

- Eberle, G., Holtz, K.-L., Kowalewski, A. & Staiger, M. (1978). Ein Beitrag zur Abklärung der Faktorenstruktur des Psycholinguistischen Entwicklungstests (PET) bei sogenannten lernbehinderten Sonderschülern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 10, 134–143.
- Filipp, S.-H. & Doenges, D. (1983). Entwicklungstests. In K.-J. Groffmann & L. Michel (Hrsg.), *Intelligenz- und Leistungsdiagnostik* (S. 202–303). Göttingen, Toronto, Zürich: Hogrefe.
- Grimm, H., Schöler, H. & Wintermantel, M. (1975). *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim: Beltz.
- Hammill, D. D. & Larsen, S. C. (1974). Relationship of selected auditory perceptual skills and reading ability. *Journal of Learning Disabilities*, 7, 429–435.
- Hammill, D. D. & Larsen, S. C. (1978). The effectiveness of psycholinguistic training: A reaffirmation of position. *Exceptional Children*, 44, 402–414.
- Issen, W. B. (1978). Der psycholinguistische Entwicklungstest (PET) aus linguistischer Sicht. In G. Peuser (Hg.), *Brennpunkte der Patholinguistik* (S. 95–114). München: Fink.
- Ingenkamp, K. (1966). *Bildertest 2–3. Intelligenztest für 2. und 3. Klassen*. Weinheim: Beltz.
- Kavale, K. A. (1981 a). The relationship between auditory perceptual skills and reading ability: A meta-analysis. *Journal of Learning Disabilities*, 14, 539–546.
- Kavale, K. (1981 b). Functions of the Illinois Test of Psycholinguistic Abilities (ITPA): Are they trainable? *Exceptional Children*, 47, 496–510.
- Kavale, K. (1982). Psycholinguistic training programs: Are there differential treatment effects? *Exceptional Child*, 29, 21–30.
- Kiese, C. & Arold, R. (1984). Die Bedeutung des Psycholinguistischen Entwicklungstests (PET) von Angermaier in der phoniatriischen Diagnostik. *Sprache-Stimme-Gehör*, 8, 55–61.
- Kirk, S. A., McCarthy, J. J. & Kirk, W. D. (1968 (rev. ed.)). *Illinois Test of Psycholinguistic Abilities*. Illinois: Urbana.
- Kuusinen, J. & Leskinen, E. (1988). Latent structure analysis of longitudinal data on relations between intellectual abilities and school achievement. *Multivariate Behavioral Research*, 23, 103–118.
- Larsen, S. C., Parker, R. M. & Hammill, D. D. (1982). Effectiveness of psycholinguistic training: A response to Kavale. *Exceptional Children*, 49, 60–66.
- Lewkowicz, N. K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology*, 72, 686–700.
- Lienert, G. A. (1967). *Testaufbau und Testanalyse*. Weinheim, Berlin: Beltz, 2. Aufl.
- Lund, K., Foster, G. & McCall-Perez, F. (1978). The effectiveness of psycholinguistic training: A re-evaluation. *Exceptional Children*, 44, 310–319.
- Lundberg, I., Frost, J. & Petersen, O. P. (1988). Effects of

an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 236–284.

- Mann, V. A. & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities*, 17, 592–599.
- Marx, H. & Marx, D. (1975). *Kognitiver Stil und Sprache bei Legasthenie*. Tübingen: unveröffentl. Diplomarbeit.
- Mattingly, I. G. (1984). Reading, linguistic awareness, and language acquisition. In J. Downing & R. Valtin (eds.), *Language awareness and learning to read* (pp. 9–25). New York: Springer.
- Minskoff, E. (1975). Research on psycholinguistic training: Critique and guidelines. *Exceptional Children*, 42, 136–144.
- Müller, R. (1966). *Diagnostische Rechtschreibtests DRT2 und 3*. Weinheim: Beltz.
- Newcomer, P., Larsen, S. C. & Hammill, D. D. (1975). A response. *Exceptional Children*, 42, 144–148.
- Raven, J. C. (1956). *The Coloured Progressive Matrices. Sets A, AB, B*. London: Lewis & Co.
- Rennen-Allhoff, B. & Allhoff, P. (1987). *Entwicklungstests für das Säuglings-, Kleinkind- und Vorschulalter*. Berlin: Springer.
- Rohwer Jr., W. D. (1970). Cognitive development and education. In P. H. Mussen (ed.), *Manual of child psychology* (3. ed.) (pp. 1424–1429). New York: Wiley.
- Salvia, J. & Ysseldyke, S. E. (1981). *Assessment in spatial and remedial education*. Boston: Houghton Mifflin.
- Schubö, W. & Uehlinger, H.-M. (1986). *SPSSx Handbuch der Programmversion 2.2*. Stuttgart, New York: Gustav Fischer Verlag.
- Torgesen, J. K., Wagner, R. K., Balthazar, M., Davis, C., Morgan, S., Simmons, K., Stage, S. & Zirps, F. (1989). Developmental and individual differences in performance on phonological synthesis tasks. *Journal of Experimental Child Psychology*, 47, 491–505.
- Valtin, R. (1974). Wie kann der Lehrer mit Hilfe von Schülern eine Legasthenie feststellen? In *Fernstudien-gang Legasthenie, Studienbegleitbrief 1*, 58–86. Weinheim: Beltz.
- Wagner, R. K. & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192–212.
- Waugh, R. P. (1973). Comparison of revised and experimental editions of the ITPA. *Journal of Learning Disabilities*, 6, 236–238.
- Waugh, R. P. (1975). The ITPA: Ballast or bonanza for the school psychologist. *Journal of School Psychology*, 13, 201–208.
- Weiß, R. H. (1972). *Grundintelligenztest CFT2 Skala 2*. Braunschweig: Westermann.
- Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, 23, 159–177.

Zur Validität der 16 Persönlichkeits-Adjektivskalen (16PA)

Validity of the 16 Personality Adjective Scales

Hermann Brandstätter, Gernot Filipp & Peter Drescher

Universität Linz

Zusammenfassung: 125 Schülerinnen und 118 Schüler von Abschlussklassen höherer Schulen beurteilten in nach Zufall in den Schulklassen zusammengestellten Untergruppen von drei bis fünf Schülern sich selbst und die anderen Mitglieder der Gruppe anhand des 16PA, mit dem die Sekundärfaktoren des 16PF über 32 bipolare Eigenschaftsskalen (pro Primärfaktor zwei Skalen) geschätzt werden. Für diese Schätzungen werden, differenziert nach Geschlecht, Verteilungskennwerte und Korrelationen berichtet. Verwendet man als Validitätskriterium die gemittelten Beurteilungen durch zwei Mitschüler(innen), ergibt sich eine mittlere Korrelation von .34 (.49 bei multipler Korrelation) für die Sekundärfaktoren. Bedingt durch die größere Varianz konsistenter Beurteilungen sind deren Validitätskoeffizienten im Vergleich zu den weniger stark streuenden inkonsistenten Beurteilungen im Mittel höher (.50 vs. .38). Die Beurteilungen einer Person gelten dann als konsistent, wenn deren Urteilsprofil auf den 16PA-Dimensionen der Form A hoch korreliert mit dem Urteilsprofil der Form B. Die Werte des Leistungsmotivationstests (LMT) sind aufgrund des 16PA bei den Schülern besser rekonstruierbar als bei den Schülerinnen. Zur Erklärung der Schulnoten tragen der Leistungsmotivationstest (vor allem Subtest L2 «Ausdauer und Fleiß») und der 16PA (vor allem der Sekundärfaktor QV- «geringe Kontaktbereitschaft») mehr bei als die beiden kognitiven Leistungstests (Textanalyse und Zahlenverbindungs-test). Die Ergebnisse der Validierungsstudie sprechen für die Verwendbarkeit des 16PA in Experimenten und Feldstudien, in denen Persönlichkeitsunterschiede der Vpn in den Grunddimensionen (Normgebundenheit, Belastbarkeit, Unabhängigkeit, Entschlußbereitschaft und Kontaktbereitschaft) von Interesse sind, aber nur wenig Zeit zu ihrer Erfassung zur Verfügung steht.

Abstract: 243 high school students (125 female, 118 male), mainly 17 to 19 years of age, rated themselves and up to four randomly selected class mates on the 16PA, an adjective version of the 16PF which provides two pairs of adjectives for each of the 16 first order factors of the 16PF and allows a reliable estimation of the 16PF second order factors. The median correlation between self ratings and peer ratings (averaged over two class mates) is $r = .34$ ($R = .49$ with multiple correlations) for the second order factors (norm orientation, emotional stability, independence, decision readiness, and extraversion). Higher consistency of self ratings, measured as correlation of a person's profile of 16PA Form A and Form B, is connected with higher validity caused by the larger variance of consistent ratings. The scores of the achievement motivation test (LMT) can be partially predicted from the 16PA second order factor scores, better for male than for female students. 16PA and LMT scales are valid predictors of the grade point average. The results suggest that the 16PA can be used in experiments and field studies for assessing basic personality dimensions in short time.

In Experimenten, Feldstudien und Umfragen wäre es oft von Interesse, nicht nur Alter, Geschlecht und Bildungsniveau, sondern auch die Persönlichkeitsstruktur der befragten oder beobachteten Personen zu kennen, um die Generalisierbarkeit der Untersuchungsergebnisse besser abschätzen zu können. Die üblichen mehrdimensionalen Persönlichkeitsfragebogen beanspruchen aber zu viel Zeit. Wählt man aus einem umfassenden Persönlichkeitsinventar nur die Items einer für den Untersuchungszweck besonders relevanten Dimension aus, läuft man

Gefahr, daß die Antworten nicht mit denen vergleichbar sind, die man unter Standardbedingungen (Vorgabe der Items dieser Skala im Kontext der Items der übrigen Skalen) erhalten hätte. Verzichtet man überhaupt auf geeichte und validierte Meßinstrumente in der Hoffnung, das relevante Persönlichkeitsmerkmal durch eine eigens für die betreffende Fragestellung konstruierte Skala präziser und eindeutiger erfassen zu können, vermehrt man nur die schon lange nicht mehr überschaubare Vielfalt von Dimensionen, deren Reliabilität und Validität

tät ungeklärt ist. Zumindest empfiehlt sich, eine speziell für eine bestimmte Studie konstruierte Skala zusammen mit einem theoretisch fundierten und empirisch bewährten umfassenden Beschreibungssystem zu verwenden. Nur so läßt sich beurteilen, ob die angeblich neue Dimension nicht aus den bekannten Grunddimensionen der Persönlichkeit rekonstruiert werden kann. Wenn in letzter Zeit häufiger von Varianten derselben fünf Grunddimensionen («Big Five») die Rede ist (McCrae & Costa, 1985; Digman, 1990; Trapnell & Wiggins, 1990; Borkenau & Ostendorf, 1991), die in Faktorenanalysen verschiedenster Persönlichkeitsfragebogen immer wieder auftauchen, kann man hoffen, daß sich endlich ein System der Persönlichkeitsbeschreibung durchsetzt, das Verständigung zwischen den Forschern und damit auch Vergleichbarkeit von Pbn-Stichproben hinsichtlich der in ihnen erfaßten Persönlichkeitsstrukturen und Vergleichbarkeit der differentialpsychologischen Befunde verschiedener Studien gewährleistet. Damit ergäbe sich auch eher die Chance, vermeintlich allgemeinspsychologische Gesetzmäßigkeiten als das zu erkennen, was sie tatsächlich oft sind: durchschnittspsychologische Aussagen, die funktionale Zusammenhänge bestenfalls für eine differentialpsychologisch identifizierbare Mehrheit von Vpn, nicht jedoch für eine ebenfalls differentialpsychologisch identifizierbare Minderheit darstellen. Auch ließen sich Widersprüche in den Befunden verschiedener Experimente zu denselben Fragestellungen besser aufklären. Nicht wenige Replikationsversuche von Hypothesenprüfungen dürften ja daran scheitern, daß eine Untersuchung mit einer hinsichtlich der Persönlichkeitsmerkmale unbemerkt anders zusammengesetzten Stichprobe wiederholt wird. Eine unterschiedliche Zusammensetzung kann sich zufällig oder als unbeachtete Folge eines bestimmten Vorgehens bei der Vpn-Gewinnung ergeben. Unter solchen Umständen sollte man sich nicht wundern, wenn die früheren Ergebnisse oft nicht replizierbar sind.

In diesem Beitrag soll über weitere Erfahrungen mit dem 16PA berichtet werden, insbesondere über einen Vergleich von Selbstbeurteilungen mit Beurteilungen durch Mitschüler. Der 16PA (16-Persönlichkeits-Adjektivskalen) ist eine von Brandstätter (1988a) entwickelte Adjektivform

des 16PF (Schneewind, Schröder & Cattell, 1983), die in allen Experimenten und Feldstudien verwendet werden kann, in denen man Persönlichkeitsmerkmale berücksichtigen möchte, ohne viel Zeit für ihre Erfassung verwenden zu können. Die Wahl fiel auf Cattell's System, weil es zu den international am häufigsten verwendeten Verfahren gehört, in einer sorgfältig adaptierten deutschen Fassung vorliegt (Schneewind et al., 1983) und Beschreibungen auf zwei Abstraktionsniveaus ermöglicht, dem von sechzehn Primär- und dem von fünf Sekundärfaktoren. Es sind insbesondere die Sekundärfaktoren, die den auch in anderen Beschreibungssystemen (Eysenck & Eysenck, 1975; Costa & McCrae, 1985) erscheinenden oder aus ihnen ableitbaren Grunddimensionen entsprechen. Für das Cattellsche System sprach auch, daß es sich in den Befindensstudien, die seit Mitte der 70er Jahre mit dem Zeitstichproben-Tagebuch nach Brandstätter (1977) durchgeführt wurden, gut bewährt hat (vgl. dazu u. a. Brandstätter, 1990). Im übrigen sprechen auch die Überprüfungen der Reliabilität und Validität des 16PF durch Borkenau (1986) und Krampen (1986) für die Brauchbarkeit des Verfahrens, auch wenn es nicht an einigen kritischen Anmerkungen fehlt (Bartussek, 1988).

Methode

1.1 Beurteilungsverfahren

Für jede der 16PF-Skalen wurden zwei Eigenschaftspaare derart konzipiert, daß sie möglichst genau die Erlebnis- und Verhaltensweisen charakterisieren, die in den zwölf Items der betreffenden Skalen zum Ausdruck kommen. Erste Entwürfe der Eigenschaftsliste wurden von mehreren Kollegen hinsichtlich ihrer sprachlichen und psychologischen Bedeutung auf Plausibilität überprüft und daraufhin modifiziert. Einige Veränderungen wurden dann noch aufgrund von Verteilungs- und Korrelationskennwerten erster Erprobungen vorgenommen. Die Eigenschaftsliste ist in Brandstätter (1988a) abgedruckt.

Die Vpn haben auf einer neunstufigen Skala anzugeben, in welchem Grad die eine oder die andere Eigenschaft auf sie zutrifft. Die Liste

umfaßt 32 Eigenschaftspaare, zwei für jeden Primärfaktor. Basierend auf Regressionsanalysen mit Daten von Vpn, die sowohl den 16PA als auch den 16PF bearbeitet haben (Brandstätter, 1988a), können für jeden Probanden die Stenwerte der Sekundärfaktoren des 16PF geschätzt werden.

1.2 Datenerhebung

Im Februar und März 1989 wurden Schülerinnen und Schüler von Abschlußklassen höherer Schulen für Zwecke der Testanalyse mit einer Versuchsform des BIS (Jäger, 1984), mit zwei weiteren kognitiven Leistungstests (Textanalyse, einem unveröffentlichten Test der Deutschen Gesellschaft für Personalwesen, der das Erfassen des wesentlichen Gehalts von komplexen Texten prüft, und Zahlenverbindungstest; Oswald & Roth, 1987) sowie mit dem Leistungsmotivationstest LMT (Hermans, Petermann & Zielinski, 1978) und dem Person-Umwelt-Strukturtest PUST (Eder & Bergmann, 1988) untersucht. Bei dieser Gelegenheit wurden die Teilnehmer auch gebeten, nach der Bearbeitung der Denkaufgaben, die mit einer Pause von 10 Minuten insgesamt ca. 180 Minuten dauerten, sich selbst und zwei bis vier Mitschüler(innen) anhand der 16PA-Skalen zu beurteilen. Dieser Bericht bezieht sich in der Hauptsache auf die 16PA-Beurteilungen. Mit berücksichtigt werden noch die Daten des LMT und die Noten des letzten Jahreszeugnisses, die von den Untersuchungsteilnehmern erfragt wurden. Die Noten-

angaben konnten zwar nicht an den Zeugnissen überprüft werden, frühere Erfahrungen sprechen aber dafür, daß man sich auf solche Angaben durchaus verlassen kann. Da die Analyse des BIS und des PUST in einer eigenen Arbeit erfolgen soll, bleiben diese Daten hier außer Betracht.

Die gegenseitigen Beurteilungen erfolgten in strikt nach dem Zufallsprinzip zusammengestellten Gruppen von drei bis fünf Schülern. Die Gruppen waren in gemischten Klassen hinsichtlich des Geschlechts nicht homogen, sondern zufallsabhängig gemischt. Die Schüler(innen) beurteilten stets sich selbst zuerst, die anderen in beliebiger Reihenfolge. Es wurde darauf geachtet, daß die Beurteilungen nicht von den anderen eingesehen werden konnten. Strikte Anonymität (Verwendung eines Zahlen- und Buchstabencodes) wurde glaubhaft gemacht und eingehalten.

Die Untersuchung erfolgte mit Zustimmung der Landesschulbehörde und der Schuldirektion. Die Eltern der Schüler wurden rechtzeitig darüber informiert. Für die Schüler war die Teilnahme freiwillig. Sieben Schulklassen nahmen in der regulären Schulzeit vollzählig daran teil. Aus sechs Klassen kamen Gruppen von Schülern zu Testterminen, die außerhalb der Schulzeit angesetzt waren. Insgesamt haben sich 254 Schülerinnen und Schüler an der Untersuchung beteiligt. In Tabelle 1 findet sich die Häufigkeitsverteilung nach Geschlecht, Alter und Schultyp. Für die folgenden Analysen wurden nur die Daten von 243 Schülern (125 Schülerinnen und 118

Tabelle 1: Häufigkeitsverteilung der demographischen Merkmale Geschlecht, Alter und Schultyp in der Gesamtstichprobe (N = 254)

Alter	Schülerinnen				Schüler			
	16 - 17	18 - 19	20 - 21	Summe	16 - 17	18 - 19	20 - 21	Summe
Schultyp								
(1) HAK	0	36	3	39	0	25	7	32
(2) HAS	33	2	3	35	2	0	0	2
(3) AHS	20	40	0	60	23	33	0	56
(4) HTL	0	0	0	0	0	25	5	30
Summe	53	78	3	134	25	83	12	120

Anmerkung. HAK = Handelsakademie HAS = Handelsschule AHS = Allgemeinbildende Höhere Schule HTL = Höhere Technische Lehranstalt

Schüler) verwendet, für die die Selbstbeurteilung und mindestens zwei Fremdbeurteilungen vorliegen. Die dritte und vierte Fremdbeurteilung, die nur bei einem Teil der Vpn vorhanden war, wurde hier nicht berücksichtigt.

2. Ergebnisse

Die Ergebnisdarstellung ist in drei Abschnitte unterteilt. Zuerst werden die Verteilungskennwerte und Paralleltestkorrelationen der aufgrund der beiden Formen des 16PA geschätzten Sekundärfaktoren berichtet. Dabei werden auch Unterschiede zwischen den Geschlechtern aufgezeigt. Dann geht es um die Übereinstimmung verschiedener Beurteiler in der Beurteilung derselben Person. In diesem Zusammenhang wird auch geprüft, ob die Validität der Selbstbeurteilungen, d.h. die Übereinstimmung zwischen Selbst- und Fremdbeurteilung, mit der Konsistenz der Selbstbeurteilungen wächst. Schließlich wird untersucht, was die 16PA-Beurteilungen zur Erklärung der LMT-Skalen und der Schulnoten beitragen.

Tabelle 2: Regressionsgewichte b und additive Konstante a zur Schätzung der 16PF-Sekundärfaktoren QI bis QV aufgrund der 16PA-Skalen (separat geschätzt nach Form A bzw. Form B) (N = 228)

	QI		QII		QIII		QIV		QV	
	A	B	A	B	A	B	A	B	A	B
A	.02	-.06	.05	.13	.02	-.11	-.11	-.00	.27	+.08
B	-.06	-.12	.20	.10	.25	-.04	-.01	-.22	-.19	-.22
C	-.01	-.02	.17	+.16	-.02	-.07	.03	-.01	-.02	-.08
E	-.12	-.11	-.02	.00	.12	+.30	.02	-.12	-.05	-.00
F	-.17	-.02	-.00	-.01	.18	+.08	-.01	-.09	.13	+.21
G	.39	+.35	.02	.18	.03	.01	.22	.05	-.18	-.05
H	.16	.18	.07	+.27	.21	+.31	.01	.13	-.01	.17
I	-.04	.10	.08	-.05	-.02	-.11	-.09	-.03	-.01	.01
L	.11	.06	-.04	-.07	.07	.12	.05	.12	-.09	-.12
M	-.11	-.16	-.08	-.00	.03	.13	-.11	-.22	-.10	-.06
N	-.06	+.01	.08	.07	.00	.00	.06	.03	.14	-.05
O	-.02	-.03	-.19	-.15	-.09	-.01	-.10	-.03	-.07	-.12
Q1	-.29	-.25	.05	.05	.05	+.07	-.19	-.06	.04	.06
Q2	-.04	-.04	.16	.01	.03	-.03	-.01	.02	-.09	-.24
Q3	.15	+.09	.16	+.08	.06	-.01	-.08	-.09	.13	-.10
Q4	.06	.08	-.10	-.12	.05	.07	-.04	-.10	.18	.04
a	3.54	3.84	1.04	.94	.16	1.50	6.45	8.49	6.00	8.75
R	.74	.69	.72	.65	.58	.59	.43	.43	.56	.58
R ²	.51	.44	.48	.37	.29	.30	.12	.13	.27	.28

Anmerkung. + - Diese Primärfaktoren haben bei Schneewind et al. (1983, S. 17) die höchsten positiven oder negativen Ladungen auf den jeweiligen Sekundärfaktoren. R² korrigierter Determinationskoeffizient.

2.1 Vergleich der beiden Formen des 16PA

Um Paralleltestkorrelationen auch für die Sekundärdimensionen bestimmen zu können, wurden nachträglich gesondert für die beiden Formen die Gewichte der Primärfaktoren des 16PA zur Schätzung der Sekundärfaktoren des 16PF mit den Daten derselben Stichprobe (n = 228) ermittelt, für die Brandstätter (1988a, S. 378) die für die Mittelwerte aus den beiden Formen des 16PA berechneten Gewichte berichtet. Form A besteht aus den ersten, Form B aus den zweiten 16 Items der Eigenschaftsliste.

Tabelle 2 enthält die Regressionsgewichte und die multiplen Korrelationen. Zum Vergleich sind jene Primärfaktoren markiert, die bei Schneewind et al. (1983, S. 17) die höchsten Ladungen auf den jeweiligen Sekundärfaktoren aufweisen. Davon weichen hier ab: N in QI, L in QII, A und Q1 in QIII sowie A und I in QIV. QIV erweist sich auch bei getrennter Analyse der beiden Formen des 16PA nur als unzulänglich rekonstruierbar.

In Tabelle 3 findet man, differenziert nach Geschlecht, die Mittelwerte, Standardabweichungen und Paralleltest-Korrelationen der

Tabelle 3: Mittelwerte, Standardabweichungen und Paralleltest-Korrelationen der 16PA-Sekundärfaktoren für Selbst- und Fremdbeurteilungen, differenziert nach Geschlecht

		Selbstbeurteilungen					Fremdbeurteilungen				
		QI	QII	QIII	QIV	QV	QI	QII	QIII	QIV	QV
Schülerinnen n = 125											
Form A	A. M.	3.5	5.0	5.7	4.2	6.6	3.7	4.8	5.6	4.5	6.6
	S. D.	1.3	1.4	1.2	.8	1.0	1.5	1.6	1.4	1.0	1.2
Form B	A. M.	3.4	5.0	5.6	4.5	6.5	4.3	4.8	5.3	5.1	6.5
	S. D.	1.2	1.3	1.3	.9	1.3	1.4	1.4	1.6	.9	1.4
r(A, B)		.46	.73	.59	.22	.53	.56	.74	.69	.32	.49
Schüler n = 118											
Form A	A. M.	3.8	5.7	5.9	4.7	6.3	3.7	4.8	5.4	4.9	6.3
	S. D.	1.2	1.2	1.3	.8	1.1	1.5	1.5	1.4	1.1	1.2
Form B	A. M.	3.7	5.5	6.0	4.6	6.1	4.2	5.1	5.7	5.4	6.5
	S. D.	1.3	1.3	1.2	.7	1.2	1.3	1.3	1.5	.9	1.4
r(A, B)		.29	.60	.72	.12	.50	.44	.69	.66	.27	.58
Schülerinnen											
gesamt	A. M.	3.3	5.0	5.7	4.1	6.5	4.0	4.6	5.4	4.6	6.6
	S. D.	1.2	1.3	1.4	.9	1.3	1.5	1.6	1.6	.9	1.3
Schüler											
gesamt	A. M.	3.7	5.6	5.9	4.6	6.3	4.0	4.7	5.5	5.1	6.5
	S. D.	1.2	1.3	1.4	.8	1.3	1.3	1.5	1.5	.9	1.3

Anmerkung. *Hier ist der Unterschied zwischen Schülerinnen und Schülern signifikant (p < .01). r(A, B) Paralleltestkorrelation. Bei den Fremdbeurteilungen wurde jeweils nur der erste Beurteiler berücksichtigt.

fünf Sekundärdimensionen der beiden Formen A und B für die Selbstbeurteilungen und die Beurteilungen durch andere. Die entsprechenden Werte der Primärdimensionen wurden, um Platz zu sparen, nicht in diesen Bericht aufgenommen. Es ist aber erwähnenswert, daß sich Schülerinnen von Schülern in einer Reihe von Primärdimensionen signifikant (p < .01) voneinander unterscheiden.

Schülerinnen sehen sich, verglichen mit gleichaltrigen Schülern, in höherem Maße als kontaktfreudig und warmherzig (A), leicht zu beunruhigen und seelisch wenig belastbar (C-), feinfühlig und zartbesaitet (I), träumerisch und phantasievoll (M), geradeheraus und natürlich (N-), an sich zweifelnd und ängstlich besorgt (O). Mit einer Ausnahme (Variable M in der Gruppe HAK) finden sich diese Unterschiede in allen Schultypen. In den Beurteilungen durch andere treten ähnliche Unterschiede, wenn auch etwas abgeschwächt, in Erscheinung. Wie Tabelle 3 zeigt, gibt es bei den Sekundärfaktoren den deutlichsten Geschlechtsunterschied in QII

(Belastbarkeit). Schülerinnen beschreiben sich als weniger belastbar (5.0 vs. 5.6).

2.2 Übereinstimmung von Urteilen über dieselbe Person

Für die in diesem Abschnitt berichteten Korrelationen wurden die einander entsprechenden Skalenwerte der beiden Formen gemittelt.

Auf der Ebene der Primärfaktoren sind die Korrelationen zwischen den Selbstbeurteilungen und den über zwei Beurteiler gemittelten Beurteilungen durch andere am höchsten (r > 0.40) in den Aspekten A, F, G und H. Auf der Ebene der Sekundärfaktoren betragen die Korrelationen r(QI) = .34, r(QII) = .20, r(QIII) = .51, r(QIV) = .32, r(QV) = .48. Die entsprechenden Korrelationen zwischen den beiden Mitschülerbeurteilungen lauten der Reihe nach .37, .30, .55, .25 und .46.

Verwendet man zur Vorhersage der von den Mitschülerurteilen gewonnenen Werte auf den Sekundärdimensionen alle fünf Selbstbeurtei-

Tabelle 4: Differentielle Vorhersagbarkeit der Beurteilungen durch andere (aufgrund der Selbstbeurteilungen) und der Selbstbeurteilungen (aufgrund der Beurteilungen durch andere)

	QIF		QIIF		QIIIF		QIVF		QVIF	
Konsistenz	R	'R ²	R	'R ²	R	'R ²	R	'R ²	R	'R ²
hoch (n = 131)	.59	.32	.29	.05	.60	.33	.52	.25	.56	.29
niedrig (n = 112)	.41	.13	.18	.00	.44	.16	.42	.14	.49	.20
gesamt	.49	.22	.21	.02	.53	.26	.46	.19	.52	.25

	QIS		QIIS		QIIIS		QIVS		QVS	
Konsistenz	R	'R ²	R	'R ²	R	'R ²	R	'R ²	R	'R ²
hoch (n = 131)	.43	.16	.33	.07	.60	.33	.39	.12	.54	.26
niedrig (n = 112)	.33	.07	.22	.00	.44	.16	.31	.05	.46	.18
gesamt	.36	.11	.27	.06	.52	.26	.33	.09	.51	.24

Anmerkung. Die Gruppen wurden am A. M. ($r = .38$) der Profilkorrelationen (Form A mit Form B der Selbstbeurteilungen) geteilt. R multiple Korrelation 'R² korrigierter Determinationskoeffizient. QIF bis QVF Sekundärfaktoren der Fremdbeurteilung als abhängige Variablen. QIS bis QVS Sekundärfaktoren der Selbstbeurteilung als abhängige Variablen.

lungsdimensionen, erhöhen sich die (multiplen) Korrelationen vor allem im Faktor QI (von .34 auf .49) und QIV (von .32 auf .46). Diese Befunde gelten im wesentlichen übereinstimmend für Schülerinnen und Schüler, deren Daten in einer Analyse zusammengefaßt werden konnten, da sich die Varianz-Kovarianz-Matrizen als homogen erwiesen.

2.3 Differentielle Vorhersagbarkeit

Da jeder der 16 Primärfaktoren des 16PF im 16PA durch zwei Eigenschaftspaare vertreten ist, kann man pro Person die Übereinstimmung des Eigenschaftsprofils der Form A mit dem Eigenschaftsprofil der Form B durch Korrelation (Pearson) der Profile bestimmen. Der Korrelationskoeffizient kann als Indikator einer personspezifischen Konsistenz der Selbstbeurteilungen interpretiert werden, die bei den Selbstbeurteilungen im Mittel $r = .38$ bei einer Standardabweichung von $s_r = .22$ beträgt. Es zeigt sich sehr deutlich (Tabelle 5), daß höhere Konsistenz der Selbstbeurteilungen mit höheren Validitätskoeffizienten verbunden ist.

Höhere Validitätskoeffizienten ergeben sich auch mit höherer Konsistenz der Fremdbeurteilungen. Wenn beide Fremdbeurteilungen niedrige Konsistenz aufweisen ($n = 62$), sind die multiplen Korrelationen für die fünf Sekundärfaktoren der Beurteilungen durch Mitschüler (vorhergesagt mit optimaler Gewichtung der fünf Faktoren der Selbstbeurteilung) der Reihe nach .37, .21, .40, .48 und .39. Haben beide Fremdbeurteilungen hohe Konsistenz ($n = 85$), lauten die Werte .64, .25, .69, .52 und .66. Ist die Konsistenz bei einer Fremdbeurteilung niedrig, bei der anderen hoch, liegen die Werte dazwischen.

Erwähnenswert ist, daß eine gemeinsame Faktorenanalyse der Konsistenzmaße der drei Beurteilungen (eine Selbstbeurteilung und zwei Fremdbeurteilungen) und der Einstufungen der Zuverlässigkeit der Urteile durch die jeweils drei Beurteiler zwei Faktoren ergibt, einen mit hohen Ladungen in den drei Konsistenzmaßen (Faktor «Konsistenz»), einen mit hohen Ladungen in den drei Zuverlässigkeitsschätzungen (Faktor «Zuverlässigkeit»). Diese eignen sich im übrigen, anders als die Konsistenzmaße, für die Dif-

Tabelle 5: Mittelwerte, Standardabweichungen und Korrelationen von Selbst- und Fremdbeurteilungen: Sekundärfaktoren in vereinfachter Schätzung

	Selbstbeurteilung		Fremdbeurteilung		r(S', S)	r(F1, F2)	r(F, S)		R(F, S1,...,S5)	
	A.M.	S.D.	A.M.	S.D.			Konsistenz niedrig	Konsistenz hoch	Konsistenz niedrig	Konsistenz hoch
QI	4.54	.86	4.88	.88	.79	.41	.24	.50	.42	.62
QII	5.43	1.07	5.34	.91	.84	.36	.10	.36	.22	.37
QIII	6.00	1.02	5.52	1.12	.89	.59	.38	.52	.40	.56
QIV	3.74	1.19	4.49	1.16	.74	.41	.41	.44	.45	.50
QV	5.86	1.26	5.52	1.18	.81	.52	.47	.56	.53	.58

Anmerkung. N = 243. r(S', S) Korrelation der vereinfachten Schätzung der Sekundärfaktoren der Selbstbeurteilungen mit der Schätzung nach Regressionsgewichten. R(F1, F2): Korrelationen der beiden Fremdbeurteilungen; r(F, S): Korrelationen der gemittelten Fremdbeurteilungen mit den Selbstbeurteilungen. r(F, S1,...,S5): Korrelation (unkorrigiert) der gemittelten Fremdbeurteilungen mit optimal gewichteten Linearkombinationen der fünf Selbstbeurteilungsskalen. Differenzierung nach Konsistenz wie in Tabelle 4.

ferenzierung der Validität ganz und gar nicht. Vermutlich hätten Aussagen der Vpn darüber, wie angemessen die einzelnen Eigenschaftspaare zur Beschreibung ihrer Verhaltensweisen erscheinen, wie stabil diese gegenüber wechselnden Umständen sind und wie gut sie von anderen beobachtet werden können, mehr zu einer Differenzierung der Validität beigetragen (vgl. dazu Amelang & Bartussek, 1990, S. 544–557 und die dort zitierte Literatur). Verwendet man den Mittelwert im Faktor «Konsistenz» zur Abgrenzung von zwei Konsistenzgruppen, ergeben sich bei niedriger Konsistenz ($n = 112$) multiple Validitätskoeffizienten (Vorhersagbarkeit der einzelnen Sekundärfaktoren der Fremdbeurteilung aufgrund der fünf Sekundärfaktoren der Selbstbeurteilung) von .38, .20, .49, .44 und .35, bei hoher Konsistenz ($n = 131$) .60, .28, .62, .51 und .63. Für alle hier verwendeten Konsistenzmaße gilt, daß sie keine signifikanten Korrelationen mit den Faktorwerten (QI bis QV) aufweisen.

Die Varianz der Faktorwerte (Selbst- und Fremdbeurteilung) sind bei hoher Konsistenz der Selbstbeurteilungen im Durchschnitt höher als bei niedriger Konsistenz ($s^2(\text{hoch})/s^2(\text{niedrig}) = 1.40$). Die Varianz der Schätzfehler (Residualvarianz) unterscheidet sich jedoch in den beiden Konsistenzstufen nicht voneinander. Dies ist ein deutlicher Hinweis darauf, daß die

Ursache für die niedrigeren Validitätskoeffizienten bei niedriger Konsistenz in der reduzierten Streuung zu suchen ist.

Amelang und Borkenau (1982) verwendeten als Maß der personspezifischen Urteilkonsistenz die intraindividuelle Varianz der Meßwerte über die Skalen mit den jeweils höchsten Ladungen in der betreffenden Dimension. Analog dazu haben wir außer dem generellen Konsistenzmaß auch noch die Brauchbarkeit eines dimensionsspezifischen Maßes überprüft. Dieses bestand in der absoluten Differenz zwischen den beiden PA-gestützten Schätzungen der fünf Sekundärfaktoren des 16PF. Die eine Schätzung erfolgte mit den durch eine Reanalyse der Daten von Brandstätter (1988a) ermittelten Regressionsgewichten für Form A des 16PA, die andere stützte sich auf die für Form B ermittelten Gewichte. Die so bestimmte dimensionsspezifische Konsistenz erwies sich nur im Faktor QIII (Unabhängigkeit) als klar validitätsdifferenzierend mit Werten von $r(81) = .63$ bei hoher, $r(95) = .39$ bei mittlerer und $r(67) = .27$ bei geringer Konsistenz. Der hier verwendete dimensionsspezifische Konsistenzindex weicht von Laning's (1988) Index für Skalierbarkeit insofern ab, als er mögliche Mittelwertsdifferenzen zwischen den beiden Formen in den geschätzten Faktorwerten, die ja in der Normgruppe definitionsgemäß gleich Null sind, außer acht läßt.

2.4 Eine alternative Gewichtung der 16PA-Skalen

Die Sekundärfaktoren des 16PF wurden mit Gewichten geschätzt, die in einer Regressionsanalyse ermittelt wurden, zu der mehrere Gruppen von Pbn, die sowohl den 16PA als auch den 16PF beantwortet hatten, die Daten beisteuerten. Im folgenden soll nun die Tragfähigkeit einer einfacheren Gewichtung erprobt werden. Für diesen Zweck wählten wir für jeden Sekundärfaktor jene Primärskalen aus, die bei Schneewind et al. (1983, S. 17) die höchsten Gewichte aufweisen, und gewichteten diese Skalenergebnisse bzw. - bei negativen Ladungen - die Differenzen «10 minus Skalenergebnis» mit 1, die übrigen jeweils mit 0. Die ausgewählten Skalen sind auch in Tabelle 2 gekennzeichnet. Die vereinfachte Schätzung der 16PF-Sekundärfaktoren erfolgte demnach auf folgende Weise:

$$\begin{aligned}
 QI &= (10 - B + G + 10 - M + N + 10 - Q_1 + Q_3)/6 \\
 QII &= (C + H + 10 - L + 10 - O + Q_3 + 10 - Q_4)/6 \\
 QIII &= (A + E + F + H + L + Q_1)/6 \\
 QIV &= (10 - A + 10 - I + 10 - M)/3 \\
 QV &= (A + F + 10 - Q_2)/3
 \end{aligned}$$

Die Korrelationen zwischen den beiden Berechnungsformen der Selbstbeurteilungs-Sekundärfaktoren sind hoch genug (zwischen .74 und .89), daß sie als äquivalent gelten können. Dies ist ein weiterer Hinweis darauf, daß die 16PA-Skalen den Gehalt der 16PF-Primärfaktoren im großen und ganzen gut getroffen haben (Tabelle 5).

Tabelle 6: Interkorrelationen der Sekundärfaktoren der Selbst- und Fremdbeurteilung (1 Beurteiler)

Selbstbeurteilung					Fremdbeurteilung						
QI	QII	QIII	QIV	QV	QI	QII	QIII	QIV	QV		
QI	-	.06	-.42	.39	-.39	QI	-	-.04	-.68	.44	-.58
QII	.03	-	.23	.15	.13	QII	-.07	-	.30	-.07	.17
QIII	-.14	.34	-	-.13	.44	QIII	-.41	.29	-	-.26	.54
QIV	.60	.15	-.19	-	-.44	QIV	.72	-.07	-.33	-	-.66
QV	-.17	.08	.27	-.06	-	QV	-.41	-.13	.24	-.33	-

Anmerkung. Regressionsanalytische Berechnung der Sekundärfaktoren unterhalb der Diagonale, vereinfachte Berechnung oberhalb der Diagonale.

Die Validitätskoeffizienten liegen nach der vereinfachten Berechnung der Sekundärfaktoren im Mittel bei .42 und sind damit etwas höher als bei der Berechnung mit Regressionsgewichten, wo der Median der Validitätskoeffizienten bei .34 liegt. Die Interkorrelationen der Sekundärfaktoren sind, wie nicht anders zu erwarten, insbesondere bei den Fremdbeurteilungen ebenfalls etwas höher (Tabelle 6). Die weiteren Analysen werden nur für die ursprüngliche Schätzung der Sekundärfaktoren berichtet, da die vereinfachte Berechnung zu ganz ähnlichen Ergebnissen führt.

2.5 Rekonstruktion der LMT-Werte aus den 16PA-Skalen

Die Möglichkeit, neue Skalen zur Messung von Persönlichkeitseigenschaften zu konstruieren, die für den einen oder anderen Zweck besonders geeignet sein sollen, erscheinen unerschöpflich. Tatsächlich herrscht ja auch kein Mangel an neuen Kreationen. Es fragt sich nur, ob dadurch nicht mehr Verwirrung als Klärung geschaffen wird. Immerhin macht sich in den letzten Jahren ein gewisser Überdruß an dem Vielerlei der Beschreibungssysteme für Persönlichkeitsmerkmale bemerkbar. Die Belege dafür, daß Persönlichkeitsbeurteilungen, wie man sie auch dreht und wendet mag, nicht über vier bis fünf Grunddimensionen hinauskommen, sind immer weniger zu übersehen (Digman, 1990).

In der Schuluntersuchung, aus der unsere Daten stammen, wurde auch der LMT (Hermans

et al., 1978) verwendet, weil wir erwartet hatten, daß Schulleistungen, die immer wieder Anstrengungen und beständige Selbstdisziplin in den Vorbereitungen auf die Schularbeiten verlangen, in höherem Maße als Testleistungen, bei denen die Situation zu zeitlich sehr begrenzten Höchstleistungen anspornt, von der Leistungsmotivation abhängen. Schulnoten repräsentieren so eher, was ein Schüler im längeren zeitlichen Verlauf durchschnittlich leistet (leisten kann und leisten will), Testleistungen dagegen, was ein Schüler bei maximaler Anstrengung leisten kann (vgl. die Unterscheidung von typical und maximal performance bei Sackett, Zedeck & Fogli, 1988).

Im folgenden soll nun versucht werden, die LMT-Skalen mit Hilfe der 16PA-Skalen (Selbstbeurteilung) zu rekonstruieren und so zu prüfen, wie weit der LMT Leistungskomponenten erfaßt, die im 16PA nicht enthalten sind.

Schon bei flüchtiger Durchsicht der Items des LMT fällt auf, daß die Skala F- (leistungshemmende Prüfungsangst; in unserer Bezeichnung LMT4) hinsichtlich der Bedeutung der Items große Ähnlichkeit mit der Skala QII- (Mangel an Belastbarkeit) aufweist. Es stellt sich darüber

hinaus die Frage, wie weit sich die übrigen im LMT erfaßten Verhaltensaspekte mit denen des 16PA decken. Daß zwischen den Sekundärdimensionen des 16PF und den Ausprägungsgraden und Strukturmustern der Motive ein Zusammenhang besteht, hat kürzlich Brandstätter (1990) anhand von Daten des Zeitstichprobentagebuchs gezeigt.

Aus Tabelle 7 geht hervor, daß der Zusammenhang zwischen LMT- und 16PA-Dimensionen bei Schülern enger ist als bei Schülerinnen. Am besten läßt sich erwartungsgemäß die Skala F- des LMT (LMT4) rekonstruieren (R = .61 bei Schülern, .46 bei Schülerinnen), am schlechtesten F+ (LMT3) (.39 bzw. .26). Der Gesamtwert der Leistungsmotivation (LMT1+LMT2+LMT3-LMT4) ist aufgrund der 16PA-Sekundärfaktoren (insbesondere QII und QIII) bei Schülern mit R = .64 und bei Schülerinnen mit R = .43 zu rekonstruieren. In der Gruppe der Schüler(innen), die sich selbst konsistent beurteilen (n = 131), ist die Vorhersagbarkeit der leistungshemmenden Prüfungsangst (F- oder LMT4) im übrigen höher als in der Gruppe (n = 112), deren Urteile eher inkonsistent (Profilkorrelation < .38) sind (R = .61 vs. R = .51).

Tabelle 7: Rekonstruktion der Leistungsmotivationswerte (LMT) mit Hilfe der Sekundärfaktoren (Selbstbeurteilung) des 16PA, differenziert nach Geschlecht (118 Schüler, 125 Schülerinnen)

Geschlecht	LMT1		LMT2		LMT3		LMT4		LMT		LMT ¹	
	m	w	m	w	m	w	m	w	m	w	m	w
QI (S)	.32**	.15	.23*	.14	-.18	.06	.13	.11	.08	.09	.12	.24
QII (S)	.21*	.08	.29**	.21*	.20*	.09	-.39**	-.22*	.41**	.26**	.04	-.00
QIII (S)	.28**	.03	.22*	.01	.21*	.19	-.35**	-.31**	.40**	.25**	.42**	.25*
QIV (S)	-.01	.04	-.06	-.01	.00	.08	-.00	-.12	-.02	.10	.18	-.17
QV (S)	-.39*	.04	-.37**	-.04	-.29**	.09	.08	.17*	-.41**	-.06	-.09	-.25*
R	.54	.19	.49	.26	.39	.26	.61	.46	.64	.43	.41	.34
R ²	.26	.00	.21	.03	.12	.03	.35	.17	.38	.15	.13	.08

Anmerkung. ¹ Als Prädiktoren wurden hier die 16PA-Sekundärfaktoren der Fremdbeurteilungen (gemittelt über zwei Beurteiler) verwendet. LMT1 (L1) = Leistungsstreben; LMT2 (L2) = Ausdauer und Fleiß; LMT3 (F+) = leistungsfördernde Prüfungsangst; LMT4 (F-) = leistungshemmende Prüfungsangst; LMT = Gesamtwert der Leistungsmotivation (LMT1 + LMT2 + LMT3 - LMT4).

R multiple Korrelation R² korrigierter Determinationskoeffizient

* p < .05 ** p < .01

Schulleistungen sind von vielen verschiedenen Bedingungen abhängig. Mehr oder weniger gegenstandsspezifische Lernfähigkeiten und Interessen der Schüler treffen auf je nach Gegenstand, Lehrer, Schulklasse und Schule unterschiedliche Anforderungen und Anregungen. Dazu kommt die Förderung oder Behinderung des Lernens durch Eltern, Geschwister und Freunde. Lernen für die Schule konkurriert in der unterrichtsfreien Zeit mit vielen anderen Betätigungsmöglichkeiten.

Schließlich ist zu bedenken, daß die Schulnoten nicht einfach die tatsächlichen Schulleistungen wiedergeben, sondern auch anzeigen, wie weit das Verhalten der Schüler insgesamt den Erwartungen der Lehrer entspricht. Diese Erwartungen beziehen sich aber nicht nur auf Leistungen, sondern auch auf das soziale Verhalten der Schüler. Außerdem stellen Noten immer eine Art von Kompromiß zwischen objektiven Leistungsbewertungen und pädagogischen Erwägungen dar.

Von den vielen Bedingungen, als deren Zusammenwirken sich die Schulnoten ergeben, werden in diesem Bericht nur die Merkmale des «Temperaments» und der Leistungsmotivation berücksichtigt und dies nur so weit, als sie durch die Selbstbeurteilungen des 16PA und LMT sowie durch Mitschülerbeurteilungen anhand der 16PA-Skalen erfaßt werden können.

Korrelationen zwischen Temperamenteigenschaften und Schulleistungen können u. a. bedeuten, daß bestimmte Eigenschaften, wie Normgebundenheit (Gewissenhaftigkeit) oder Belastbarkeit (emotionale Stabilität) zu den Bedingungen guter Schulleistungen gehören, oder daß sich gute oder schlechte Schulleistungen auf die Selbstbeurteilungen und Beurteilungen durch Mitschüler auswirken, d. h. daß beide Arten von Beurteilungen mehr oder weniger treffende Ursachenattributionen guter oder schlechter Schulleistungen darstellen. Wir vermuten, daß das Persönlichkeitsbild, das die Schüler und Schülerinnen von sich selbst und von anderen, die sie aus längerer Klassengemeinschaft recht gut kennen, entwerfen, mehr

Tabelle 8: Gewichte (standardisierte partielle Regressionskoeffizienten) der Sekundärfaktoren des 16PA (Selbst- und Fremdbeurteilung) und der LMT-Subtests in der Vorhersage des Notendurchschnitts (Deutsch, Englisch und Mathematik), differenziert nach Geschlecht (m/w)

Geschlecht	m	w	m	w	m	w	m	w			
QIS	-.01	.03	QIF	-.01	.05	QIS	.00	.04	LMT1	-.22*	.10
QIIS	.00	-.10	QIIF	.10	-.22*	QIIS	.01	-.09	LMT2	-.22	-.41**
QIIIS	-.26**	-.08	QIIIF	-.34**	-.23*	QIIIS	-.27**	-.07	LMT3	-.09	-.02
QIVS	-.15	-.00	QIVF	-.23*	-.09	QIVS	-.16	-.01	LMT4	-.03	.15
QVS	.49**	.20*	QVF	.47**	.28**	QVS	.50**	.23*	TA	.07	-.22*
								ZVT	-.01	-.16	
R	.46	.24		.55	.43		.47	.38		.43	.39
R ²	.18	.02		.27	.15		.17	.09		.16	.13

Anmerkung. QIS bis QVS Sekundärfaktoren der Selbstbeurteilung als Prädiktoren; QIF bis QVF Sekundärfaktoren der Fremdbeurteilungen als Prädiktoren; TA Textanalyse; ZVT Zahlenverbindungstest; LMT1 (L1) = Leistungsstreben; LMT2 (L2) = Ausdauer und Fleiß; LMT3 (F+) = leistungsfördernde Prüfungsangst; LMT4 (F-) = leistungshemmende Prüfungsangst; R multiple Korrelation; R² korrigierter Determinationskoeffizient.

* p < .05 **p < .01

über die Bedingungen als über die Wirkungen der Schulleistungen bzw. der Schulnoten aussagt. Diese Annahme dürfte in besonderem Maße für die Selbstbeurteilungen gelten, da in diese ja mehr und vielfältigere außerschulische Erfahrungen einbezogen sind. Aber auch die Beurteilungen durch Mitschüler(innen) können, selbst wenn Ursachenattributionen von Schulleistungen dabei stärker hereinspielen sollten, zur Bedingungsanalyse von Schulleistungen herangezogen werden; denn Ursachenattributionen mögen zwar in bestimmten Fällen und unter bestimmten Bedingungen nur Tautologien darstellen (Gute-Noten-Haben «heißt» Begabt- oder Strebsam- oder Angepaßtsein) oder einfach falsch sein; oft werden aber Attributionen einen Kern von (nicht nur begriffslogischer, sondern empirisch überprüfbarer) Wahrheit enthalten.

Als förderlich für gute Schulleistungen hatten wir Normgebundenheit (QI) und Belastbarkeit (QII) angenommen. Wie Tabelle 8 zeigt, sind jedoch nur Kontaktbereitschaft (QV) und Unabhängigkeit (QIII) Prädiktoren der Schulleistung: je geringer die Kontaktbereitschaft und je höher die Unabhängigkeit, desto besser die Schulleistungen. Dies gilt auch für die Beurteilungen durch die Mitschüler(innen). Kombiniert man die PA- mit den LMT-Skalen, tragen übereinstimmend bei Selbstbeurteilungen und Mitschülerbeurteilungen nur LMT2 (Ausdauer und Fleiß) und QV (Kontaktbereitschaft) zur Erklärung der Schulnoten bei. Unterschiede zwischen den Fächern Deutsch, Englisch und Mathematik in den Korrelations- und Regressionskoeffizienten, die hier nicht näher berichtet werden, können als geringfügig vernachlässigt werden.

Ergänzend sei hier noch darauf hingewiesen, daß sich bei den HTL-Schülern, anders als bei den übrigen Schülern, sowohl leistungsfördernde (F+ bzw. LMT3) als auch leistungshemmende (F- bzw. LMT4) Prüfungsangst förderlich auf Schulleistungen auswirken, während die anderen beiden Skalen keine nennenswerten Korrelationen mit den Schulnoten aufweisen.

Überraschenderweise tragen die beiden kognitiven Leistungstests Textanalyse (TA) und Zahlenverbindungstest (ZVT) nur bei den Schülerinnen zur Erklärung des Notendurchschnitts bei. Die Leistungen in der Textanalyse korrelieren

im übrigen nicht mit den Sekundärdimensionen des 16PA. Eine signifikant positive Korrelation ($r(243) = .23$; $p < 0.05$; zweiseitig) findet sich zwischen QV (Kontaktbereitschaft) und der Leistung im Zahlenverbindungstest.

3. Diskussion

Die Verteilung der Korrelationen zwischen den einander entsprechenden Skalen von Form A und B (Tabelle 3 und 4) stimmt im großen und ganzen mit der überein, die Brandstätter (1988 a) berichtet. Da bei der Wahl der Adjektivpaare Synonyme bewußt vermieden wurden – es wurde vielmehr darauf geachtet, verschiedene Facetten der in den 16PF-Items ausgedrückten Erlebnis- und Verhaltensweisen mit passenden Adjektiven zu charakterisieren –, reicht eine mittlere Korrelation von $r(A, B) = .35$ aus, um sagen zu können, daß die beiden Eigenschaftspaare nicht den gleichen, wohl aber verwandte Verhaltensaspekte repräsentieren. Wie bei Brandstätter (1988 a) korreliert auch hier in der Regel die Skala der einen Form am höchsten mit der entsprechenden Skala der anderen Form, so daß die nötige konvergente und diskriminative Validität gegeben ist¹.

Die Retest-Reliabilitäten der über Form A und B gemittelten 16PA-Skalen liegen nach Brandstätter (1988 a) bei einem Zeitabstand von einem Monat zwischen .53 und .82 (Median .71). Dies weist darauf hin, daß die mäßigen Korrelationen zwischen den beiden Testformen untere Schranken der Wiederholungsreliabilität darstellen und nicht auf einen Mangel an Stabilität der Messungen zurückzuführen sind.

Wenn sich Schülerinnen in manchen Skalen des 16PA im Durchschnitt anders charakterisieren als Schüler, sollte man in diesen Aussagen nicht einfach Widerspiegelungen von allgemein verbreiteten Vorurteilen sehen und ihnen damit jede Validität absprechen. Es ist vielmehr wahrscheinlich, daß diese Schülerinnen nicht nur meinen und sagen, sie seien kontaktfreudiger, feinfühlicher, phantasievoller, natürlicher und ängstlicher als Schüler, sondern daß sie sich im Alltag

¹ Die vollständige Korrelationstabelle ist aus Platzgründen nicht abgedruckt.

auch tatsächlich so, d. h. eben anders als gleichaltrige Schüler derselben sozialen Schicht, fühlen und verhalten, sei es weil sie ein mit der traditionellen Geschlechtsrolle konformes Selbstbild erworben haben und diesem «nachleben», sei es weil biogenetisch bedingte Geschlechtsunterschiede Erleben und Verhalten bestimmen und im Selbstbild reflektiert werden.

Wenn die Geschlechtsunterschiede in der Beurteilung durch andere in ähnlicher Weise in Erscheinung treten, ist zwar nicht von vornherein auszuschließen, daß Vorurteile (Geschlechtsstereotype) wirksam sind; die Wahrnehmung könnte sich der Erwartung angeglichen haben (Assimilation). Ein Gegenargument wäre, daß die Beurteilung durch andere eher zu einem Kontrastphänomen führen müßte, wenn das tatsächliche Verhalten nicht der Erwartung entspricht: Wer von Schülerinnen erwartet, daß sie kontaktfreudiger und feinfühlicher sind als Schüler, müßte sie im Vergleich zu Schülern und im Kontrast zur Erwartung als weniger kontaktfreudig und feinfühlig einschätzen, wenn sie tatsächlich genau so kontaktfreudig und feinfühlig wären wie Schüler.

Die Übereinstimmung zwischen zwei Personen in den Urteilen auf den Primärfaktoren über dieselbe dritte Person liegt im Mittel (Median) bei .31 (von .22 bis .58). Sie ist deutlich höher als die Übereinstimmung zwischen Selbst- und Fremdbeurteilung (ein Beurteiler), deren Median bei .24 (.08 bis .49) liegt. In den Sekundärfaktoren ist die Übereinstimmung zwischen Selbstbeurteilungen und Beurteilungen durch andere, von der Dimension QII (Belastbarkeit) abgesehen, genügend hoch, daß sie als Beleg für die Validität der 16PA-Selbstbeurteilungen gelten kann. Wenn die Validitätskoeffizienten etwas niedriger ausfielen als vergleichbare Werte bei Amelang und Borkenau (1982, S. 135), mag dies teils auf die größere Heterogenität der von diesen Autoren verwendeten Stichprobe, teils auf die Mittelung der Fremdbeurteilungen über drei Verwandte oder Bekannte (anstatt über zwei Mitschüler) zurückzuführen sein. Des weiteren könnte die Verankerung der 32 Eigenschaftspole der Ratingskalen in jeweils drei bis vier Adjektiven (statt in einem wie beim 16PA) zu einer Erhöhung der Reliabilität und damit der Validität beigetragen haben.

Unabhängigkeit (QIII) und Kontaktbereit-

schaft (QV) der Selbstbeurteilung korrelieren im wesentlichen nur mit den entsprechenden Faktoren der Beurteilung durch andere. Weniger gut steht es mit der *diskriminativen* Validität der Selbstbeurteilung im Hinblick auf die Kriterienfaktoren QI und QIV. Hier sind die multiplen Korrelationen deutlich höher als die Korrelationen der korrespondierenden Skalen. Dies deutet darauf hin, daß die Beurteilungsaspekte, die in diese Skalen gemäß den von Brandstätter (1988 a) ermittelten Gewichten (zur Schätzung der 16PF-Sekundärfaktoren aufgrund der 16PA-Einstufungen) vorwiegend eingehen, in der Fremdbeurteilung eine etwas andere Bedeutung haben als in der Selbstbeurteilung.

Die interne Konsistenz der Beurteilungen einer Person, als Korrelation der Eigenschaftsprofile der beiden Formen des 16PA pro Person bestimmt, scheint eine recht deutliche Differenzierung der Validität der Urteile zu ermöglichen. Die Varianzen der Sekundärfaktoren sind aber sowohl bei den Selbstbeurteilungen als auch bei den Beurteilungen durch die Mitschüler(innen) bei geringer Konsistenz niedriger als bei hoher Konsistenz, während sich die Varianzen der Schätzfehler kaum voneinander unterscheiden; diese sind in keinem Faktor bei hoher Konsistenz niedriger als bei niedriger Konsistenz. Die eigentliche Ursache für die (tatsächlich nur scheinbar) höhere Validität konsistenter Urteile liegt daher in der höheren Streuung, die konsistente Urteile zwangsläufig aufweisen; denn nur wenn häufig beide der einander entsprechenden Skalen hohe oder niedrige Werte aufweisen, ergibt sich hohe Konsistenz und damit auch eine große Streuung der Werte in den Primär- und schließlich auch in den Sekundärfaktoren.

Dies spricht nicht gegen das Konzept und die Art der Operationalisierung der Urteilkonsistenz, sondern weist nur auf die Gefahr naheliegender Fehlinterpretationen von Unterschieden in Validitätskoeffizienten hin, die sich im Vergleich von hoch und niedrig konsistenten Urteilen ergeben. Dieses Problem besteht für jede Art von Konsistenzbestimmung, die sich in der einen oder anderen Weise auf die Varianz von Urteilen über semantisch ähnliche Verhaltensaspekte bezieht.

Es stellt sich nun die Frage, wie weit hohe personenspezifische Urteilkonsistenz der Verhaltens-

ebene zuzuschreiben ist (vgl. dazu Brandstätter, 1983, S. 120 f.). Auch wenn dies anhand der vorliegenden Daten nicht eindeutig zu entscheiden ist, spricht doch einiges dafür, daß Urteilkonsistenz etwas mit der Eigenart der beurteilten Person zu tun hat, da die drei Konsistenzmaße (für die Selbstbeurteilung und die beiden Fremdbeurteilungen) im Mittel mit $r(243) = .22$ korreliert sind und auf demselben Faktor laden. Wie weit Konsistenz auch ein Merkmal der Beurteiler ist, könnte im Rahmen einer Analyse nach dem Modell von Kenny (1988) geprüft werden.

Gegen die plausible Erklärung, inkonsistente Beurteilungen seien vor allem durch mangelnde Sorgfalt der Beurteiler bedingt, spricht die Tatsache, daß die Restvarianz (die Varianz der Abweichungen vom vorhergesagten Wert) bei inkonsistenten Urteilen nicht größer ist als bei konsistenten. Hohe Konsistenz ist andererseits keine Gewähr für Sorgfalt. Gerade wenn die Bereitschaft zu genauer und deshalb oft anstrengender Differenzierung nicht sehr hoch ist, man sich aber auch nicht nur zum Schein (durch wahlloses Markieren von Ziffern) der Aufgabe entledigen will, liegt es nahe, sich bei den Urteilen mehr an der Bedeutungsähnlichkeit der Begriffe zu orientieren als an den Nuancen ihrer Unterschiede und damit an den Unterschieden der Situationen und Verhaltensweisen, für die sie gedacht sind. Eine Übereinstimmung zwischen Selbst- und Fremdbeurteilung kommt in einem solchen Fall allerdings nur dann zustande, wenn in beiden Urteilspektiven die «semantische Ergänzung» in der betreffenden Dimension von denselben auffallenden und daher ohne Anstrengung feststellbaren Verhaltensmerkmalen ausgeht, z. B. davon, daß jemand lebhaft spricht, und den gleichen semantischen Regeln («wer lebhaft spricht, soll kontaktfreudig heißen») oder den gleichen Hypothesen über Kovariation von Verhaltensweisen («wer lebhaft spricht, ist mit großer Wahrscheinlichkeit auch kontaktfreudig») folgt. Soll auf diese Weise eine Person (von mehreren anderen oder von ihr selbst und anderen) in sagen wir fünf Grunddimensionen übereinstimmend beurteilt werden, bedarf es mindestens eines Merkmals (d. h. eines leicht merkbaren Kennzeichens des Verhaltens) pro Dimension, von dem aus dann die weiteren Schlüsse im semantischen Netzwerk und/oder im Geflecht der Annahmen über Ko-

variation von Verhaltensweisen ohne weitere Überprüfung an (erinnerten) Beobachtungen gezogen werden.

Sollte sich herausstellen, daß konsistente Selbstbeurteilungen auch eine höhere Validität in Bezug auf objektivere Kriterien und nicht nur in Bezug auf Beurteilungen durch andere haben, spräche dies für die Annahme, daß konsistente Urteile die Folge von treffend erfaßtem konsistentem Verhalten und nicht von semantischen Extrapolationen sind. Eine hinsichtlich hoher und niedriger Konsistenz differenzierte Analyse nach dem Muster von Tabelle 7 und 8 weist tatsächlich in diese Richtung.

Daß eine vereinfachte Gewichtung der 16PA-Primärfaktoren zur Schätzung der 16PF-Sekundärfaktoren der regressionsanalytisch bestimmten Gewichtung hinsichtlich der Validität der Selbstbeurteilungen sogar etwas überlegen ist, könnte bedeuten, daß die von Brandstätter (1988 a) ermittelten Gewichte noch der Überprüfung und Revision anhand von Daten weiterer Stichproben bedürfen.

Die Vermutung, die Werte des Leistungsmotivationstests ließen sich weitgehend aus den Daten des 16PA rekonstruieren, hat sich in der Gesamtgruppe nur für die Summe der vier LMT-Skalen und für die Komponente F- (leistungshemmende Prüfungsangst) bestätigt. Höhere Vorhersagbarkeit findet sich bei den Schülern (.54, .49, .39 und .61), die bei L1 und L2 auf die Faktoren QI, QII, QIII und QV, bei F- auf -QII und -QIII zurückgeht. Daß Motivstärken relativ unabhängig von Temperamentsfaktoren variieren, entspricht der Auffassung von R. B. Cattell, der Fähigkeits-, Temperaments- und Motivationsdimensionen zur umfassenden Beschreibung und Klassifikation von Persönlichkeitsmerkmalen für nötig hält (vgl. Cattell & Child, 1975; Cattell & Johnson, 1986).

Die Schulnoten lassen sich durch eine gewichtete Summe von LMT- und 16PA-Dimensionen besser vorhersagen als durch die eine oder andere Variablengruppe allein. Unerwartet kam, daß nicht QI (Normgebundenheit) und QII (Belastbarkeit), sondern QV (Kontaktbereitschaft) mit den Schulleistungen korreliert: Je geringer die Kontaktbereitschaft (d. h. je deutlicher die Introversion ausgeprägt ist), desto besser sind die Schulleistungen. Auch in der Beurteilung durch andere sind es die beiden Faktoren QV (Kon-

taktbereitschaft) und QIII (Unabhängigkeit), die bemerkenswert hohe Korrelationen mit dem Notendurchschnitt aufweisen. Von den LMT-Skalen trägt nur L2 (Ausdauer und Fleiß) zur Erklärung der Schulnoten bei.

Introversion (geringe Kontaktbereitschaft) könnte den Schulleistungen zugute kommen, weil ablenkende Aktivitäten, insbesondere gesellige Unternehmungen mit Freunden, für introvertierte Schüler(innen) weniger verlockend sind. Mit dieser Deutung stimmt überein, daß sich Extravertierte in der Freizeit viel wohler fühlen als in der Arbeitszeit, während sich Introvertierte bei der Arbeit kaum weniger wohlfühlen als in der Freizeit (Brandstätter, 1988 b). In einer noch nicht veröffentlichten Arbeit zeigte sich außerdem, daß die Differenz in der Einstufung der Begriffe Arbeit und Freizeit auf der Valenzdimension eines semantischen Differentials bei Extravertierten, die der Begriff Freizeit besonders positiv anmutet, viel größer ist als bei Introvertierten. Daß der Faktor QIII (Unabhängigkeit; könnte auch als Dominanz bezeichnet werden) eher mit guten Noten verbunden ist, entspricht nicht der weit verbreiteten Meinung, mit guten Noten würden vor allem Wohlverhalten und Anpassung belohnt.

Abschließend läßt sich sagen, daß die Eigenschaftsskalen des 16PA, die mit geringem Zeitaufwand eine Schätzung der 16PF-Sekundärdimensionen für experimentelle Zwecke ermöglichen sollen, eine befriedigende Validität aufweisen. Von den Subskalen des Leistungsmotivations-tests ist vor allem die Skala der leistungshemmenden Angst rekonstruierbar. Gut vorhersagbar ist auch der Gesamtwert der Leistungsmotivation. Zur Erklärung der Schulnoten trägt aber der Subtest L2 (Fleiß und Ausdauer) über den 16PA hinaus bei.

Auch wenn zu vermuten ist, daß neue Persönlichkeits- oder Selbstkonzeptsskalen² den durch die fünf Grunddimensionen der Personwahr-

nehmung aufgespannten Raum kaum überschreiten können, und daß der 16PF und damit auch der 16PA diese Grunddimensionen einschließen, somit eine ausreichende Rekonstruktion beliebiger anderer Persönlichkeitsdimensionen ermöglicht, ist dies für einige der in letzter Zeit entwickelten Skalen (z. B. für den Psychotizismus-Faktor des EPQ (Eysenck & Eysenck, 1975) oder für den Fragebogen zur Handlungs- und Lageorientierung; Kuhl, 1983) erst noch zu zeigen. Daß gängige Persönlichkeitstests wie der EPI (Eysenck Personality Inventory; Eysenck & Eysenck, 1964) oder das Freiburger Persönlichkeitsinventar (FPI; Fahrenberg & Selg, 1970) nicht anders als der 16PF auf dieselben Grunddimensionen reduzierbar und damit auch weitgehend ineinander überführbar sind, dürfte aus vorliegenden Arbeiten hinreichend klar hervorgehen (u. a. Amelang & Borkenau, 1982; Noller, Law & Comrey, 1987; Brandstätter, 1991). Es ist zu hoffen, daß diese Einsicht und die Bereitstellung einer ökonomischen Form der Erfassung dieser Grunddimensionen dazu beiträgt, daß künftig Persönlichkeitsunterschiede insbesondere in der sozialpsychologischen Forschung häufiger berücksichtigt werden.

Literatur

- Amelang, M. & Bartussek, D. (1990). *Differentielle Psychologie und Persönlichkeitsforschung*. 3. Aufl. Stuttgart: Kohlhammer.
- Amelang, M. & Borkenau, P. (1982). Über die faktorielle Struktur und externe Validität einiger Fragebogen-Skalen zur Erfassung von Dimensionen der Extraversion und emotionalen Labilität. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 119–146.
- Bartussek, D. (1988). Beurteilung der deutschen Form des 16PF-Tests. *Diagnostica*, 34, 367–379.
- Borkenau, P. (1986). Untersuchungen zur internen Konsistenz und externen Validität der deutschsprachigen Version des 16PF-Test von Schneewind, Schröder & Cattell. *Diagnostica*, 32, 100–110.
- Borkenau, P. & Ostendorf, F. (1991). Ein Fragebogen zur Erfassung fünf robuster Persönlichkeitsfaktoren. *Diagnostica*, 37, 29–41.
- Brandstätter, H. (1977). Wohlbefinden und Unbehagen. Entwurf eines Verfahrens zur Messung situationsabhängiger Stimmungen. In W. H. Tack (Hrsg.), *Bericht über den 30. Kongreß der DGfPs in Regensburg 1976*, Band 2 (S. 60–62). Göttingen: Hogrefe.
- Brandstätter, H. (1983). *Sozialpsychologie*. Stuttgart: Kohlhammer.

- Brandstätter, H. (1988 a). Sechzehn Persönlichkeits-Adjektivskalen (16PA) als Forschungsinstrument anstelle des 16PF. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 35, 370–391.
- Brandstätter, H. (1988 b). Persönlichkeitsunterschiede in Gefühlen und Stimmungen während der Arbeit und Freizeit. In W. Schönflug (Hrsg.), *Bericht über den 36. Kongreß der Deutschen Gesellschaft für Psychologie in Berlin 1988*, Band 1 (S. 23–24). Göttingen: Hogrefe.
- Brandstätter, H. (1990). Motives in everyday life situations: An individual difference approach. In F. Halisch & J. van den Bercken (Eds.), *International perspectives on achievement and task motivation* (pp. 327–349). Lisse: Swets & Zeitlinger.
- Brandstätter, H. (1991). Social exchange orientation versus social reinforcement orientation in controversial discussions. In: S. Worchel, W. Wood & J. A. Simpson (Eds.), *Productivity and process in groups*. Sage: Beverly Hills.
- Cattell, R. B. & Child, D. (1975). *Motivation and dynamic structure*. New York: Wiley.
- Cattell, R. B. & Johnson, R. C. (Eds.) (1986). *Functional psychological testing. Principles and instruments*. New York: Brunner/Mazel.
- Costa Jr., P. T. & McCrae, R. R. (1985). *The NEO Personality Inventory*. Odessa, FL.: Psychological Assessment Resources.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440.
- Eder, F. & Bergmann, C. (1988). Der Person-Umwelt-Struktur-Test. *Psychologie in Erziehung und Unterricht*, 35, 299–309.
- Eysenck, H. J. & Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University Press.
- Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder & Stoughton.
- Fahrenberg, J. & Selg, H. (1970). *Das Freiburger Persönlichkeitsinventar FPI. Handanweisung*. Göttingen: Hogrefe.

Anschrift: Prof. Dr. Hermann Brandstätter, Mag. Gernot Filipp und Mag. Peter Drescher, Abteilung für Sozial- und Wirtschaftspsychologie, Johannes-Kepler-Universität Linz, A-4040 Linz/Donau

- Hermans, H., Petermann, R. & Zielinski, W. (1978). *LMT-Leistungsmotivations-test*. Amsterdam: Swets & Zeitlinger.
- Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychologische Rundschau*, 19, 21–25.
- Kenny, D. A. (1988). Interpersonal perception: A social relations analysis. *Journal of Social and Personal Relationships*, 5, 247–261.
- Krampen, G. (1986). Zur Validität der deutschen Form des 16PF. Faktorielle Validität und Beziehungen zum FPI. *Diagnostica*, 32, 91–99.
- Kuhl, J. (1983). *Motivation, Konflikt und Handlungskontrolle*. Heidelberg: Springer.
- Lanning, K. (1988). Individual differences in scalability: An alternative conception of consistency for personality theory and measurement. *Journal of Personality and Social Psychology*, 55, 142–148.
- McCrae, R. R. & Costa Jr., P. T. (1985). Updating Norman's «adequate taxonomy»: Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49, 710–721.
- Noller, P., Law, H. & Comrey, A. L. (1987). Cattell, Comrey, and Eysenck personality factors compared: More evidence for the five robust factors? *Journal of Personality and Social Psychology*, 53, 775–782.
- Oswald, W. D. & Roth, E. (1987). Der Zahlen-Verbindungs-Test. *ZVT*. C. J. Göttingen: Hogrefe.
- Sackett, P. R., Zedeck, S. & Fogli, L. (1988). Relations between measures of typical and maximum performance. *Journal of Applied Psychology*, 73, 482–486.
- Schneewind, K. A., Schröder, G. & Cattell, R. B. (1983). *Der 16-Persönlichkeits-Faktoren-Test. (16PF)*. Bern: Huber.
- Trapnell, P. D. & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the big five dimensions of personality. *Journal of Personality and Social Psychology*, 59, 781–790.