

DER SAARBRÜCKER PERSÖNLICHKEITSFRAGEBOGEN SPF (IRI) ZUR MESSUNG VON EMPATHIE

Psychometrische Evaluation
der deutschen Version
des Interpersonal Reactivity Index (Davis, 1980)

PD Dr. Christoph Paulus
FR Erziehungswissenschaft
Universität des Saarlandes, Saarbrücken

Der Interpersonal Reactivity Index (IRI) ist einer der am häufigsten in der Praxis eingesetzte Fragebogen zur Messung von Empathie. Die wenigen deutschen Übersetzungen weisen allerdings ähnliche testtheoretischen Schwächen auf wie das Original (vgl. dazu z.B. Enzmann, 1996; Beven, 2004; Lauterbach & Hosser, 2007). Nach einer mehrfachen Überarbeitung aufgrund faktorenanalytischer Ergebnisse und marginalen Umformulierungen in der Übersetzung konnten wir eine deutsche Version der IRI erstellen, die gute Kennwerte in den Bereichen der Reliabilität, der Validität und der Itemtrennschärfen aufweist. Der Saarbrücker Persönlichkeitsfragebogen SPF(IRI) ist auch online zu bearbeiten unter www.cpaulus.de. Durch das Weglassen aller negativ formulierten Items konnte die Auswertung erheblich verbessert werden, so dass nun eine einfachere Scorebildung möglich wurde.

The Interpersonality Reactivity Index (IRI) is one of the most used questionnaires in practice to the measurement of empathy. However, all German translations show similar test-theoretical weaknesses (e.g. Enzmann, 1996; Beven, 2004; Lauterbach & Hosser, in 2007). After a multiple reworking on account of factor-analytic results and marginal changes in the translation we could provide a German version of the IRI which shows good values in the reliability, factorial validity and the item analyses. The Saarbrücken personality questionnaire SPF (IRI) is given also to be worked on-line. The evaluation could be improved by leaving out all negatively formulated Items considerably, so that now an easier score education became possible.

Einleitung

Empathie als „the tendency to apprehend another person’s condition or state of mind” (Johnson, Cheeks & Smither, 1983) oder etwas enger gefasst “the ability to understand and share in another’s emotional state or context” (Cohen & Strayer, 1996) verstanden, wird inzwischen nach langen Diskussionen als ein Konzept verstanden, das sowohl affektive als auch kognitive Komponenten beinhaltet. Der Interpersonal Reactivity Index (IRI) von Davis (1980, 1983) berücksichtigte diesen mehrdimensionalen Ansatz schon recht früh. Er unterteilte bei der Konstruktion des IRI Empathie in vier zueinander in Beziehung stehende Subgruppen: *perspective taking* (PT), *fantasy* (FS), *empathic concern* (EC) und *personal distress* (PD). PT misst dabei die Fähigkeit, spontan eine Sache aus der psychologischen Perspektive eines Anderen sehen zu können; die Tendenz, diese Fähigkeit einzusetzen, geht nach Davis einher mit besserer sozialer Akzeptanz und einem daraus folgenden höheren Selbstwertgefühl. Die *fantasy scale* (FS) erfasst die Tendenz, sich in die Gefühlswelt von Figuren in Romanen oder Filmen zu versetzen und ist insofern eher unter emotionalen Aspekten zu betrachten (Davis & Franzoi, 1991). Dabei wird keine Beziehung zu interpersonalen Fähigkeiten erwartet, insbesondere nicht zum Konstrukt Selbstwert. Die FS stellt allerdings auch ein Maß zur Stärke von Emotionalität dar. Die Original-FS enthält drei Items aus der *fantasy-empathy-Scale* von Stotland et al. (1978); dort konnte gezeigt werden, dass höhere Fantasy-Werte mit höheren Emotionsstärken („greater physical arousal“) einhergingen. Weiterhin wird von Davis (ebd.) keine Beziehung zwischen der FS und Messung zur Sensitivität gegenüber Anderen erwartet. Die restlichen beiden Subskalen stellen Operationalisierungen von typischen emotionalen Reaktionen eines Beobachters dar: Die EC-Skala dient zur Messung fremd-orientierter Gefühle wie Mitleid oder Sorge um Personen in Not. Bei dieser Skala ist theoretisch nicht klar, welcher Zusammenhang zu anderen Maßen der Emotionalität oder auch Sozialibilität besteht. Infolgedessen kann auch nicht ohne weiteres davon gesprochen werden, dass hohe Empathiefähigkeit mit hohem Selbstwert einhergeht (ebd.). Wenn auch die direkte Beziehung zwischen EC und sozialem Handeln weiterhin unklar bleibt, so konnte z. B. Paulus (1997) zeigen, dass insbesondere Empathiefähigkeit in Hilfesituationen ausschlaggebend werden kann, in denen das „Opfer“ keine Hilfe möchte, sie sogar explizit ablehnt. Die PD-Skala dagegen soll eigenfokussierte Gefühle wie Unruhe oder Unwohlsein in engen interpersonalen Situationen messen. Damit ist sie klar negativ korreliert mit Maßen, die soziale Fertigkeiten messen. Die PD-Skala wird von manchen Autoren (z.B. Baron-Cohen & Wheelwright, 2004) auch als ein Maß zur Emotionsregulation interpretiert. Ein Gesamtscore zur Bestimmung allgemeiner Empathie ist im Original von Davis nicht vorgesehen, da die Komponenten, die gemessen werden, als weitgehend unabhängig von einander angesehen werden (Rogers et al., 2007). Damit enthält der IRI drei emotionale (EC, FS und PD) und einen kognitiven Empathie-Faktor (PT).

Der Interpersonal Reactivity Index (IRI) ist einer der am häufigsten in der Praxis eingesetzte Fragebogen zur Messung von Empathie. Gut ein Drittel der in der Metaanalyse von Jolliffe & Farrington (2004) einbezogenen Studien zum Zusammenhang zwischen Empathie und Gewalt benutzten den IRI oder Teile daraus zur Empathiemessung, obwohl die psychometrischen Daten des Fragebogens, vor allem in der deutschen Übersetzung von Enzmann (1996), nicht immer als ausreichend angesehen wurden. So berichten Beven et al. (2004) und Ireland (1999) von unzureichenden Reliabilitätswerten; Curwen (2003) vermutet ebenfalls einen negativen Einfluss von sozialer Erwünschtheit auf die Zuverlässigkeit. Sogar Davis (1996) selbst spricht davon, dass insbesondere bei Studien an aggressiven Pbn die *fantasy scale* nicht aussagekräftig sein könnte. In der Studie von Dziobek et al. (2005), die sich mit der Frage beschäftigten, ob Menschen mit der Fähigkeit, psychische Signale ihres Gegenüber besser erkennen zu können als andere, dies vielleicht deshalb können, weil sie empathischer waren, konnte lediglich ein

hoch-signifikanter Effekt auf der FS im Verhältnis zu ihrer Kontrollgruppe nachgewiesen werden, auf allen anderen drei Skalen des IRI zeigten sich dagegen keine Unterschiede.

Lauterbach & Hosser (2007) konnten zeigen, dass in einer von ihnen anhand der Daten von 893 Vpn durchgeführten Faktorenanalyse die beiden Skalen FS und PD auf dem gleichen Faktor luden. Alle im Original negativ formulierten Items luden auf einem gemeinsamen Faktor, obwohl sie aus unterschiedlichen Subtests stammten. Auch eine Veränderung von einer Vier- zu einer Drei- oder Fünf-Faktoren-Lösung erbrachte kein besseres Bild. Die immer wieder auftauchenden Doppelladungen mehrerer Items verhinderten ebenfalls eine saubere Faktorenlösung. Erst nach dem Ausschluss der negativen Items war es annähernd möglich, eine klare Struktur zu erreichen. Einzelne Doppelladungen ließen sich aber auch hiermit nicht vermeiden. Pulos, Elison & Lennon (2004) konnten zeigen, dass neben der Vier-Faktoren-Struktur zwei übergeordnete Faktoren extrahiert werden können, die sie „general empathy“ bzw. „emotional control“ nannten. Während der erste der beiden Faktoren auch theoretisch mit Empathie in starkem Zusammenhang stand, blieb die Bedeutung und Interpretation des zweiten Faktors (zusammengesetzt aus PT und PD) unklar. Siu & Shek (2005) konnten in ihrer chinesischen Stichprobe lediglich drei Faktoren extrahieren: „personal distress“, „fantasy – empathy“ und „cognitive and emotional empathy“.

Die Reliabilitätswerte werden dagegen in allen Studien als „gut“ angegeben, liegen allerdings nur zwischen .63 und .77 (Davis, 1983; Christopher et al., 1993; Jolliffe et al., 2004; Lauterbach et al., 2007).

Im Folgenden sollen nun die Analysen zur Faktorenstruktur, zur Reliabilität und zur Validität zusammen mit einigen Itemanalysen dargestellt werden, die zeigen werden, dass die hier vorgestellte Übersetzung des IRI – in der deutschen Fassung als „Saarbrücker Persönlichkeitsfragen SPF (IRI)“ bezeichnet – die bisher geschilderten Probleme so nicht mehr aufweist.

Methode

Stichprobe

Die folgenden Analysen beruhen auf einem Datensatz von 339 Vpn. Der Altersdurchschnitt lag bei 26,64 Jahren ($s = 8,84$ J.) und streut zwischen 19 und 64 Jahren. 53,1% der Pbn waren männlich.

Finden einer Faktorenstruktur

Die Vollversion der deutschen Übersetzung ergibt, ähnlich wie bei Lauterbach et al. (2007), ein uneinheitliches Bild. Der Screeplot legt keine eindeutige Faktorenzahl nahe, die Eigenwerte liegen bis zur Faktorzahl 8 über 1,00. Allerdings beträgt der Zuwachs an Varianzaufklärung ab Faktor 6 bis zum theoretischen 8. Faktor nur noch knapp 5%, so dass als erstes eine 6 Faktoren-Struktur getestet wurde. Das Ergebnis der ersten FA zeigte schon eine recht brauchbare Struktur, in der die theoretisch angenommen vier Faktoren schon erkennbar waren, jedoch mit einer großen Anzahl von Doppelladungen. Im nächsten Schritt wurden sukzessive Items mit Doppelladungen oder zu niedrigen Ladungswerten gestrichen. Dabei zeigte sich auch, dass die negativ formulierten Items offensichtlich den Vpn Probleme zu bereiten schienen. Lauterbach & Hosser (2007) verweisen in diesem Zusammenhang darauf, dass sprachliche Fähigkeiten und auch Intelligenz der Vpn einen hohen Einfluss auf die negativ formulierten IRI-Items hatten und dass der Reduktion der Skala um diese Items wieder die ursprüngliche Faktorenstruktur hergestellt werden konnte. Ähnlich argumentieren Magazine et al. (1996) wenn sie vermuten,

dass die negative Formulierung von Aussagen einen hohen Anspruch an die kognitiven Fähigkeiten der Pbn voraussetzt.

n=339	FS	PD	PT	EC
12-f	,81			
15-f	,73			
07-f	,72			
02-f	,68			
03-d		,73		
06-d		,73		
13-d		,70		
08-d		,65		
10-p			,81	
04-p			,77	
16-p			,72	
14-p	,36		,51	
01-e				,78
05-e				,70
11-e		,45		,60
09-e	,40			,59

Nach Eliminierung der negativ formulierten Items 3, 4, 7, 12-15, 18 und 19 sowie der Items mit unklarer Faktorzuordnung (1, 11 und 27) ergab sich eine deutlichere Faktorenstruktur (vgl. Tabelle 1). Betrachtet man die Items mit Doppelladungen genauer, so zeigt sich, dass diese bei den Items Nr. 14 („Wenn mir das Verhalten eines anderen komisch vorkommt, versuche ich mich für eine Weile in seine Lage zu versetzen.“), Item Nr. 11 („Ich würde mich selbst als eine ziemlich weichherzige Person bezeichnen.“)

n=819	FS	PD	PT	EC
12-f	,79			
07-f	,76			
15-f	,73			
02-f	,64			
06-d		,84		
08-d		,83		
03-d		,60		
13-d		,52		
10-p			,78	
04-p			,77	
16-p			,67	
14-p			,66	
11-e				,72
01-e				,72
05-e				,65
09-e				,52

bzw. Item 09 („Oft berühren mich Dinge, die ich nur beobachte“) theoretisch erklärt werden können.

Tabellen 1: Rotierte Faktormatrix nach Kürzung der unklaren Items¹ (links: n=339; (Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Dargestellt werden nur Ladungen >.35)

Der Eindruck, „eine weichherzige Person zu sein“, entsteht subjektiv gesehen ja besonders dann, wenn man in distress-auslösende Situationen gerät und das insbesondere natürlich bei empathischen Personen; die angesprochene Beobachtung in den Items 9 bzw. 14 suggeriert durch die Wortbedeutung eine Ähnlichkeit mit der FS, die auf visuelle Anreize als Auslöser ausgerichtet ist. Es lassen sich also die vier theoretischen Empathiefaktoren EC, PT, FS und PD auch in einer verkürzten Fassung eindeutig nachweisen. Damit wird die Skala im Gegensatz zum 28-Items-Original faktoriell klarer. Die so extrahierten Faktoren lassen sich nach Vergrößerung der Stichprobe auf inzwischen 819 Vpn immer noch replizieren, wobei die Doppelladungen inzwischen verschwunden sind (s. Tabelle rechts).

Reliabilität

Als nächstes stellt sich die Frage, ob sich bei dieser gekürzten Version die Reliabilitätskennwerte weiterhin in den bisher bekannten Bereichen befinden. Davis (1980) beschreibt Cronbach-Alpha-Werte zwischen .71 und .78 bei gleichzeitig sehr hohen Retest-Werten zwischen .62 und .80 für einen Zeitraum von 8-10 Wochen (Davis & Franzoi, 1991). In Tabelle 2 sind verschiedene in der Literatur beschriebene Reliabilitätskennwerte im Vergleich aufgeführt. Auf Grund der Daten ergibt sich ein Cronbach Alpha von .78, das damit trotz oder gerade wegen (?) der geringeren Itemzahl weiterhin sehr zufriedenstellend ist. Auch der split-half-Koeffizient mit anschließender Spearman-Brown-Korrektur weist mit $r_{tt} = .80$ einen sehr guten Wert auf. Insgesamt gesehen kann die Reliabilität also durch die Reduktion der

¹ In der ursprünglichen Fassung des IRI sind die Subtests mit jeweils 7 Items vertreten, von denen nach der Faktorenanalyse noch 4-5 pro Faktor übrig blieben. Aus Gründen der Gleichgewichtigkeit wurden die Items mit der jeweils schwächsten Faktorladung noch gestrichen, so dass sich eine ausgewogene 16-Item-Skala ergab.

Itemzahl also verbessert werden. Die einzelnen Subtests sind dabei in sich sehr konsistent, wie die folgende Tabelle zeigt:

Sub-test	SPF (IRI)	Christopher et al. (1993)	Lamsfuss et al. (1990)	Lauterbach & Hosser (2007)
EC	.71	.76	.74	.77
PT	.71	.74	.56	.77
FS	.74	(nicht gemessen)	.75	.66
PD	.66	.70	.73	.63

Tabelle 2: Cronbach-Alpha-Werte der Subtests des SPF(IRI)

Die Korrelation der Subtests untereinander bleibt mäßig hoch, allerdings in einigen Fällen höher als bei Davis (1980) in seinen Samples beschrieben. Die dort genannten Korrelationen werden (gemittelt) in der folgenden Tabelle in Klammern angegeben:

		FS	PT	PD
C	E	,38 (,33)	,32 (,33)	,26 (,09)
	F		,28 (,12)	,18 (,11)
T	P			,05 (-,25)

Tabelle 3: Interkorrelationen der Subtests (Korrelationen über ,10 sind signifikant auf $p < ,01$)

Dabei muss ein besonderes Augenmerk auf die PD-Skala gerichtet werden, denn die Ergebnisse der distress-Skala sind auch in anderen Arbeiten nicht unumstritten. So werden bei Beven et al. (2004) nur die anderen drei Skalen in deren Untersuchungen berücksichtigt, da die distress-Skala keine eindeutigen Bezüge zu ihrer Stichprobe herstellen konnte. Lauterbach & Hosser (2007) berichten ebenfalls von einer unklaren Zuordnung der PD-Skala, die in deren Studie zusammen mit der FS einen gemeinsamen Faktor bildet bzw. mit PT korrelierte ($r = ,34; p < ,001$). Davis (1983), Davis & Franzoi (1991) oder auch D'Orazio (2004) fanden dagegen keine oder eine leicht negative Korrelation zwischen PT und PD. Hassenstab et al. (2007) vermuten hinter der PD-Skala ein weiterreichendes Konzept als das der reinen emotionalen Empathie. In Anlehnung an Baron-Cohen et al. (2004) erklären sie PD-Scores als die Fähigkeit, Emotionen zu regulieren und zu hemmen, wobei dies in keinem direkten Gegensatz zur allgemeinen empathischen Fähigkeit des Nachvollziehens von Emotionen stehen soll. In ihrer Studie erklären sie signifikant niedrigere PD-Werte bei ausgebildeten Therapeuten im Sinne einer Fähigkeit, in emotional angespannten (Therapie-)Situationen ihren distress-Level durch intellektuelle Prozesse auf moderatem Niveau halten zu können. Dadurch sollen sie sich besser von emotionalen Belastungen distanzieren können. Für diese Art der Affektregulation sprechen auch die Befunde u.a. von Lauterbach & Hosser (2007), die zeigten, dass es keine signifikanten Unterschiede zwischen einer Stichprobe von Mehrfach-Kriminellen und Nicht-Kriminellen auf dem PD-Faktor gab (alle andere IRI-Skalen unterschieden sich signifikant).

Itemanalysen

1.1.1 Trennschärfe und Schwierigkeit

Die Item-Score-Korrelationen zeigen gute Werte:

Item	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
01-e	51,17	52,75	,50	,74
05-e	51,07	53,61	,40	,74
09-e	51,33	51,78	,52	,73
11-e	51,37	52,55	,42	,74
02-f	50,61	55,12	,38	,75
07-f	51,68	52,89	,33	,75
12-f	50,86	54,36	,35	,75
15-f	50,96	52,91	,40	,74
03-d	51,55	54,74	,28	,75
06-d	51,76	54,34	,27	,75
08-d	52,02	54,05	,31	,75
13-d	52,14	57,39	,09	,77
04-p	50,93	54,89	,30	,75
10-p	51,10	54,40	,32	,75
14-p	51,70	53,03	,39	,74
16-p	51,87	51,97	,44	,74

Tabelle 4: Item-Subscore Korrelation des SPF(IRI)

Die Trennschärfen der einzelnen Items können überwiegend als brauchbar angesehen werden, lediglich Item 13-d liegt mit .09 sehr tief. Item 13² fällt auch bei der Betrachtung der Itemschwierigkeit in der Höhe etwas aus der allgemeinen Tendenz heraus. Möglicherweise ist die in diesem Item geschilderte Situation („Notfälle“) einerseits zu unspezifisch (welche Notfälle sind gemeint?), andererseits aber auch zu ungewöhnlich (wer erlebt schon häufig echte Notfälle?), um sie reliabel beantworten zu können. Deshalb schlagen wir vor, das Item umzuformulieren in „In heiklen Situationen neige ich dazu, die Kontrolle über mich zu verlieren.“

Mit der Itemschwierigkeit bezeichnet man bei Einstellungsmessungen das Antwortverhalten im Sinne der Fragebogenthematik. Im Falle der Empathiemessung wurden alle Items so codiert, dass eine hohe Zustimmung einen hohen Ausprägungsgrad im Sinne des Subtests (z.B. der Perspektivenübernahme) darstellt. Bestimmt wird die Itemschwierigkeit im klassischen Sinne nach der Formel:

$$P = \frac{\bar{x}_i - x_{\min}}{x_{\max} - x_{\min}}$$

Damit ergibt sich ein Maß für die „Leichtigkeit einer Zustimmung“ zur betreffenden Aussage. Liegt dieses Maß in extrem hohen bzw. niedrigen Bereichen, so kann davon ausgegangen werden, dass dieses Item nur schlecht geeignet ist, zwischen „guten“ (= empathischen) und „schlechten“ (unempathischen) Vpn zu differenzieren. Für den SPF(IRI) ergeben sich damit folgende Schwierigkeitsindices:

² Ich neige dazu, in Notfällen die Kontrolle über mich zu verlieren

I-tem Nr.	1	2	3	4	5	6	7	8	9	10
p	,69	,76	,56	,78	,77	,62	,45	,48	,74	,79

I-tem Nr.	1	2	3	4	5	6	C	E	F	P	P
p	,68	,66	,30	,63	,63	,65	,72	,63	,71	,49	0

Tabelle 5: Schwierigkeitsindizes der Items und der Subscores der SPF(IRI)

Externe Validität³

Zur Validierung haben wir weitere Fragebögen zur Erfassung von Empathie hinzugezogen, die bereits als valide gelten. Um dabei alle Subtests des SPF abdecken zu können, war es notwendig, die Items aus zwei verschiedenen Empathiefragebögen zu kombinieren, da kein vergleichbarer äquivalenter Fragebogen in deutscher Sprache existiert, der die Faktoren perspective taking, fantasy, empathic concern und personal distress beinhaltet.

E-Skala: Fragebogen zur Erfassung von Empathie (Silbereisen & Schulz, 1977)

Laut Silbereisen & Schulz (1977) ist Empathie die Bereitschaft und Fähigkeit eines Individuums, Emotionen über das Ausdrucksverhalten anderer Personen nachzuempfinden und durch die Zuordnung zu situativen Hinweisreizen zu begreifen. Sie ist gekennzeichnet durch die Wahrnehmung affektiver Signale bei anderen Individuen, die Wahrnehmung dafür maßgeblicher situativer Hinweisreize und die Wahrnehmung damit korrespondierender eigener Emotionen. Dem Fragebogen liegen zwei basale Dimensionen von Empathie zugrunde, eine auf reale Situationen und eine mit fiktiven Situationen assoziierte Empathie:

- Faktor 1: Einfühlungsbereitschaft (Phantasie-Empathie, Empathie in fiktiven Situationen)
Der Faktor Einfühlungsbereitschaft klärt 28,8% der Gesamtvarianz auf und wird repräsentiert durch die Items 1-13. Er erfasst die Fähigkeit und Bereitschaft einer Person, sich in das Erleben und Verhalten einer anderen Person in fiktiven Inhalten hineinzuversetzen.
- Faktor 2: Betroffenheit (Kommunikationsempathie, Empathie in realen Situationen)
Der zweite Faktor Betroffenheit klärt 9,4% der Gesamtvarianz auf und wird durch die Items 14-25 repräsentiert. Er erfasst das Ausmaß, in dem eine Person dazu tendiert, empathisches Verhalten in realen Situationen zu erleben.

Empathiefragebogen nach Schmitt (1982)

Empathie wird hier als instinktive Nachahmung verstanden, die zu identischen oder ähnlichen propriozeptiven Empfindungen des Subjekts mit seinem Gegenüber führt, dadurch einen ähnlichen oder gleichen emotionalen Zustand zwischen beiden Personen bewirkt und als Konsequenz die soziale Verständigung erleichtert.

- Empathie als soziale Kognition der Perspektiven- oder Rollenübernahme, die eine Befähigung voraussetzt, die Welt aus den Augen des anderen zu betrachten und dadurch dessen

³ Ich bedanke mich bei Andreas Korbach, Matthias Jacques und Melanie Scazzari für die Mithilfe bei diesem Teil der Untersuchung.

Handlungsbereitschaften zu antizipieren. Dieses Empathieverständnis wurzelt im Symbolischen Interaktionismus und der kognitiven Entwicklungstheorie von Piaget.

- Empathie als Fähigkeit und Bereitschaft, die Gefühle und Bedürfnisse des Gegenübers richtig zu erkennen und nachzuempfinden. Diese Auffassung von Empathie findet sich vor allem in Theorien wieder, die Empathie als Voraussetzung für Hilfsbereitschaft und angemessenes Hilfehandeln postulieren.

Eine Überprüfung des Fragebogens ergab eine interne Konsistenz von .70, eine Faktorenanalyse konnte die Items 1, 7, 12, 14, 17, 18 (>.50) der Dimensionen Bereitschaft zur Rollenübernahme und die Items 10, 11, 15 (>.50) der Dimension „Empathieabwehr“ zuordnen.

Die folgende Tabelle 6 zeigt für die Faktoren EC, FS und PD signifikant positive Korrelationen mit den beiden anderen Fragebögen zur Empathiemessung. Die EC-Skala korreliert mit beiden Vergleichsmessungen um .58 ($p < .01$); die FS zeigt die höchste Korrelation ($r = .78$, $p < .01$), PD korreliert nur mäßig ($r = .31$, $p < .05$) mit der Distress-Skala des Empathiefragebogens vom Schmitt (1982). Insgesamt gesehen sind die Korrelationen wegen der recht kleinen Stichprobe ($n = 51$) aber umso bedeutsamer, wie die teilweise sehr hohen Effektstärken zeigen.

	Empathiefragebogen Schmitt (1982)			E-Skala	
	Empathie	distress	Perspektivenübernahme	Empathie-Score	Fantasy-score
EC	,581(**) ES = 1,43			,583(**) ES = 1,43	
FS					,783(**) ES = 2,52
PD		,306(*) ES = 0,64			
PT			-,167		

Tabelle 6: Interkorrelation und Effektstärken der Subskalen des SPF(IRI) mit dem Empathiefragebogen und der E-Skala ($p < .01$; * $p < .05$)**

Lediglich die Skala PT weist keinen Zusammenhang zur vergleichbaren Skala im Empathiefragebogen auf. Die bei Schmitt verwendeten Items im Subtest Perspektivenübernahme lauten:

- Ich versuche immer, mich in meinen Gegner hineinzuversetzen, bevor ich etwas unternehme (Item 1).
- Wenn ich einem gemütskranken Menschen begegne, frage ich mich, wie ich mich an seiner Stelle fühlen würde. (Item 7)
- Wenn ich einen sehr alten Menschen sehe, frage ich mich, wie ich mich an seiner Stelle fühlen würde. (Item 12)
- Ich habe oft versucht mir vorzustellen, wie sich jemand fühlt, der dauernd Hunger leidet. (Item 14)
- Selbst wenn ich mich mit jemandem streite, versuche ich mir vorzustellen, wie er zu seiner Sichtweise steht. (Item 17)
- Wenn ich ein geistig behindertes Kind sehe, versuche ich mir vorzustellen, wie es die Dinge sieht. (Item 18)

Wie sich feststellen lässt, sind diese Items konkreter auf bestimmte Situationen formuliert („Gegner“, „gemütskrank“, „geistig behindert“), die damit im Gegensatz zu einer eher allgemeineren Aussage bei

Davis stehen. Zudem wird vorausgesetzt, dass die beschriebenen Situationen entweder schon mal erlebt wurden oder bekannt seien, was möglicherweise in unserer Stichprobe nicht der Fall war.

Selbsteinschätzung

Bei 745 Vpn (Alter zwischen 10 und 63 Jahren) wurden inzwischen Selbsteinschätzungen bzgl. Ihrer Empathie durch folgende 5 Fragen erhoben:

- Würden Sie sich selbst als einfühlsamen Menschen beschreiben?
- Können Sie anderen Menschen geduldig zuhören?
- Haben Sie Verständnis für Eitelkeiten, persönliche Steckenpferde und bestimmte Verhaltensweisen wie Geltungssucht, Profiliergehabe, Rechthaberei, Überheblichkeit o.ä. bei Freunden oder Bekannten?
- Verstehen Sie die Gefühle anderer und achten auf ihre Bedürfnisse?
- Die Körpersprache anderer Menschen sagt oft mehr aus als ihre Worte

Die Fragen stammen aus Beschreibungen empathischer Fähigkeiten in Beratungs- und Therapieberufen. Der Schwerpunkt liegt in einer Kombination aus perspective taking und empathic concern.

	FS	EC	PT	PD
Selbsteinschätzung	,233**	,527**	,469**	-,050

Tabelle 7: Korrelation der Selbsteinschätzung mit den Empathie-Faktoren (: $p < .01$)**

Die Selbsteinschätzung bringt bei den Faktoren FS, EC und PT eine befriedigende Übereinstimmung, im Hinblick auf die beiden direkt erwähnten Faktoren sogar gute Korrelationswerte; der Faktor distress kann keine Übereinstimmung bringen, da er in den Fragen nicht repräsentiert ist.

Zusammenfassung

Der Interpersonal Reactivity Index (IRI) ist einer der am häufigsten in der Praxis eingesetzte Fragebogen zur Messung von Empathie. Die wenigen deutschen Übersetzungen weisen allerdings ähnliche testtheoretischen Schwächen auf wie das Original (vgl. dazu z.B. Enzmann, 1996; Beven, 2004; Lauterbach & Hossler, 2007). Nach einer mehrfachen Überarbeitung aufgrund faktorenanalytischer Ergebnisse und marginalen Umformulierungen in der Übersetzung konnten wir eine deutsche Version des IRI erstellen, der gute Kennwerte in den Bereichen der Reliabilität, der faktoriellen Validität und der Itemtrennschärfe aufweist. Durch das Weglassen aller negativ formulierten Items konnte die Auswertung erheblich verbessert werden, so dass nun eine einfachere Scorebildung möglich wurde.

Die interne und externe Validität konnte durch Heranziehen von Vergleichswerten aus existierenden Empathiefragebögen sowie einer Selbsteinschätzung der Pbn gesichert werden.

Die aktuelle Fassung des Fragebogens findet man hier: www.cpaulus.de

Literatur

- Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high-functioning autism and normal sex differences. *Journal of Autism and Developmental Disorder*, 34, 163-175.
- Beven, J.P., O'Brien-Malone, A. & Hall, G. (2004). Using the Interpersonal Reactivity Index to assess empathy in violent offenders. *International Journal of Forensic Psychology*, 1, 33-41.
- Christopher, F. S., Owens, L.A. & Stecker, H. (1993). Exploring the dark side of courtship: A test of a model of male premarital sexual aggressiveness. *Journal of Marriage and the Family*, 55, 469-479.
- Curwen, T. (2003). The importance of offense characteristics, victimisation history. Hostility, and social desirability in assessing empathy of male adolescent sex offenders. *Sexual Abuse*, 15, 347-364.
- Davis, M. (1980). A multidimensional approach to individual differences in empathy. *JSAS catalogue of Selected Documents in Psychology*, 10, 85.
- Davis, M. (1983): Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126.
- Davis, M. & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality*, 25, 70-87.
- D'Orazio, D. M. (2004)- Letter to the Editor. *Sexual Abuse: A Journal of Research and Treatment*, 16, 173-174.
- Dziobek, I., Rogers, K., Fleck, S., Hassenstab, J., Gold, S., Wolf, O., & Convit, A. (2005). In search of "matser mindreaders": Are psychics superior in reading the language of the eyes?. *Brain and Cognition*, 58, 240-244.
- Enzmann, D. (1996). *Gestreßt, erschöpft oder ausgebrannt? Einflüsse von Arbeitssituation, Empathie und Coping auf den Burnoutprozeß*. München: Profil
- Hassenstab, J., Dziobek, I., Rogers, K., Wolf, O. & Convit, A. (2007). Knowing what others know, feeling what others feel. A controlled study of empathy in psychotherapists. *Journal of Nervous and Mental Disease*, 195, 277-281.
- Ireland, J.L. (1999). Provictim attitudes and empathy in relation to bullying behaviour among prisoners. *Legal and Criminological Psychology*, 4, 41-66.
- Jolliffe, D. & Furlong, D.P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, 9, 441-476.
- Lamsfuss, S. M., Silbereisen, R. K. & Boehnke, K. (1990). *Empathie und Motive prosozialen Handelns*. Arbeitsbericht Nr. 10 (1990) der FU Berlin.
- Lauterbach, G. & Hosser, D. (2007). Empathy in prisoners. *Swiss Journal of Psychology*, 66, 91-101.

-
- Paulus, C., & Zumkley-Münkel, C. (1992). Empathie, Kompetenz und Altruismus. In L. Montada (Hrsg.): Bericht über den 38. Kongress der DGPs in Trier 1992. Göttingen: Hogrefe, S. 750.
- Paulus, C. (1997). Empathie, Kompetenz und Altruismus. WWW-Dokument: <http://bildungswissenschaften.uni-saarland.de/personal/paulus/empathy/empathie.htm>
- Pulos, S., Elison, J. & Lennon, . (2004). The hierarchical structure of the interpersonal reactivity index. *Social Behavior and Personality*, 32, 355-360.
- Schmitt, M. (1982). Empathie: Konzepte, Entwicklung, Quantifizierung. Bericht aus der Arbeitsgruppe „Verantwortung, Gerechtigkeit, Moral“ Forschungsprojekt: Entwicklung interpersonaler Verantwortlichkeit und interpersonaler Schuld (P.I.V.),
- Silbereisen, R. K. & Schulz, W. (1977). Prüfung der Testgüte einer „Empathie-Skala“. *Diagnostica*, 23, 179-187.
- Siu, A. M., & Shek, D. T. L. (2005). validation of the interpersonal reactivity index in a Chinese context. *Research on Social Work Practise*, 15, 118-126.
- Stotland, S., Mathews, K., Sherman, S., Hansson, R. & Richardson, B. (1978). *Empathy, fantasy, and helping*. Beverly Hills, Ca.: Sage.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur Psychologischen Methodenlehre*. Göttingen: Hogrefe.
- Mischo, C. (2003). Wie valide sind Selbsteinschätzungen der Empathie? *Gruppendynamik und Organisationsberatung*, 34(3), 187-203.