mutter zur jeweils gleichen Frage) und zum Untertest SPEF zum Gesamtwert aufgeführt. Hinsichtlich der Reliabilität des Verfahrens werden Angaben gemacht, welche sich für einige Subskalen im Bereich der Werte von .60 (EWU) und .73 (PFST-S-Diff) für Cronbachs Alpha bewegen. Hiermit wurden Werte im durchschnittlichen Bereich erzielt. Der Wert für PFST-A-Diff ist mit .34 problematisch, und es ist nicht nachvollziehbar, warum diese Signierung dennoch beibehalten wurde. Wie bereits zur Normierung angeführt wurde, ist nicht immer nachvollziehbar, welche der im Handbuch angeführten Untersuchungen (z. B. sieben Untersuchungen mit N zwischen 39 und 653 zum PFST; dann wieder die Formulierung einer ersten Teilstichprobe von $N = 559$ und einer Validierungsstichprobe von $N = 337$) zur Berechnung der Werte herangezogen wurden.

Gültigkeit

Bezüglich der Validität des Verfahrens werden Angaben zur konvergenten Validität in der Normierungsstichprobe gemacht. Des Weiteren wurde der SURT zur konvergenten Validierung mit dem in der familienrechtlichen Gutachtung häufig verwendete FRT-R und in einer weiteren Untersuchung zum FIT korreliert. Problematisch sind erneut die Werte für den PFST-A-Wert: Hier waren die Korrelationen extrem niedrig. Des Weiteren zeigten sich deutliche Unterschiede in den Korrelationen zwischen den Subtests des SURT und dem FRT-R, wenn zwischen den Altersgruppen 4–6 Jahre und 7–8 Jahre unterschieden wird. Dennoch waren die Korrelationen signifikant und wiesen in die gleiche Richtung (wurde der Vater/ die Mutter im SURT positiver dargestellt, geschah das genauso im FRT-R). Diskutiert wird, dass der FRT-R und der SURT nicht identisches messen. Das Ergebnis macht zumindest deutlich, dass diese Verfahren Ähnliches bei der Fragestellung zu Beziehungen eines Kindes zu den Eltern erfassen, und bestätigt den validen Einsatz im familien-

rechtlichen Kontext. Die Konstruktvalidität wird nur ansatzweise untersucht. Als divergentes Kriterium werden nur zwei Untertests einer Intelligenztestbatterie benutzt, und dies nur für eine kleine Teilstichprobe. Wie bereits zuvor angeführt wurde, ist nicht immer einfach nachzuvollziehen, welche Probanden aus welchen der angeführten Untersuchungen für die Berechnung der Gütekriterien herangezogen wurden, so auch zur Validität des Verfahrens.

Weitere Gütekriterien

Durch die distanzierte Vorgabe (Tiere, „ein“ Kind) und die Benennung zuerst eines und dann des anderen Elternteils im Wechsel wird versucht, suggestive Effekte zu minimieren. Für die Kinder ist das Material ansprechend, der Aufwand ist gering und der Konzentrationskompetenz des Kindes entsprechend. Bei der Durchführung des Tests im Kontext der familienpsychologischen Begutachtung können Störungen wie z. B. emotionale Erregung des Kindes oder Abbruch aufgrund einer zu hohen Belastung für das Kind auftreten. Es wird nicht thematisiert, wie darauf reagiert werden soll. Sollte das Kind zur Untersuchung einmal von einem und ein weiteres Mal vom anderen Elternteil gebracht werden? Wie soll auf Rückfragen der Kinder reagiert werden? Einige Ausführungen im Handbuch (7. Gütekriterien und Normierung) sind beim ersten Durchlesen nicht nachvollziehbar, sodass mehrfaches Lesen oder eigenes Nachlesen der Quellen notwendig ist.

Abschlussbewertung/Empfehlung

Mit dem SURT wurde für die familienpsychologische Begutachtung ein praktikables Diagnoseinstrument konstruiert, welches auch einer Prüfung nach den klassischen Testgütekriterien standhält. Seine Anwendung ermöglicht, aufgrund einer formulierten Theorie die Beziehungsqualität des zu untersuchenden Kindes zu seinen Eltern und diese als Vergleich zwischen den Eltern objektiv zu erfassen. Beim SURT handelt es sich eigentlich um drei hinsichtlich der quantitativen Auswertung voneinander unabhängige Tests. Man kann alle Tests anwenden oder jeden einzeln einsetzen. Im Handbuch sind Beispiele aufgeführt, bei denen der Testautor selbst nur zwei der drei Tests verwandt hat. In der gutachterlichen Praxis stellt insbesondere der Subtest EWU ein äußerst hilfreiches Instrument im Sinne einer adäquaten Weiterentwicklung und Alternative des FRT-R dar. Bei der Anwendung der EWU muss bei vier- bis fünfjährigen Kindern überprüft werden, ob das System der Smileys für die Kinder der Altersgruppe zu schwierig ist. Da die Begutachtung in einem sehr spezifischen Setting stattfindet, sind einige Problemlagen Folgen dieser spezifischen Situation. Hierzu zählt, dass es schwerlich möglich sein wird, eine repräsentative

Stichprobe zu gewinnen. Es wäre hilfreich gewesen, auf diese Problematik zu verweisen und diese zu diskutieren. Es fehlen zudem differenzierte Untersuchungen zur Reliabilität und zur prognostischen Validität des Verfahrens an Teilstichproben mit und ohne Bezug zum Anwendungsgebiet. Im Handbuch werden zwar bei der Normierung Teilstichproben erwähnt, in denen einzelne Subtests ausgefüllt wurden, es fehlt jedoch der explizite Hinweis, inwiefern es legitim ist, den SURT nicht als Ganzes durchzuführen. Aufgrund der im Handbuch angedeuteten unterschiedlichen Ergebnisse der Altersstichproben wäre es hilfreich gewesen, in den Anhang auch die Prozentränge nicht nur nach Geschlecht getrennt, sondern auch für die Altersgruppen aufzunehmen.

SURT	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung		X		
Objektivität		X		
Zuverlässigkeit			X	
Validität		X		

Diese Testrezension wurde im Auftrag des Testkuratoriums der Föderation deutscher Psychologeneinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2009, 2010) erstellt.

Testkuratorium. (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologeneinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie*, 34, 470–478.

Testkuratorium. (2010). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologeneinigungen. Revidierte Fassung vom 09. September 2009. *Psychologische Rundschau*, 61, 52–56.

Testinformationen

Hommers, W. (2009). *Sorge- und Umgangsrechtliche Testbatterie (SURT)*. Bern: Huber.

Bezugsquelle: Testzentrale Göttingen, Herbert-Quandt-Str. 4, 37081 Göttingen. Test komplett 227 €, 10 Protokollbogen 10,20 €, 10 Auswertebogen 10,20 €.

Bitte zitieren Sie diesen Artikel wie folgt: Salzgeber, J., Bach, J. & Wiedemann, M. (2017). TBS-TK Rezension: „Sorge- und Umgangsrechtliche Testbatterie (SURT)“. *Psychologische Rundschau*, 68, 235–237.

DOI: 10.1026/0033-3042/a000369