

Ergebnisse zur Äquivalenzprüfung einer deutschen Übersetzung der „Personality Research Form“ (PRF) von Jackson

H. Stumpf¹, A. Angleitner¹ & F.W. Steege²

1 Psychologisches Institut der Universität Bonn

2 Bundesministerium der Verteidigung, Abt. P II 4

Einleitung

Die „Personality Research Form“ (PRF) von Jackson (1974) hat wegen ihrer für einen Persönlichkeitsfragebogen besonders günstigen psychometrischen Eigenschaften (Anastasi, 1972; Kelly, 1972; Wiggins, 1972) im anglo-amerikanischen Bereich weite Verbreitung gefunden und ist bereits in mehrere Sprachen übertragen worden (Skinner & Rampton, 1973; Schulken, van den Broek & Diekstra, 1974; Skinner, Jackson & Rampton, 1976). Auch in deutscher Sprache wurden bereits PRF-Items zu Forschungszwecken verwendet (Hoffmann, 1969; Crott & Baltes, 1973; Breuer, Keil, Kleiber, Meier & Piontkowski, 1975; Keil, Meier & Piontkowski, 1977).

Das Inventar basiert auf Konstrukten der Personologie Murrays (1938) und umfaßt in seiner Gesamtversion (mit zwei parallelen Formen, AA und BB) folgende 20 Inhaltsskalen:

- | | |
|-----------------------------|-----------------------------|
| 1) Abasement (Ab) | 11) Exhibition (Ex) |
| 2) Achievement (Ac) | 12) Harmavoidance (Ha) |
| 3) Affiliation (Af) | 13) Impulsivity (Im) |
| 4) Aggression (Ag) | 14) Nurturance (Nu) |
| 5) Autonomy (Au) | 15) Order (Or) |
| 6) Change (Ch) | 16) Play (Pl) |
| 7) Cognitive Structure (Cs) | 17) Sentience (Se) |
| 8) Defence (De) | 18) Social Recognition (Sr) |
| 9) Dominance (Do) | 19) Succorance (Su) |
| 10) Endurance (En) | 20) Understanding (Un). |

Hinzu kommen zwei Validitätsskalen, Desirability (Dy) und Infrequency (In).

Die vorliegende Arbeit berichtet über ein Forschungsvorhaben zur Adaptierung des Gesamtfragebogens für den deutschen Sprachraum (vgl. auch Angleitner, Stumpf & Wieck, 1976; Stumpf & Angleitner, 1977).

1. Zur Adaptierung

1.1. Methodologische Vorbemerkungen

Bei der Übertragung eines Persönlichkeitsfragebogens in eine andere Sprache steht der Untersucher vor dem doppelten Problem, eine inhaltlich und sprachlich brauchbare Übersetzung zu erstellen und deren psychologische Äquivalenz mit der Originalversion zu überprüfen. Unter „psychologischer Äquivalenz“ wird hier verstanden, daß ein In-

ventar bei Personen vergleichbarer Persönlichkeitsstruktur in beiden Sprachversionen vergleichbare Response-Muster hervorruft.

Von den bei Butcher & Pancheri (1976, S. 37–42) angeführten Itemübersetzungstechniken ist die *direkte Übersetzung* gewiß die am wenigsten geeignete Methode, beiden Aspekten des Problems zugleich gerecht zu werden, da eine Überprüfung der Äquivalenz der Übersetzung dabei nicht stattfindet. Der „*committee approach*“, d. h. die Kooperation mehrerer Übersetzer, ist nach Butcher & Pancheri (1976, S. 37) „a vast improvement over the solitary translator method since it reduces errors and diminishes biases“. Wendet man ihn allein an, bleibt jedoch auch bei diesem Ansatz die psychologische Äquivalenz ungeprüft.

Die *Rückübersetzungsmethode* hat sicher den Vorteil, daß sie bei erfolgreicher Anwendung einen evidenten Nachweis erbringt, daß der Inhalt eines Items getreu in die Zielsprache übertragen wurde. Sie verlangt aber Ermessensentscheidungen, wenn das rückübersetzte Item nicht wortwörtlich mit dem Original übereinstimmt. Zudem kann selbst bei vollkommen befriedigendem Resultat die psychologische Äquivalenz zwar als weitgehend plausibel, nicht aber als nachgewiesen gelten, da wichtige Konnotationen des Items in der Zielsprache selbst bei inhaltlich vollkommen getreuer Rückübersetzung u. U. wieder preisgegeben werden müssen.

Es ist daher naheliegend – egal, welche der genannten Methoden man wählt –, die Übersetzung gesondert noch auf ihre psychologische Äquivalenz zu überprüfen. Als Verfahren dazu bieten sich v. a. die „*pretest field study*“ und verschiedene Formen der „*bilingual retest technique*“ (Butcher & Pancheri, 1976, S. 39–42) an. Die „*pretest field study*“ kann darin bestehen, Personen aus der Zielkultur bestimmte Itemeigenschaften (z. B. Akzeptabilität, Verständlichkeit) beurteilen zu lassen, um so krasse Übertragungsmängel zu vermeiden, oder sie kann darauf hinauslaufen, psychometrische Eigenschaften der Original- und übersetzten Version zu vergleichen (wie es bei einigen PRF-Übersetzungen (Skinner & Rampton, 1973; Schulken, van den Broek & Diekstra, 1974; Skinner, Jackson & Rampton, 1976) geschah).

Bei der vorliegenden Untersuchung wurde einer einfachen Variante der „*bilingual retest technique*“ der Vorzug gegeben, da mit ihr Item für Item die Responses einer und derselben Personengruppe auf beide Sprachversionen korreliert werden können; zudem können auch auf der Skalenebene des Fragebogens Korrelationen berechnet werden. Zwar hat die einfache „*bilingual retest technique*“, d. h. die Vorgabe aller Items der Original- und übersetzten Version an eine Stichprobe zweisprachiger Probanden, den Nachteil, daß sie eine quasi-Wiederholungstestung (auf deren Problematik besonders Loevinger (1957) hingewiesen hat) erforderlich macht, aber der Nachteil einer unmittelbaren Testwiederholung konnte bei der PRF-Adaptierung zumindest z. T. durch eine besondere Staffelung der Parallelformen und Sprachversionen ausgeglichen werden (vgl. 1.2.).

Die von Prince & Mombour (1967) vorgeschlagene Variante der Technik, zwei Zufallsgruppen jeweils nur eine Hälfte der Original- und übersetzten Items bearbeiten zu lassen, wurde in Anbetracht dieser Ausgangslage nicht angewendet, da man

mit ihr keine Korrelationen zwischen den Itembeantwortungen beider Versionen berechnen kann.

Da vorgesehen war, die Ergebnisse der Äquivalenzprüfung für eine Revision der ersten Übersetzungsversion zu verwenden und mit der revidierten Fassung erneut eine Äquivalenzprüfung vorzunehmen, wurde für die Erstellung der ersten Fassung die am wenigsten aufwendige Technik, die direkte Übersetzung, gewählt.

Die Anwendung besonderer Übersetzungsverfahren, wie z. B. des „decentering“ (Werner & Campbell, 1970), kam wegen der Länge des Inventars nicht in Frage. Ein „decentering“ hätte zudem den sorgfältig ausbalancierten inneren Aufbau des Fragebogens zerstört.

1.2. Erste Äquivalenzprüfung: Vorgehensweise und Stichprobe

Die von Angleitner, Angleitner-Binns & Gephardt übersetzten PRF-Formen AA und BB sowie die englischen Originalversionen wurden einer ausgewählten Stichprobe von Probanden, die beide Sprachen beherrschten (fortgeschrittene Anglistik-Studenten; Auszubildende zum Fachdolmetscher; Anglophone, die seit längerem im deutschen Sprachraum ansässig waren), zur Bearbeitung in folgender Reihenfolge vorgegeben: Form AA (engl.) – Form BB (dtsch.) – Form AA (dtsch.) – Form BB (engl.). Ein derartiger Bearbeitungsdurchgang sollte in einer Woche abgeschlossen und nach zwei Wochen wiederholt werden. Zusätzlich wurde ein kurzer Ergänzungsfragebogen zu demographischen Variablen der Pbn und zu ihrer Einschätzung der „Personality Research Form“ vorgegeben.

Die Testprotokolle von insgesamt 92 Pbn (115 Formulare waren ausgegeben worden) konnten ausgewertet werden. Das durchschnittliche Alter dieser 92 Pbn (53 W; 39 M) lag bei 26.6 Jahren (Streubereich: 19–50). Als Muttersprache wurde in 76 Fällen Deutsch und in 16 Fällen Englisch angegeben. Die meisten Pbn gaben zudem Kenntnisse in weiteren Sprachen an.

1.3. Ergebnisse

Insgesamt fielen die Phi-Korrelationen zwischen den Beantwortungen der englischen und der entsprechenden deutschen Items in Anbetracht der herkömmlichen Erfahrungen bezüglich der Gütekriterien von Fragebogenitems überraschend hoch aus, zumal die Übersetzungen oft sehr frei und nicht in jedem Fall inhaltlich optimal waren. Nur etwa 15 Prozent aller 880 Items hatten in beiden Untersuchungsdurchgängen Äquivalenzkoeffizienten unter .41.

Bei der Beurteilung der psychologischen Übersetzungsäquivalenz ist die Unreliabilität des Items eine natürliche Orientierungsmarke. In Anbetracht der hohen durchschnittlichen Stabilität der Itembeantwortungen wurden für die Revision der Übersetzung die Items mit Phi-Koeffizienten von .40 und darunter als „revisionsbedürftig“ (d. h. eine Neuübersetzung erfordernd) und diejenigen mit Phi-Werten zwischen .41 und .60 als „überprüfungsbedürftig“ klassifiziert.

Trotz der Revisionsbedürftigkeit einer Reihe von Items ergaben sich für die Übersetzungsäquivalenz auf Skalenebene überraschend hohe (Produkt-Moment-)Korrelationen (vgl. Tabelle 1).

Tabelle 1
Äquivalenz*, Homogenität* (nach Cronbach-„alpha“),
Stabilität und Parallelität* der Skalen der vorläufigen deutschen PRF-Übersetzung
(ohne Dezimalpunkte)

PRF-Skalen	Form AA			Form BB			Form AAxBB
	Äqui- valenz	Homo- genität	Stabi- lität	Äqui- valenz	Homo- genität	Stabi- lität	Parallelität
Ab	77	63	84	69	45	73	66
Ac	85	64	92	85	49	89	79
Af	86	85	92	87	79	94	78
Ag	79	67	85	82	68	85	74
Au	88	67	93	81	58	88	69
Ch	86	70	91	80	66	88	71
Cs	84	72	90	83	61	88	77
De	78	59	88	80	56	80	59
Do	88	79	92	87	73	90	76
En	80	74	92	86	78	86	71
Ex	89	83	94	89	80	91	86
Ha	94	87	96	95	87	95	84
Im	79	69	88	93	81	93	80
Nu	86	68	91	88	78	94	79
Or	89	83	95	88	75	92	84
Pl	82	73	88	86	73	87	75
Se	87	69	82	87	67	88	77
Sr	88	86	92	87	84	89	83
Su	85	77	91	91	79	93	77
Un	82	65	89	87	68	89	68
In	52	61	75	82	60	72	54
Dy	71	50	76	70	51	80	45
Mittelwert	84	73**	90	86	71**	89	75

* Werte des ersten Testdurchgangs

** Ohne die Validitätsskalen In und Dy

Die Homogenitäten und Paralleltestreliabilitäten der deutschen PRF-Skalen (vgl. Tabelle 1) erwiesen sich bis auf wenige Ausnahmen als vergleichbar mit den von Jackson (1974), Stricker (1974) und Tetenbaum (1974) für anglo-amerikanische Stichproben berichteten Werten. Die Skalenstabilitäten liegen in den meisten Fällen etwas über denen früherer Untersuchungen (Bentler, 1964; Hoffmann, 1971) mit der Originalversion; bei diesen Studien fanden die Zweittestungen aber auch unter anderen äußeren Bedingungen statt als die Ersttestungen.

2. Zur Revision der vorläufigen Übersetzung

Bei der Revision der Übersetzungen wurden alle Items, bei denen die Äquivalenz in beiden Durchgängen unter .40 geblieben war (69 aus der Form AA und 62 aus BB), von drei Anglisten, die die vorläufige deutsche Version nicht kannten, neu übersetzt. Von den drei Übersetzungen pro Item, die oft erheblich divergierten, wurde von den Autoren der revidierten Fassung (Stumpf & Wieck, 1976) die ihrer Meinung nach beste ausgewählt, wobei sich in einigen Fällen Umformulierungen oder Zusammensetzungen von Itemteilen aus mehreren Übersetzungsversuchen als unumgänglich erwiesen. Zugleich wurden alle übrigen der insgesamt 880 Items durchgesehen. Verbesserungen der Übersetzungen wurden in der Regel vorgenommen, wenn der Äquivalenzkoeffizient eines Items im ersten Durchgang der Voruntersuchung weniger als .60 betragen hatte, wenn stilistische oder Übertragungsmängel festgestellt wurden oder wenn sich die Autoren durch eine genauere Übersetzung oder bessere Formulierung eine höhere Übertragungstreue versprochen.

Insgesamt erfuhren 359 der Items der alten deutschen Form AA in einer im folgenden als „Zwischenform“ bezeichneten Neufassung Veränderungen in irgendeiner Form. Bei 254 dieser Items erfolgten die Änderungen wegen zu geringer Äquivalenzkoeffizienten, bei 105 aus semantischen oder stilistischen Gründen. In der Form BB wurden zur Zwischenform hin 307 Items geändert, und zwar 205 wegen der Äquivalenzkoeffizienten, 102 aus semantischen oder stilistischen Gründen. Die Inspektion der Korrelationen zwischen den englischen und deutschen Items ergab, daß die Höhe der Koeffizienten in erster Linie von der Übersetzungsgenauigkeit abzuhängen scheint. Selbst Items mit schlechter Idiomatik im Deutschen erbrachten hohe Korrelationen, wenn sie inhaltlich nur eng am Original orientiert waren. Sie wurden aber als für die Endform nicht tragbar angesehen, da diese nicht mehr an zweisprachigen Stichproben verwendet werden soll. Versuche, Items semantisch etwas freier, dafür aber in typisch deutsche Redewendungen zu übersetzen, resultierten nicht selten in mäßigen Korrelationen.

Bei der Revision zur Zwischenform hin waren den Autoren die Skalenzugehörigkeit und Schlüsselrichtung der Items bekannt, und es wurden im Sinne einer Steigerung der inhaltlichen Komponente der Validität (Loevinger, 1957) bewußt bei mehreren Übersetzungsmöglichkeiten diejenige bevorzugt, die dem jeweiligen Konstruktinhalt am besten zu entsprechen schien.

Bei den im folgenden dargestellten weiteren Schritten der Adaptierung wurde von der Überlegung ausgegangen, daß selbst bei genauester Übertragung der denotativen Bedeutung eines Items in eine andere Sprache noch keine Gewähr dafür besteht, daß auch die konnotative Bedeutung erhalten bleibt (vgl. u. a. Schwebcke, Lück & Jandron, 1973, S. 444). Betrachtet man die Gesamtmenge der Items, so ist mit einer von Fall zu Fall variierenden Verschiebung des semantischen Umfeldes bei der Übersetzung zu rechnen, die unsystematisch in alle erdenklichen Richtungen erfolgen kann. Diese Tendenz müßte sich, bezogen auf die einzelnen Skalen, u. a. in einer Verminderung ihrer Homogenität und der durchschnittlichen Trennschärfen manifestieren.

Ein weiteres Moment, das bei der Übersetzung von Items schwer zu kontrollieren ist, ist der Unterschied im Grad der sozialen Erwünschtheit, die die in ihnen beschriebenen Handlungen oder Sachverhalte in der Zielpopulation gegenüber der Ursprungspopulation haben (Schwebcke, Lück & Jandron, 1973, S. 445–446). Bei der „Personality Research Form“ stellt sich dieses Problem in ähnlicher Weise wie das eines evtl. Homogenitätsverlustes: Bei der PRF-Konstruktion wurde nämlich außer einer Homogenisierung der Skalen auch ihre Reinigung von „social desirability“ angestrebt (Jackson, 1970), so daß bei der Übersetzung für den Durchschnitt der Items ein höheres Ausmaß an „social desirability“ zu erwarten ist.

Um Items, die hinsichtlich ihres Inhalts zu weit vom Konstrukt ihrer Skala abwichen und solche mit einem hohen Gehalt an sozialer Erwünschtheit in der deutschen Übersetzung ausfindig zu machen und, soweit möglich, in einer anderen Formulierung wiederzugeben, wurden sämtliche Items der Zwischenform von fünf zu diesem Zweck trainierten Beurteilern eingeschätzt. Die Rater (fortgeschrittene Psychologiestudenten) wurden mit dem Konzept der „social desirability“ von Items nach Edwards (1957) und mit den relevanten Konstrukten der Personologie Murrays (1938) sowie deren Rezeption bei Jackson (1974) vertraut gemacht. Sie hatten zunächst in einer Einübungsphase 40 nach den vorliegenden Ergebnissen besonders inhaltstreu übersetzte und trennscharfe Items je einer der 20 Inhaltsskalen zuzuordnen. Die Ergebnisse dieser Zuordnung wurden in einer Gruppendiskussion durchgesprochen, um evtl. Mißverständnisse der Konstrukte zu beheben und die Beurteilungsprozedur einzuüben. Sodann waren alle 880 Items der Zwischenform, die zuvor in eine Zufallsfolge gebracht worden waren, in der beschriebenen Weise einzuordnen. Die Schlüsselrichtung jedes Items war vorgegeben.

Die Einschätzung nach sozialer Erwünschtheit erfolgte anhand einer eigenen Übersetzung der Instruktion zur „SDSV“-Skala von Edwards (1957) in einem neunstufigen Rating. Das „matching“ der Items zu ihren Skalen gelang in ca. 70 Prozent der Fälle allen fünf Ratern ohne Fehler, allerdings erwies sich die Zuverlässigkeit der Zuordnungen als skalenabhängig. Items aus Skalen, die bei der normalen Anwendung des Inventars hoch korrelieren (insbesondere Ac und En), wurden nur in wenigen Fällen von allen korrekt zugeordnet.

Alle Items mit weniger als vier korrekten Zuordnungen und mit extremen durchschnittlichen „SDSV“-Werten (2 und weniger bzw. 8 und mehr) wurden erneut durchgesehen. Es wurde versucht, diesen Items im Deutschen eine neue Formulierung zu geben oder alternative Ausdrücke zu den Übersetzungen der Zwischenform zu wählen, um den Konstruktgehalt besser wiederzugeben bzw. die „social desirability“-Komponente in den Hintergrund treten zu lassen. Dabei wurde streng darauf geachtet, daß die neue Formulierung der alten als Übersetzung ebenbürtig war. Deshalb mußte in manchen Fällen, insbesondere bei kurzen Items (wie z. B. „I rarely swear“), der Versuch, aufgegeben werden, eine bessere Wendung zu finden.

Im Rahmen dieser Prozedur wurden 38 Items der Zwischenform AA geändert; von diesen waren sechs aus der ursprünglichen Fassung beibehalten und 32 bereits zur Zwischenform hin verändert worden. Bei BB wurden 60 Items erneut revidiert, von denen

45 schon für die Zwischenform geändert worden waren. Die so entstandene „revidierte Übersetzung“ (Stumpf & Wieck, 1976), die in der im folgenden zu berichtenden zweiten Äquivalenzprüfung Verwendung fand, enthält in der Form AA noch 75 und in BB noch 118 Items, die gegenüber der ursprünglichen Fassung keine Veränderung erfahren haben (vgl. Figur 1).

3. Zweite Äquivalenzprüfung: Vorgehensweise und Stichprobe

Die in der ersten Untersuchung vorgegebene Reihenfolge der zu bearbeitenden Fragebögen wurde beibehalten, ein zweiter Durchgang erfolgte diesmal nicht. Als Bearbeitungszeit wurde eine Woche angesetzt. Von den 110 ausgegebenen Testformularen wurden 93 zurückerhalten, von denen aus terminlichen Gründen aber nur 91 in die Verrechnung der Daten einbezogen werden konnten. Von diesen 91 Vpn (62 weibl., 29 männl.) waren die meisten fortgeschrittene Anglistikstudenten; hinzu kamen Anglophone und Personen, die durch Auslandsaufenthalte oder familiäre Umstände besondere Kenntnisse der englischen Sprache besaßen.

Der Altersmittelwert dieser zweiten Stichprobe beträgt 23.7 Jahre (Streubereich: 18–52).

3.1. Ergebnisse

Durch die Revision der Übersetzung konnte die Übersetzungsäquivalenz der Items im allgemeinen erheblich gesteigert werden.

Bei der Form AA fielen pro Skala höchstens zwei Phi-Koeffizienten niedriger als .40 aus; bei der Form BB ist das Gesamtergebnis nicht minder günstig (wenn man von einem Ausnahmefall, der Skala Au, absieht). Die Ergebnisse sind im einzelnen bei Stumpf & Angleitner (1977) dokumentiert.

Da im Vergleich zur Voruntersuchung Stichprobenunterschiede als Ursache für Abweichungen zwischen den Übersetzungsäquivalenzkoeffizienten weitestgehend ausschließen, kann geschlossen werden, daß nach den Ergebnissen unserer Untersuchungen schon relativ geringe inhaltliche oder stilistische Änderungen ausreichen können, um das Antwortverhalten der Vpn in weitem Umfang den Reaktionen auf die Originalitems anzugleichen. Zwar sind nach den Ergebnissen der Voruntersuchung oft schon mit mäßigen Übersetzungstexten relativ hohe Äquivalenzkoeffizienten zu erreichen, es kann aber dennoch als Resultat unserer Arbeiten festgehalten werden, daß der hohe Arbeitsaufwand in die sprachliche Gestaltung von Itemübersetzungen sich gelohnt hat.

Die Korrelationen der Skalenwerte der revidierten deutschen PRF-Version mit den entsprechenden Skalenwerten des englischen Originals (vgl. Tabelle 2) haben sich, verglichen mit den Werten der Voruntersuchung (vgl. Tabelle 1), nur in wenigen Fällen so deutlich verändert, wie dies nach den Ergebnissen auf Itemebene eigentlich zu erwarten gewesen wäre. Immerhin konnte aber für alle Inhaltsskalen beider Formen das Desiderat realisiert werden, daß der Koeffizient für die Übersetzungsäquivalenz zumindest bei .80 liegen soll.

Bei der Infrequenzskala müssen die in Tabelle 2 angeführten Werte mit dem Vorbehalt interpretiert werden, daß nur eine sehr geringe Streuung vorliegt, wie es der Natur

der Skala entspricht. Bei der Validitätsskala Dy scheint durch die Revision in beiden Formen – was die Gesamtskala angeht – die größte Verbesserung gelungen zu sein.

Die meisten Skalen der deutschen PRF-Version sind hinsichtlich ihrer internen Konsistenz mit denen der Originalversion vergleichbar. Die für anglo-amerikanische Stichproben berichteten KR-20 (Jackson, 1974, Tab. 8) und Cronbach-„alpha“-Werte (Stricker, 1974; Tetenbaum, 1974) der Skalen der Form AA liegen mit Ausnahme der Skalen Ab und De (und evtl. Ag) sehr dicht bei den an unserer Stichprobe ermittelten Werten der englischen Form AA, so daß kein Anlaß besteht, unsere Stichprobe bezüglich dieses Merkmals als untypisch anzusehen.

Bei der Skala In ist zu beachten, daß in diesem Fall die Homogenitätsforderung nicht in der gleichen Weise gestellt werden kann wie bei den anderen Skalen, da sie einem anderen Zweck dient als der Erfassung eines vermeintlich einheitlichen Persönlichkeitsmerkmals.

Tabelle 2
Äquivalenz, Homogenität (Cronbach-„alpha“)
und Parallelität der revidierten deutschen PRF-Skalen

PRF-Skalen	Form AA		Form BB		Form AA x BB
	Äquivalenz	Homogenität	Äquivalenz	Homogenität	Parallelität
Ab	83	50	80	48	60
Ac	87	77	88	69	76
Af	84	74	92	74	77
Ag	84	74	88	77	73
Au	83	70	89	70	64
Ch	87	69	92	70	76
Cs	88	78	91	72	77
De	88	65	86	49	67
Do	93	88	95	89	89
En	90	84	93	82	85
Ex	93	85	95	87	87
Ha	95	86	96	84	87
Im	86	69	94	79	73
Nu	90	72	89	67	74
Or	94	82	94	81	87
Pl	88	76	91	72	78
Se	92	68	83	46	60
Sr	89	85	92	83	84
Su	89	82	93	81	84
Un	84	68	92	68	71
In	79	52	40	18	22
Dy	88	56	87	60	68
Mittelwert	89	75**	91	73**	74

** ohne Validitätsskalen In und Dy

Im Vergleich zu den internen Konsistenzen der Skalen der ursprünglichen deutschen Übersetzung konnten durch die Revision zwar die meisten, aber nicht alle Skalenhomogenitäten verbessert werden. Praktisch bedeutsame Verbesserungen konnten nur in wenigen Fällen (Form AA: En; Form BB: Ac, Au, Cs, Do) erreicht werden; zieht man zudem in Betracht, daß zwei der eingetretenen Verschlechterungen (Form BB: Nu, Se) auch als praktisch bedeutsam angesehen werden können, so muß man zu dem Schluß kommen, daß bezüglich einer Homogenitätserhöhung die bei der Revision durchgeführten Item-Skalen-Zuordnungen nur wenig erfolgreich gewesen sind. Auch die Parallelität der Skalen konnte nicht wesentlich verbessert werden (vgl. Tabelle 2).

4. Diskussion

Da die deutsche Version der „Personality Research Form“ ursprünglich nur zu einem Instrument für die Persönlichkeitsforschung und -diagnostik im deutschen Sprachraum entwickelt werden sollte, war das wichtigste Motiv zur Durchführung der Äquivalenzprüfung ein ökonomisches: Es galt einen mit hohem Aufwand zusammengestellten Itempool unter minimalem Verlust an psychometrischer Qualität ins Deutsche zu übertragen statt ein deutsches Inventar von Grund auf neu zu entwickeln. Ein Urteil darüber, inwieweit dieses Unternehmen erfolgreich gewesen ist, wird durch den Mangel an vergleichbaren Untersuchungen mit zweisprachigen Stichproben erschwert, zumal bei den sich der „bilingual retest technique“ bedienenden Arbeiten die Auswertung in der Regel erst auf der Skalenebene begann (Spreen & Spreen, 1963; Lonner, 1967; Glatt, 1969; Schwebcke, Lück & Jandron, 1973; Butcher & Gur, 1974; Spielberger, Gonzalez-Reigosa, Martinez-Urrutia, Natalicio & Natalicio, 1971; Prince & Mombour, 1967; Mote, Natalicio & Rivas, 1971).

Da die in der vorliegenden Untersuchung gefundenen Skalenäquivalenzen im Durchschnitt aber über den von Glatt (1969) für verschiedene MMPI-Sprachversionen berichteten Werten liegen und zudem die Itemäquivalenzen in der Mehrzahl, verglichen mit den Reliabilitäten, recht hoch ausfielen, erscheint es zumindest gerechtfertigt, die Weiterverfolgung des gesteckten Zieles als aussichtsreich anzusehen.

Nach den Ergebnissen der zweiten Äquivalenzprüfung ist die einfache „bilingual retest technique“ außerdem in der Lage, Mängel in der Übersetzung von Items ausfindig zu machen, deren Bereinigung besonders auf der Itemebene zu einer Steigerung der psychologischen Äquivalenz führt.

Des weiteren stellte sich heraus, daß die psychologische Äquivalenz in sehr enger Beziehung zur sprachlichen Übersetzungsgüte steht, so daß es für zukünftige Adaptierungsversuche ratsam erscheint, schon vor der ersten Datenerhebung an mehrsprachigen Stichproben einen größeren Aufwand in die inhaltliche und sprachliche Gestaltung der Übersetzung zu investieren – etwa durch Anwendung des „committee approach“ oder (wie z. B. bei Schwebcke, Lück & Jandron, 1973) vorherige Überprüfung an einer Rückübersetzung.

Ein Grundproblem, das durch Äquivalenzstudien wie die vorliegende nicht abschließend geklärt werden kann, ist die Frage, ob es sinnvoll ist, die der Originalversion zu-

grundlegenden persönlichkeits-theoretischen Konstrukte auch auf die Zielkultur anzuwenden (Schwebcke, Lück & Jandron, 1973, S. 440; Butcher & Pancheri, 1976, p. 29). Auf diese Frage kann erst eine Antwort gegeben werden, wenn die wichtigsten Forschungsergebnisse zur Originalversion (insbesondere was ihre Validität angeht) in der Zielkultur überprüft worden sind.

Es wird aber gerade bei Verwendung der „bilingual retest technique“ ein Ansatz geschaffen, wesentliche Teilaspekte dieses Problems zu lösen, da innerhalb der Zielkultur zu beiden Sprachversionen Daten anfallen. So zeigte sich z. B. in unserer Untersuchung, daß die Skalenhomogenität der englischen PRF-Version in einer vorwiegend aus dem deutschen Sprachraum stammenden Stichprobe weitgehend mit den bei anglo-amerikanischen Vpn festgestellten internen Konsistenzen vergleichbar sind; dieser Umstand und die relative Übereinstimmung unserer Skalenstabilitäten mit den Werten von Bentler (1964) und Hoffmann (1971) ist ein erster Hinweis darauf, daß die meisten von Jackson (1974) verwendeten Konstrukte in beiden Kulturbereichen ähnlich brauchbar sind. Wenn diese Ergebnisse auch ihre Parallelen in Untersuchungen an frankophonen (Skinner & Rampton, 1973; Skinner, Jackson & Rampton, 1976) und niederländischen (Schulken et al., 1974) Vpn haben, so bedarf es dennoch eingehender Untersuchungen mit verbesserten PRF-Übersetzungen an weniger ausgelesenen Stichproben, um die Behauptung der transkulturellen Brauchbarkeit zu erhärten. Insbesondere muß eine Validierung erfolgen sowie eine Verbreiterung der Datenbasis für faktorenanalytische Untersuchungen.

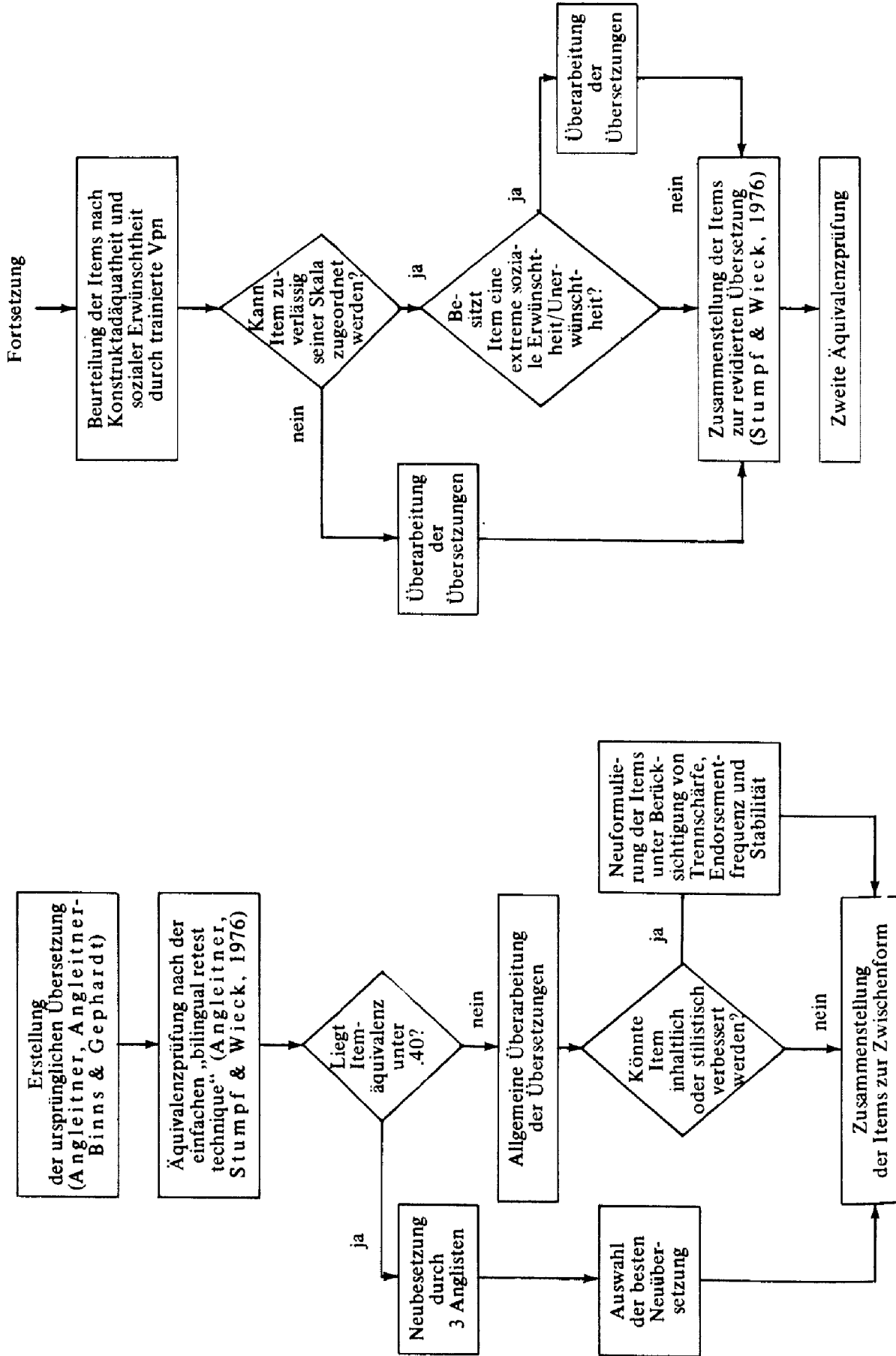
Zusammenfassung

Als erste Schritte zur Adaptierung der „Personality Research Form“ für den deutschen Sprachraum wurden die Items der Formen AA und BB ins Deutsche übertragen und einer Stichprobe zweisprachiger Personen (N = 92) zusammen mit den Originalitems zur Beantwortung vorgegeben. Die korrelationsstatische Datenauswertung auf Item- und Skalenebene ergab ein relativ hohes Maß an Übereinstimmung im Antwortverhalten in bezug auf beide Sprachversionen, jedoch auch zahlreiche Hinweise auf revisionsbedürftige Stellen in der Übersetzung. Die deutsche Version wurde daraufhin völlig überarbeitet, wobei zusätzlich die Ergebnisse von Expertenratings zur Konstruktadäquatheit und sozialen Erwünschtheit der Iteminhalte einbezogen wurden. Eine erneute Äquivalenzprüfung nach der „bilingual retest technique“ (N = 91) ließ auf eine Steigerung der Übersetzungsäquivalenz, insbesondere auf der Itemebene, schließen.

Summary

In order to obtain a German version of the "Personality Research Form", preliminary German translations of the items were prepared and submitted (together with the originals) to a sample (N = 92) of bilinguals. Correlations between the two language versions indicated a relatively high degree of translation equivalence, but also pointed to numerous weaker passages in the translation. Using the results of this study and additional data from expert ratings of content saturation and social desirability, the translations were completely revised and tested again by the bilingual retest technique (N = 91). Results indicated an increase of translation equivalence, especially on the item level.

Figur 1
Flußdiagramm der bisher durchgeführten Adaptionsschritte



Literaturverzeichnis

- Anastasi, A.: Personality Research Form. In: O. K. Buros (Ed.), *The Seventh Mental Measurements Yearbook*. Highland Park: Gryphon Press, 1972.
- Angleitner, A., Stumpf, H. & Wieck, Th.: Die „Personality Research Form“ von Jackson: Konstruktion, bisheriger Forschungsstand und vorläufige Ergebnisse zur Äquivalenzprüfung einer deutschen Übersetzung. *Wehrpsychologische Untersuchungen*, 1976, 11, Heft 3.
- Bentler, P. M.: Response variability: Fact or artifact? Unpublished doctoral dissertation, Stanford University, 1964.
- Breuer, F., Keil, W., Kleiber, D., Meier, F. & Piontkowski, U.: *Psychologie des wissenschaftlichen Lernens*. Münster: Aschendorff, 1975.
- Butcher, J. N. & Gur, R.: A Hebrew translation of the MMPI. An assessment of translation adequacy and preliminary validation. *Journal of Cross-Cultural Psychology*, 1974, 5, 220–227.
- Butcher, J. N. & Pancheri, P.: *A handbook of cross-national MMPI research*. Minneapolis: University of Minnesota Press, 1976.
- Crott, H. W. & Baltes, P. B.: Desirability of personality dimensions: Auto- and heteroperceptions by American and German college students. *Journal of Social Psychology*, 1973, 91, 15–27.
- Edwards, A. L.: *The social desirability variable in personality assessment and research*. New York: Holt, Rinehart & Winston, 1957.
- Glatt, K. M.: An evaluation of the French, Spanish, and German translations of the MMPI. *Acta Psychologica*, 1969, 29, 65–84.
- Hoffmann, H.: Dimensions of desirability ratings in a German and Austrian sample: A cross-cultural study. *Proceedings of the XVIth International Congress of Applied Psychology*, Amsterdam, 1968. Amsterdam: Scretz & Zeitlinger, 1969.
- Hoffmann, H.: Personality changes of hospitalized alcoholics after treatment. *Psychological Reports*, 1971, 29, 948–950.
- Jackson, D. N.: *Manual for the Personality Research Form*. Goshen: Research Psychologists Press, 1967 (2nd ed. 1974).
- Jackson, D. N.: A sequential system for personality scale development. In: C. D. Spielberger, (Ed.), *Current topics in clinical and community psychology*, Vol. 2. New York: Academic Press, 1970.
- Keil, W., Meier, F. & Piontkowski, U.: Dimensionen der Selbstbeschreibung in Jacksons Personality Research Form. *Diagnostica*, 1977, im Druck.
- Kelly, E. L.: Personality Research Form. In: O. K. Buros (Ed.), *The Seventh Mental Measurements Yearbook*. Highland Park: Gryphon Press, 1972.
- Loevinger, J.: Objective Test as instruments of psychological theory. *Psychological Reports*, 1957, 3, 635–694.
- Lonner, W. J.: Cross-cultural measurement of vocational interests. Unpublished doctoral dissertation, University of Minnesota, 1967.
- Mote, Th. A., Natalicio, L. F. S. & Rivas, F.: Comparability of the Spanish and English Editions of the Spielberger State-Trait Anxiety Inventory. *Journal of Cross-Cultural Psychology*, 2, 1971, 205–206.

- Murry, H. A.: Explorations in Personality. New York: Oxford University Press, 1938.
- Prince, R. & Mombour, W.: A technique for improving linguistic equivalence in cross-cultural surveys. *International Journal of Social Psychiatry*, 1967, 13, 229–237.
- Schulkens, J., van den Broek, A. & Diekstra, R. F. W.: De personality research form. *GEDRAG – tijdschrift voor psychologie*, 1974, 2, 65–101.
- Schwebcke, A., Lück, H. E. & Jandron, E.: Probleme und Erfahrungen der Adaption fremdsprachiger Tests. *Psychologische Beiträge*, 1973, 15, 434–470.
- Skinner, H. A. & Rampton, G. M.: Evaluation of the Personality Research Form (PRF) in a military environment. *Canadian Forces Personnel Applied Research Unit*, April, 1973. CFB Toronto, Downsview, Ontario.
- Skinner, H. A., Jackson, D. N. & Rampton, G. M.: The Personality Research Form in a Canadian context: Does language make a difference? *Canadian Journal of Behavioral Science*, 1976, 8, 156–168.
- Spielberger, C. D., Gonzales-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. S. & Natalicio, D. S.: Development of the Spanish edition of the State-Trait Anxiety Inventory. *Interamerican Journal of Psychology*, 1971, 5, 145–158.
- Spreen, O. & Spreen, G.: The MMPI in a German-speaking population: standardization report and methodological problems of cross-cultural interpretations. *Acta Psychologica*, 1963, 21, 365–373.
- Stricker, L. J.: Personality Research Form: Factor Structure and Response Style Involvement. *Journal of Consulting and Clinical Psychology*, 1974, 42, 529–537.
- Stumpf, H. & Angleitner, A.: Äquivalenz und Gütekriterien einer revidierten Übersetzung der „Personality Research Form“ von Jackson. *Wehrpsychologische Untersuchungen*, 1977, 12, Heft 3.
- Stumpf, H. & Wieck, Th.: Douglas N. Jackson: Personality Research Form (PRF). Revidierte Übersetzung. *Psychologisches Institut der Universität Bonn*, 1976.
- Tetenbaum, T. J.: Student needs, teacher orientations and student ratings of teachers. Unpublishing doctoral dissertation, New York University, 1974.
- Werner, O. & Campbell, D.: Translating, working through interpreters, and the problem of decentering. In: R. Navroil and R. Cohen (Ed.), *A handbook of methods in cultural anthropology*. New York: American Museum of Natural history, 1970.
- Wiggins, J. S.: Personality Research Form. In: O. K. Buros (Ed.), *The Seventh Mental Measurements Yearbook*. Highland Park: Gryphon Press, 1972.

Anschrift der Verfasser:

Dipl.-Psych. Heinrich Stumpf
Dr. Alois Angleitner
Psychologisches Institut der Universität Bonn
An der Schloßkirche 1
D-5300 Bonn

Regierungsdirektor Dr. F. W. Steege
Bundesministerium der Verteidigung
Abt. P II 4
Postfach 161
D-5300 Bonn