

## SELF-BF – EIN INDIREKTES VERFAHREN ZUR ERFASSUNG DER BIG FIVE<sup>1</sup>

Seit einiger Zeit wird der Persönlichkeit eine ebenso große Bedeutung für Eignung und Leistung im Beruf beigemessen wie der Intelligenz, zumindest im Bereich anspruchsvollerer Expertentätigkeiten (Borkenau et al., 2005) sowie bei Managementfunktionen (Sarges, 2000, 2006), und eine noch größere Rolle spielt die Persönlichkeit bei Fragen der Passung von Personen zu Organisationen, Vorgesetzten, Kollegen, Mitarbeitern u.a. Eines der in diesen Zusammenhängen am häufigsten herangezogenen Strukturmodelle der Persönlichkeit ist das sog. Big Five-Konzept mit den Persönlichkeitsfaktoren Neurotizismus (bzw. umgepolt: Emotionale Stabilität), Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit (John & Srivastava, 1999). Nach einer Zusammenschau diverser Metaanalysen im Hinblick auf die Prädiktion beruflicher Leistung kommen Barrick, Mount und Judge (2001) zu dem Schluss, dass von diesen fünf großen Persönlichkeitsfaktoren allen voran der Gewissenhaftigkeit und mit einigen Abstrichen auch noch der Emotionalen Stabilität eine generelle Validität zugesprochen werden kann. Darüber hinaus aber dürfte es etliche anforderungsspezifische prognose-dienliche Validitäten geben (wozu eine Übersicht allerdings noch aussteht) und last but not least die Chance vielfältiger Hinweise zur Beantwortung individueller Passungsfragen.

Wegen solch augenfälliger Bedeutsamkeit des Big Five-Konzepts für die Praxis der Eignungsdiagnostik haben wir mit der Entwicklung eines Verfahrens begonnen, das im Vergleich zu klassischen Persönlichkeitsinventaren zur Messung der Big Five sowohl eine höhere *Akzeptanz* erzielen als auch die *Verfälschbarkeit* in Richtung positiver Selbstdarstellung (Kubinger, 2002; Viswesvaran & Ones, 1999) reduzieren soll.

Auch wenn – wie entsprechende Studien zeigen (Marcus, 2003) – die Validität von Persönlichkeitsinventaren durch positive Verfälschungstendenzen nicht reduziert wird, so ist mit einer ausschließlichen Argumentation über die korrelative Validität das Problem der Verfälschbarkeit vielleicht für manchen Forscher, nicht aber für die Praxis erledigt (vgl. hierzu auch Kanning, in diesem Band): Kandidaten in Auswahl-situationen, die sozial erwünscht antworten, landen gewöhnlich auf den oberen Rangplätzen (Hough, 1998; Rosse, Stecher, Miller & Levin, 1998) und werden somit auch bevorzugt eingestellt. Das mag manchmal gut sein, doch längst nicht immer.

Eine Akzeptanzerhöhung von konstruktorientierten Verfahren der Persönlichkeitserfassung ist wünschenswert, weil bei praktisch allen gängigen Inventaren die Operationalisierungen

---

<sup>1</sup> erschienen in: Sarges, W. & Scheffer, D. (Hrsg.). (2008), Innovative Ansätze für die Eignungsdiagnostik (S. 215–227). Göttingen: Hogrefe.

der Konstrukte über Statements oder Fragen laufen, deren Inhalte zumindest in Auswahl-situationen nicht selten als anforderungsfern, langweilig, dummlich- bzw. ärgerlich-durchschaubar, merkwürdig, trivial-penetrant oder gar als Zumutung empfunden werden – jedenfalls im Vergleich zu Verfahrensklassen mit leichter herstellbarem Anforderungsbezug wie Simulationen oder Interviews.

Um dadurch hervorgerufene Einschränkungen der sozialen Validität von Persönlichkeitsverfahren zu überwinden, gingen wir einen anderen Weg: Mit unserem SELF-BF (SELbsteinschätzungs-Fragebogen-Big Five) wird – im Gegensatz zu üblichen Persönlichkeitsinventaren – nicht das Verhalten und/oder Erleben in bestimmten Situationen erfragt, sondern das Selbstkonzept, d.h. das kognitive Abbild einer Person von sich selbst, und zwar über Eigenschaftsworte. Die *affektive* Komponente der Selbstwahrnehmung dagegen, die Bewertung dieses kognitiven Bildes (= Selbstwertschätzung bzw. Selbstwertgefühl), haben wir nicht erfasst, da sich einerseits in bisherigen Forschungen herausgestellt hat, dass globale Maße des Selbstwerts nicht mit spezifischeren Verhaltenskriterien oder Persönlichkeitsmerkmalen korrelieren (z.B. Baumeister, Campbell, Krueger & Vohs, 2003) und andererseits diese durchschaubaren und für relevant gehaltenen Maße in Auswahl-situationen ganz besonders bedrohunggefährdet sein dürften. Der für unsere Zwecke hier wichtigste Unterschied zwischen Selbstwert (feelings of self-worth) und Selbstkonzept (beliefs about the self) liegt schon allein in der Spezifität: Selbstwert-Einschätzungen beziehen sich als Globalmaße (auch wenn sie bereichsspezifisch differenziert werden, z.B. nach mental, emotional-motivational, sozial-interaktiv oder physisch) lediglich auf einen oder wenige Gesamteindrücke des persönlichen Wertes, Selbstkonzepte dagegen auf viele einzelne Qualitäten, die Menschen sich zuschreiben (z.B. gesellig, dominant, kreativ, künstlerisch, intelligent, genau, interessant etc.). Selbstkonzepte erschienen uns daher im Hinblick auf die gewünschten Differenzierungen viel versprechend (s.a. Swann, Chang-Schneider & Larsen McClarty, 2007).

In der Tat hatten wir mit Vorläuferversionen des SELF-BF ermutigende Erfahrungen gesammelt: einerseits mit Bezug auf die Akzeptanz der verwendeten Erhebungsmethode (Charakterisierung der eigenen Person über Eigenschaftsworte), andererseits mit Bezug auf die Diskrimination von Probanden: (1. von arbeitslosen Personen, die starke vs. schwache Motivation zu Erwerbsarbeit erkennen ließen, 2. von Beschäftigten im Einzelhandel, die durch Stehlen auffällig vs. nicht-auffällig geworden waren, 3. von Personen einer dispersen Stichprobe, die als fleißig vs. nicht fleißig wahrgenommen wurden). Wenn man schon diese für bestimmte praktische Zwecke zwar nützlichen, aber doch auch eher engen Konstrukte „Motivation zu Erwerbsarbeit“, „Neigung zum Stehlen“ und „Fleiß“ mit dem SELF erfassen konnte, so meinten wir, müsste es doch auch möglich sein, breiter gefasste Konstrukte wie die der Big Five mit dieser von Probanden besser akzeptierten und weniger durchschaubaren Methode messbar zu machen.

Ausgehend von einer Vorläuferversion unseres Verfahrens war es nun unser Ziel, einerseits die Akzeptanz weiter zu erhöhen und andererseits die Durchschaubarkeit noch zu erschweren, und zwar

- durch Zulassen auch von offenen (vs. geschlossenen) Antworten sowie
- durch deutliche Erhöhung von Anzahl und Inhaltvielfalt der Variablen.

Die Erfassung des Selbstkonzepts erfolgte dann auf drei verschiedene Arten (Subtests) mit zugleich abnehmender Offenheit der Antworten im Reaktionsschema:

- durch offene Antworten (Eigenschaftsworte selbst aufschreiben),
- durch Freiwahlen aus einer Liste von vorgegebenen Eigenschaftsworten und
- durch Zwangswahlen bei vorgegebenen Eigenschaftswort-Paaren.

Diese Selbsteinschätzungen verstanden wir natürlich nicht als verhaltensmäßig zutreffend (= „behavioral-veridikal“) – diesbezügliche Mängel von Selbstberichten sind ja hinlänglich bekannt –, vielmehr interessierten uns mögliche Zusammenhänge zwischen solcherart erfassten kognitiven Repräsentationen des Selbst einer Person und den Big Five in etablierten Persönlichkeitsinventaren (hier konkret: dem NEO-FFI).

Selbstkonzept-Variablen (= Elemente) und ihre metakognitiven Aspekte dürften für diesen Zweck einiges an Potenzial besitzen. Ein zentrales metakognitives Charakteristikum eines Selbstkonzept-Elements beispielsweise ist seine Stärke. Indikatoren für Stärke sind Wichtigkeit, Sicherheit, Klarheit, Extremität, Zugänglichkeit, zeitliche Stabilität, Zielbezogenheit u.a. (Swann et al., 2007, S. 86). Selbstkonzept-Elemente (self-beliefs) beispielsweise, die hochwichtig sind, bleiben mit höherer Wahrscheinlichkeit auch stabiler über die Zeit.

Falls es am Ende gelingt, substantielle Zusammenhänge zwischen den beiden Instrumenten SELF-BF und NEO-FFI zu erreichen, könnte man zumindest in akzeptanzgefährdeten Auswahl- und Beratungssituationen besser den SELF-BF zur Messung der Big Five heranziehen. Und wenn sich zusätzlich noch empirisch nachweisen ließe, dass die Vielfalt der SELF-Variablen (über 500, s.u.) weniger durchschaubar und somit schlechter verfälschbar ist als die 60 durchschaubaren Items des NEO-FFI, wären wir unseren beiden Zielen näher gekommen.

## ZUM AUFBAU DES SELF-BF

Der SELF-BF besteht – wie erwähnt – aus drei Subtests:

- *Eigenschaften aufschreiben*: einem Selbstbeschreibungsbogen mit offenen Antworten, dem von uns so genannten Operanten Selbstkonzept-Test (OSKT) (vgl. Abbildung 1),
- *Eigenschaften ankreuzen*: einem Selbstbeschreibungs-Inventar mit Freiwahl (SI-FW) (vgl. Abbildung 2),
- *sich zwischen je zwei Eigenschaften entscheiden*: einem Selbstbeschreibungs-Inventar mit Zwangswahl (SI-ZW) (vgl. Abbildung 3).

Alle drei Subtests verlangen sowohl eine Positiv-Beschreibung (= wie man glaubt, am ehesten zu sein), als auch eine kontrastierende Negativ-Beschreibung (= wie man glaubt, am wenigsten zu sein).

Beim OSKT haben wir in einer Vielzahl von Anwendungen festgestellt, dass die im offenen Antwortformat erhaltenen Eigenschaften sich durchaus nicht immer deckten mit den in einem Freiwahlformat angekreuzten. Insofern erhofften wir uns, mit dem OSKT zusätzliche (inkrementelle) diagnostische Varianz zu erfassen, die also nicht deckungsgleich mit der aus dem Freiwahlformat des nachfolgenden SI-FW war (vgl. auch Delia & Grossberg, 1977); und

weiterhin hofften wir, auch mit dem auf das SI-FW folgenden Forced-Choice-Format (SI-ZW) noch relevante Varianz zu erfassen, die nicht schon in den beiden vorangegangenen Subtests enthalten war.

Einige Bemerkungen noch zu den Items des SI-FW: Sie entsprachen zu einem großen Teil den am häufigsten genannten Eigenschaften im OSKT aus einer Vielzahl von kleineren, mittleren und auch großen (= personenreichen) Studien an recht verschiedenen Personengruppen, darüber hinaus haben wir noch etliche Items zur Erhöhung der inhaltlichen Breite in Richtung der Big Five im lexikalischen Ansatz eingefügt (aus Adjektivlisten von Goldberg, 1992, Gough, 1997, und John, 1990). Analoges gilt für die Konstruktion des Forced-Choice-Verfahrens (SI-ZW): Zwar lehnte sich dieses an eine Idee von Ghiselli (1954) an – seine Item-Konstruktion nach der forced-choice-technique galt schon in der damaligen Forschung als „disguised measure and .. less sensitive to faking“ (Schippmann & Prien, 1985, S.1176; siehe hierzu auch Christiansen, Burns & Montgomery, 2005); wir haben aber auch hier viele Items selbst entwickelt, bestehende ersetzt und übernommene ergänzt – auch ausgehend von in den erwähnten Studien im OSKT vorgekommenen Positiv- und Negativ-Eigenschaften und immer darauf bedacht, die soziale Erwünschtheit der Antwort-Alternativen pro Item auszubalancieren.

**Abbildung 1: OSKT (Eigenschaften aufschreiben)**

|  |   |
|--|---|
| Bitte schreiben Sie 5 Eigenschaften auf,<br>die Sie <b>am ehesten</b> beschreiben: |   |
| 1.   |   |
| 2.   | Bitte schreiben Sie 5 Eigenschaften auf,<br>die Sie <b>am allerwenigsten</b> beschreiben: |
| 3.   | 1.  |
| 4.   | 2.  |
| 5.   | 3.  |
|  | 4.  |
|  | 5.  |

Abbildung 2: SI-FW (Eigenschaften ankreuzen)

Bitte kreuzen Sie auf der nächsten Seite diejenigen Worte an, die Sie **am ehesten** beschreiben. Sie können so viele Worte ankreuzen, wie Sie wollen!

Manche Worte haben nichts mit Ihrem Beruf zu tun, einige Worte sind Ihnen sogar vielleicht nicht bekannt. Das macht aber nichts, denn es gibt keine richtigen oder falschen Antworten.

Versuchen Sie bitte einfach, sich ganz spontan so zutreffend wie möglich zu beschreiben.

|                    |                          |                 |                          |                  |                          |
|--------------------|--------------------------|-----------------|--------------------------|------------------|--------------------------|
| hilfsbereit .....  | <input type="checkbox"/> | angesehen ..... | <input type="checkbox"/> | ruhig .....      | <input type="checkbox"/> |
| humorvoll .....    | <input type="checkbox"/> | gelassen .....  | <input type="checkbox"/> | gründlich .....  | <input type="checkbox"/> |
| strukturiert ..... | <input type="checkbox"/> | teamfähig ..... | <input type="checkbox"/> | höflich .....    | <input type="checkbox"/> |
| versichtlich ..... | <input type="checkbox"/> | konstant .....  | <input type="checkbox"/> | ausgewogen ..... | <input type="checkbox"/> |

**Achtung! Änderung!**  
Bitte kreuzen Sie nun diejenigen Worte an, die Sie **am wenigsten** beschreiben.

|                       |                          |                    |                          |                      |                          |
|-----------------------|--------------------------|--------------------|--------------------------|----------------------|--------------------------|
| falsch .....          | <input type="checkbox"/> | langsam .....      | <input type="checkbox"/> | gerissen .....       | <input type="checkbox"/> |
| draufgängerisch ..... | <input type="checkbox"/> | egoistisch .....   | <input type="checkbox"/> | verständnislos ..... | <input type="checkbox"/> |
| chaotisch .....       | <input type="checkbox"/> | kaltherzig .....   | <input type="checkbox"/> | geduldig .....       | <input type="checkbox"/> |
| schallend .....       | <input type="checkbox"/> | unaufmerksam ..... | <input type="checkbox"/> | berflächlich .....   | <input type="checkbox"/> |

Abbildung 3: SI-ZW (sich zwischen je 2 Eigenschaften entscheiden)

Kreuzen Sie bitte bei jedem der folgenden Eigenschaftspaare nun diejenige **Eigenschaft** von beiden an, die Sie **am meisten** beschreibt.

Das ist zwar manchmal schwierig, aber bitte entscheiden Sie sich trotzdem. Es gibt keine richtigen oder falschen Antworten. Versuchen Sie deshalb, sich so zutreffend wie möglich zu beschreiben.

Beispiel:  klug  
 einfallsreich

|  |   |   |
|--|---|---|
| 1. <input type="checkbox"/> fähig, kompetent | 14. <input type="checkbox"/> energiegeladen | 27. <input type="checkbox"/> begeisterungsfähig |
| <input type="checkbox"/> umsichtig, diskret  | <input type="checkbox"/> ausgeglichen       | <input type="checkbox"/> ausgeglichen           |
| 2. <input type="checkbox"/> verständnisvoll  | 15. <input type="checkbox"/> kommunikativ   | 28. <input type="checkbox"/> ehrgeizig          |
| <input type="checkbox"/> genau               | <input type="checkbox"/> gewissenhaft       | <input type="checkbox"/> ehrlich                |
| 3. <input type="checkbox"/> introvers        | 16. <input type="checkbox"/> energiegelad   | 29. <input type="checkbox"/> selbstverleugn     |
| 4. <input type="checkbox"/> energiegelad     | 17. <input type="checkbox"/> selbstverleugn | 30. <input type="checkbox"/> selbstverleugn     |

**Achtung! Änderung!**  
Kreuzen Sie bitte jetzt bei jedem der folgenden Eigenschaftspaare diejenige **Eigenschaft** von beiden an, die Sie **am wenigsten** beschreibt.

|  |  |   |
|--|--|---|
| 40. <input type="checkbox"/> intolerant  | 54. <input type="checkbox"/> unzuverlässig         | 68. <input type="checkbox"/> gleichgültig |
| <input type="checkbox"/> faul            | <input type="checkbox"/> unfreundlich              | <input type="checkbox"/> kompliziert      |
| 41. <input type="checkbox"/> unpünktlich | 55. <input type="checkbox"/> Verantwortung meidend | 69. <input type="checkbox"/> ziellos      |
| <input type="checkbox"/> ungeduldig      | <input type="checkbox"/> verkrampt                 | <input type="checkbox"/> gerissen         |
| 42. <input type="checkbox"/> egoistisch  | 56. <input type="checkbox"/> egoistisch            | 70. <input type="checkbox"/> stritsch     |

Die Zahl der Variablen, d.h. das Reservoir der nennbaren Worte, betrug beim OSKT nicht 5 (jeweils positiv und negativ) – wie der erste Anschein suggerieren mag, sondern theoretisch weitaus mehr, empirisch jedenfalls haben wir jeweils (positiv und negativ) weit über 100 erhalten; die Variablenzahl bei dem SI-FW betrug 100 für die positiven Adjektive und 100 für die negativen; die Variablenzahl beim SI-ZW betrug 40 für die positive und 40 für die negative Beschreibung; insgesamt hatten wir also über 500 Variablen mit je zwei Ausprägungsgraden (genannt bzw. angekreuzt bzw. oben angekreuzt vs. nicht genannt bzw. nicht angekreuzt bzw. unten angekreuzt). Letztlich bleibt beim Ausfüllen des SELF-BF für die Testpersonen wenig durchschaubar, was genau erfasst werden soll („Wir möchten Sie als Person näher kennenlernen“). Gleichwohl dürfte – kurz gesagt – die verhaltensleitende Formel beim Beantworten der drei Subtests sein: *„Was mir nahe liegt, fällt mir schnell ein, erkenne ich sofort, kann ich spontan entscheiden“*. Immerhin erreichten wir bei all unseren Untersuchungen mit dieser Art der Verwendung von Eigenschaftsworten eine so nicht erwartete Erhöhung der Akzeptanz: Die Testpersonen berichteten nach Test-Absolvierung oft ungefragt über ihr durch den SELF-BF angeregtes Interesse und ihre Befriedigung darüber, dass sie sich selbst einmal in dieser Breite selbstreflektorisch hätten betrachten und beschreiben können – ganz nach der Devise „Wer bin ich eigentlich?“.

Diese Erfahrungen passen gut zu einem erst kürzlich von Steinmayr und Kersting (in diesem Band) gefundenen Zusammenhang zwischen Akzeptanz und Verfälschungsgefährdung: Je weniger ein Verfahren als verfälschbar bzw. verfälschungswürdig wahrgenommen wird, umso mehr Akzeptanz erfährt es.

## **DATENERHEBUNG**

Bei einer namhaften Firma für Finanzdienstleistungen konnten wir 178 Berater in anonymer Forschungssituation mit den zwei Instrumenten SELF-BF und NEO-FFI befragen. Den Testpersonen wurde wahrheitsgemäß gesagt, dass die Erhebung dazu diene, für neue Bewerber, die recht häufig aus anderen Branchen kommen, einen brauchbaren Normkorridor für wichtige persönliche Voraussetzungen eines Finanzberaters zu gewinnen. Die Leistungswerte der befragten Finanzberater streuten von niedrig bis hoch, in etwa gemäß der entsprechenden Verteilung bei der Gesamtheit der Berater dieser Firma; 81% der Befragten waren Männer, 19% Frauen; über das Alter konnten wir aus Datenschutzgründen keine Information bekommen; es lag aber nach unserer Einschätzung bei im Mittel 35 Jahren.

## **BILDUNG DER MERKMALSSCORES FÜR DEN SELF-BF**

Auf Basis der so erhaltenen Daten bildeten wir zunächst Extremgruppen, und zwar für jedes einzelne der fünf Persönlichkeitsmerkmale des NEO-FFI: jeweils die 30% Personen mit den niedrigsten und die 30% mit den höchsten Werten im betrachteten Merkmal.

Für jedes Merkmal unterzogen wir dann die entsprechenden Extremgruppen (die „graue Mitte“ der verbliebenen 40% ließen wir immer außer Betracht) einem kontrastierenden Vergleich, und zwar für jedes unserer drei Instrumente OSKT, SI-FW und SI-ZW getrennt.

Dabei wurden pro Instrument sowohl ein linearer Score (gewichtete Summe prädiktiver Items), ein konfiguraler Score (gewichtete Summe kriteriumsdiskriminativer Itemkonfigurati-

onen) und ein Summenscore aus beiden gebildet. Die Besonderheit dabei war der Einbezug eines konfiguralen Scores: Das Entdecken und Heranziehen von auf Konfigurationen basierenden Trennungen zwischen den Extremgruppen sollte die Möglichkeit kriteriumsrelevanter Varianzerhöhung eröffnen. Dies war auch durchgehend der Fall, so dass wir jeweils mit dem Summenscore (linearer Score plus konfiguraler Score) deutlich bessere Diskriminationen der Extremgruppen erzielen konnten als mit dem linearen Score allein.

Schließlich fassten wir pro Persönlichkeitsmerkmal die Summenscores aller drei Instrumente (OSKT, SI-FW, SI-ZW) zu einem Gesamtscore zusammen, der nun als Testscore des SELF-BF für das betrachtete Persönlichkeitsmerkmal herangezogen wurde, um ihn bei von dieser Stichprobe unabhängigen anderen Stichproben (s.u.) einzusetzen. Dies erschien uns gewinnbringend, da in der Anwendung auf die hier in Rede stehende Stichprobe insgesamt (N = 178) bei jedem Persönlichkeitsmerkmal alle drei Instrumente (als Prädiktoren) untereinander geringer korrelierten (im Mittel etwa .30) als mit dem zugehörigen Persönlichkeitsmerkmal (Kriterium) (im Mittel etwa .60).

## ZUR KONSTRUKTVALIDITÄT DES SELF-BF

Um zu sehen, ob die fünf so gewonnenen Merkmalsscores des SELF-BF mit den 5 entsprechenden Persönlichkeitsmerkmalen des NEO-FFI hinreichend gut zusammenhängen, zogen wir eine Stichprobe aus echten Auswahl-situationen derselben Firma für dieselbe Funktion (Finanzberater) heran, nämlich N = 223 angenommene Bewerber. Diese Personen sind mit denen der Forschungssituation (N = 178) insoweit vergleichbar, als sie nach dem gleichen Auswahlprozedere (AC) eingestellt wurden wie die Personen der Forschungssituation ein paar Jahre zuvor, bei im übrigen weitgehend unverändertem Bewerbermarkt. Die Korrelationen der Scores des SELF-BF und des NEO-FFI für jedes der 5 Merkmale sind in Tabelle 1 aufgeführt.

**Tabelle 1:** Korrelationen des SELF-BF mit dem NEO-FFI für jedes der 5 Persönlichkeitsmerkmale über die N = 223 Angenommenen

|                           | <i>echte<br/>Auswahl-<br/>situation</i> |
|---------------------------|---|
|                           | <i>N+ = 223<br/>(Angenommene)</i>       |
| <i>Neurotizismus</i>      | <i>.43</i>                              |
| <i>Extraversion</i>       | <i>.53</i>                              |
| <i>Offenheit</i>          | <i>.41</i>                              |
| <i>Verträglichkeit</i>    | <i>.54</i>                              |
| <i>Gewissenhaftigkeit</i> | <i>.56</i>                              |

Zwar ist die Höhe dieser Korrelationen noch nicht befriedigend, immerhin aber liegen schon drei der fünf Korrelationen über .50, nämlich die für Gewissenhaftigkeit, Verträglichkeit und Extraversion.

Zum Zwecke einer Abschätzung der Stichprobenspezifität dieser Befunde haben wir zwei weitere Stichproben untersucht: Einmal ermittelten wir auch für die N = 1116 Abgelehnten die entsprechenden Korrelationen (s. Tabelle 2, mittlere Spalte). Zum anderen haben wir in einer anonymen Forschungssituation eine disperse, bei einem anderen Projekt angefallene Stichprobe jüngerer Erwachsener (zwischen 21 und 26 Jahren, ca. 2/3 Männer, 1/3 Frauen, ca. 1/5 Studenten) mit dem SELF-BF und dem BFI-K, der Kurzform eines anderen Persönlichkeitsinventars für die Big Five (Rammstedt & John, 2005), befragt und die entsprechenden Korrelationen bestimmt (s. Tabelle 2, rechte Spalte).

**Tabelle 2:** Korrelationen des SELF-BF mit dem NEO-FFI für jedes der 5 Persönlichkeitsmerkmale jeweils über die N = 223 Angenommenen und die N = 1116 Abgelehnten sowie des SELF-BF mit dem BFI-K über N = 485 junge Erwachsene

|                           | <i>echte<br/>Auswahl-<br/>situation</i> | <i>echte<br/>Auswahl-<br/>situation</i> | <i>anonyme<br/>Forschungs-<br/>situation</i> |
|---------------------------|---|---|--|
|                           | <i>N+ = 223</i>                         | <i>N- = 1116</i>                        | <i>N = 485</i>                               |
| <i>Neurotizismus</i>      | <i>.43</i>                              | <i>.45</i>                              | <i>.41</i>                                   |
| <i>Extraversion</i>       | <i>.53</i>                              | <i>.58</i>                              | <i>.54</i>                                   |
| <i>Offenheit</i>          | <i>.41</i>                              | <i>.46</i>                              | <i>.23</i>                                   |
| <i>Verträglichkeit</i>    | <i>.54</i>                              | <i>.47</i>                              | <i>.43</i>                                   |
| <i>Gewissenhaftigkeit</i> | <i>.56</i>                              | <i>.59</i>                              | <i>.58</i>                                   |

Alle drei Datensätze führen zu einem recht ähnlichen Bild: Gewissenhaftigkeit und Extraversion scheinen noch am besten erfasst zu werden, Verträglichkeit, Offenheit und Neurotizismus dagegen schneiden schwächer ab.

## ZU VERFÄLSCHUNGSTENDENZEN

Aus Tabelle 3 lässt sich ersehen, wie sich für die Big Five des NEO-FFI die Mittelwerte in der Auswahl-situation im Vergleich zur Forschungssituation alle signifikant und z.T. auch substantiell in sozial erwünschter Richtung ändern – eine weitere Bestätigung des aus der Forschungsliteratur bekannten Befunds (Pauls & Crost, 2005). Da nun vor allem Gewissenhaftigkeit und mit Einschränkungen auch Neurotizismus die beiden Big Five-Faktoren sind, die über alle Berufe gesehen am ehesten mit beruflicher Leistung korrelieren (Barrick, Mount & Judge, 2001), sind hier die Beschönigungstendenzen von besonderem Interesse. Am stärksten ist die Beschönigungstendenz bei Gewissenhaftigkeit, am geringsten bei Neuroti-

zismus (was allerdings in diesen Personenstichproben von Finanzberatern bzw. entsprechenden Bewerbern vermutlich mit der starken Varianzeinschränkung durch die niedrigen Mittelwerte bei Neurotizismus zusammenhängt).

**Tabelle 3: Vergleich von NEO-FFI und SELF-BF im Hinblick auf die Big Five-Mittelwerte und deren Differenzen zwischen Forschungssituation (N=178) und Auswahl-situation (N+ = 223)**

|                                      |                           | <i>anonyme<br/>Forschungs-<br/>situation</i> | <i>echte<br/>Auswahl-<br/>situation</i> | <i>Differenzen:<br/>Signifikanzen und<br/>Effektgrößen</i> |                |
|--------------------------------------|---------------------------|--|---|--|----------------|
|                                      |                           | <i>N = 178</i>                               | <i>N+ = 223</i>                         |  |                |
| <i>Big Five<br/>nach<br/>NEO-FFI</i> | <i>Neurotizismus</i>      | <i>1,02</i>                                  | <i>0,87</i>                             | <i>**</i>  | <i>d= -.30</i> |
|                                      | <i>Extraversion</i>       | <i>2,66</i>                                  | <i>2,88</i>                             | <i>***</i>   | <i>d= -.56</i> |
|                                      | <i>Offenheit</i>          | <i>2,43</i>                                  | <i>2,60</i>                             | <i>***</i>   | <i>d= -.39</i> |
|                                      | <i>Verträglichkeit</i>    | <i>2,64</i>                                  | <i>2,94</i>                             | <i>***</i>   | <i>d= -.72</i> |
|                                      | <i>Gewissenhaftigkeit</i> | <i>2,95</i>                                  | <i>3,29</i>                             | <i>***</i>   | <i>d= -.80</i> |
| <i>Big Five<br/>nach<br/>SELF-BF</i> | <i>Neurotizismus</i>      | <i>18,13</i>                                 | <i>14,20</i>                            | <i>***</i>   | <i>d= -.47</i> |
|                                      | <i>Extraversion</i>       | <i>39,32</i>                                 | <i>44,60</i>                            | <i>**</i>  | <i>d= -.39</i> |
|                                      | <i>Offenheit</i>          | <i>19,25</i>                                 | <i>18,57</i>                            | <i>n.s.</i>  | <i>d= -.08</i> |
|                                      | <i>Verträglichkeit</i>    | <i>35,04</i>                                 | <i>36,30</i>                            | <i>n.s.</i>  | <i>d= -.10</i> |
|                                      | <i>Gewissenhaftigkeit</i> | <i>41,98</i>                                 | <i>49,75</i>                            | <i>***</i>   | <i>d= -.49</i> |

Nun meinen Rauch et al. (in diesem Band) unter Rekurs auf Bing et al. (2004), dass der „Bewerbungseffekt“ (= Beschönigungstendenz, also höhere Mittelwerte in für wichtig angesehenen Skalen) dadurch zustande komme, dass Persönlichkeitsfragen in der Bewerbungssituation anders interpretiert würden als in der Forschungssituation: Während bei letzterer kein spezifischer Kontext zur Beurteilung des eigenen Verhaltens vorgegeben werde, werde in der Bewerbungssituation das eigene Verhalten antizipierend auf den Arbeitsplatz bezogen. Aber: Dieses Argument dürfte zumindest auf unsere Untersuchungen nicht zutreffen: Wir haben auch in der ursprünglichen Forschungssituation den Testpersonen gesagt, dass wir ihre Daten benötigten, um für zukünftige Bewerber (zum selbständigen Finanzberater) bessere Normwerte ermitteln zu können.

Um Beschönigungstendenzen zu begegnen, so wird gern argumentiert, könne man durch Heranziehen entsprechender Normwerte die Mittelwertsverschiebung leicht kompensieren. Dem aber hält Kanning – wie eingangs schon vermerkt und unseres Erachtens zu Recht – den

Einwand mangelnder Fairness entgegen: dass wahrscheinlich eben nicht *alle* Bewerber ihre Antworten in gleichem Maße „beschönigen“ und dass somit ihre Chancen angenommen zu werden unbilligerweise verringert werden.

Ganz frei von Erwünschtheitstendenzen ist der SELF-BF allerdings auch nicht, doch ist seine diesbezügliche Anfälligkeit geringer als bei dem inhaltsdirekten Fragebogen NEO-FFI, wie die schwächeren Signifikanzen und kleineren Effektgrößen zeigen.

Ein interessanter Befund mit Bezug auf die Auswahl-situation noch am Rande: Weder im NEO-FFI noch im SELF-BF trat ein einziger signifikanter Unterschied in den Mittelwerten der Big Five zwischen den 223 angenommenen und den 1116 abgelehnten Bewerbern auf (s. Tabelle 4). Im Hinblick auf diese Persönlichkeitsmerkmale, ob so oder so gemessen, wurden die Bewerber per AC demnach nicht selektiert. Diese Information aus *Bewerbungssituationen* kontrastiert mit der Information aus *Bewährungssituationen*, dass nämlich die Big Five das Potenzial zu prädiktiver Validität hätten (Barrick, Mount & Judge, 2001). Vielleicht ist es doch so, dass die Konstrukte der Big Five zu breit sind für spezifische Berufe und diesbezüglich die vorgängige Selbstselektion (via Attraction sensu Schneider, 1987) allein schon das Ihre getan hat.

**Tabelle 4: Mittelwerte der 223 Angenommenen und der 1116 Abgelehnten im NEO-FFI und im SELF-BF**

|                              |                           | <i>Echte Auswahl-situation</i> |                  |
|------------------------------|---------------------------|--------------------------------|------------------|
|                              |                           | <i>N+ = 223</i>                | <i>N- = 1116</i> |
| <i>Big Five nach NEO-FFI</i> | <i>Neurotizismus</i>      | 0,87                           | 0,88             |
|                              | <i>Extraversion</i>       | 2,88                           | 2,86             |
|                              | <i>Offenheit</i>          | 2,60                           | 2,64             |
|                              | <i>Verträglichkeit</i>    | 2,94                           | 2,90             |
|                              | <i>Gewissenhaftigkeit</i> | 3,29                           | 3,29             |

|                              |                           |       |       |
|------------------------------|---------------------------|-------|-------|
| <i>Big Five nach SELF-BF</i> | <i>Neurotizismus</i>      | 14,20 | 14,71 |
|                              | <i>Extraversion</i>       | 44,60 | 43,67 |
|                              | <i>Offenheit</i>          | 18,57 | 18,48 |
|                              | <i>Verträglichkeit</i>    | 36,30 | 36,03 |
|                              | <i>Gewissenhaftigkeit</i> | 49,75 | 49,32 |

## ABSCHLIESSENDE BEMERKUNGEN

Schon von der Self-Efficacy-Forschung war zu erfahren (Bandura, 1986), dass Selbst-Konzepte (mit)erklären, warum Leistung und Zufriedenheit im Beruf auch bei Menschen, die gleiches Wissen und gleiche Fertigkeiten besitzen, so verschieden sein können. Menschen entwickeln ihre Selbst-Konzepte, indem sie ihr eigenes Verhalten und das Verhalten anderer vergleichend dazu beobachten und die Reaktionen der anderen ihnen gegenüber registrieren. Mit dem, was ein Individuum schließlich von sich denkt (und fühlt), mit seinem Selbstkonzept also, ordnet es sich in seine physische und soziale Erfahrungswelt ein, so dass darüber auch Aussagen über wichtige globale Persönlichkeitsfaktoren wie die Big Five möglich werden sollten.

Wir haben bei unserem Verfahren bewusst darauf verzichtet, zur Erfassung des Selbst-Konzepts lediglich eine vorhandene Adjektiv-Checkliste (wie etwa die mit 100 Items von Goldberg, 1992) heranzuziehen, sozusagen als Schnellverfahren zur Erfassung der Big Five. Denn einerseits wollten wir mit deutlich mehr Items (über 500) die Durchschaubarkeit erschweren und andererseits die Chancen für bessere Differenzierungen vergrößern.

Dabei haben wir die linearen Diskriminationen ergänzt durch kriteriumsrelevante Konfigurationen, die diverse, qualitativ unterschiedliche, aber in sich homogene Personengruppen identifizieren (vgl. dazu Montel, in diesem Band). Dies liegt zwar nicht auf der Linie methoden-ideologisch imprägnierter Überzeugungen von „korrekten“ Datenanalysen. Wir ließen uns aber von der Hoffnung leiten, dass in den Daten mehr stabile trennscharfe Informationen zu finden sein könnten als das, was die kompensatorischen linearen Modelle zu entdecken in der Lage sind – ganz nach der Einsicht des Pragmatisten John Dewey (1910), dass wir den fraglichen Gegenstand umso besser erkennen, je mehr Wechselwirkungen wir feststellen; in dieselbe Richtung weist übrigens die Überzeugung Cronbachs (1975), dass Interaktionen höherer Ordnung oftmals bedeutsamer sind als die Haupteffekte.

Derzeit sind wir dabei, Verbesserungen der Konstruktvalidität des SELF-BF für den angepeilten Personenkreis „Bewerber für qualifizierte Jobs“ zu erreichen. Dazu gehören im Wesentlichen:

- Modifikationen des Itempools,
- Untersuchung etlicher großer Stichproben, die verschiedene Gruppen von Bewerbern für qualifizierte Jobs (Professionals, Management-Nachwuchs, Vertriebsfunktionen) repräsentieren,
- neue und umfangreichere Kalibrierung der Merkmalsscores für die Big Five,
- Reliabilitäten – dabei auch: für welche Untergruppen von Personen welche Items besonders stabil über die Zeit sind – nicht nur, wie solche Kennwerte über eine gesamt-populationsrepräsentative Stichprobe aussehen,
- Faktorenstruktur, konvergente und diskriminante Validität,
- Heranziehen neuerer Analyseverfahren wie Künstliche Neuronale Netze zur umfassenderen Entdeckung nicht-linearer Zusammenhänge (Krumm & Schmidt-Atzert, in diesem Band).

Kritisch bleibt anzumerken, dass auch der SELF-BF ein Fragebogen ist und damit nur solche Informationen erfassen kann, die explizit, also der Introspektion zugänglich sind. Wenn man das gesamte Wissen (inkl. Bewerten) einer Person mit Bezug auf ihr Selbst erfahren möchte, um z.B. die Konvergenz/Divergenz von explizitem und implizitem selbstbezogenem Wissen abzuschätzen (Gawronski & Bodenhausen, 2005), muss man zusätzlich auf Verfahren zur impliziten Messung der Selbstwahrnehmung zurückgreifen (z.B.: IATs, s. Rudolph, Schröder & Schütz, in diesem Band; Operante Tests, s. Scheffer, in diesem Band) – sobald derartige Verfahren diese Informationen hinreichend valide zu liefern imstande sind.

## LITERATUR

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barrick, M.R., Mount, M.K. & Judge, T. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9-30.
- Baumeister, R.F., Campbell, J.D., Krueger, J.I., & Vohs, K.D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4, 1-44.
- Bing, M.N., Whanger, J.C., Davison, H.K. & VanHook, J.B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology*, 89, 150-157.
- Borkenau, P., Egloff, B., Eid, M., Hennig, J., Kersting, M., Neubauer, A.C. & Spinath, F.M. (2005). Persönlichkeitspsychologie: Stand und Perspektiven. *Psychologische Rundschau*, 56, 271-290.
- Christiansen, N., Burns, G.N. & Montgomery, G.E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance*, 18, 267-307.
- Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.
- Delia, J.G. & Grossberg, L. (1977). Interpretation and evidence. *Western Journal of Speech Communication*, 41, 37.
- Dewey, J. (1910). *How we think*. Boston: Heath.
- Gawronski, B. & Bodenhausen, G.V. (2005). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731.
- Ghiselli, E.E. (1954). The forced-choice technique in self description. *Personnel Psychology*, 7, 201-208.
- Goldberg, L.R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Gough, H.G. (1997). *The Adjective Check List (ACL)*. Palo Alto, CA: Consulting Psychologists Press.
- Hough, L.M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, 11, 209-244.
- John, O.P. (1990). The „big five“ factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L.A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford Press.
- John, O.P. & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L.A. Pervin & O.P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102-138). New York: Guilford.
- Kanning, U. (2003). Sieben Anmerkungen zum Problem der Selbstdarstellungen in der Personalauswahl. *Zeitschrift für Personalpsychologie*, 2, 193-197.
- Kubinger, K.D. (2002). On faking personality inventories. *Psychologische Beiträge*, 44, 10-16.
- Marcus, B. (2003). Das Wunder sozialer Erwünschtheit in der Personalauswahl. *Zeitschrift für Personalpsychologie*, 2, 129-132.
- Pauls, C.A. & Crost, N.W. (2005). Effects of different instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences*, 39, 297-308.

- Rammstedt, B. & John, O.P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit. *Diagnostica*, 51, 195–206.
- Rheinberg, F. & Stiensmeier-Pelster, J. (Hrsg.). (2003). *Diagnostik von Motivation und Selbstkonzept*. Göttingen: Hogrefe.
- Rosse, J.G., Stecher, M.D., Miller, J.L. & Levin, R.A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634–644.
- Sarges, W. (Hrsg.). (2000). *Management-Diagnostik* (3. Aufl.). Göttingen: Hogrefe.
- Sarges, W. (2006). Management-Diagnostik. In F. Petermann & M. Eid (Hrsg.), *Psychologische Diagnostik* (S. 739–746). Göttingen: Hogrefe.
- Schippmann, J.S. & Prien, E.P. (1985). The Ghiselli Self-Description Inventory: A psychometric appraisal. *Psychological Reports*, 57, 1171–1177.
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437–453.
- Swann, W.B., Chang-Schneider, C. & Larsen McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist*, 62, 84–94.
- Viswesvaran, C. & Ones, D.S. (1999). Meta-analysis of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59, 197–210.

### ***Ergänzender Hinweis:***

*Für den Einsatz in der Wirtschaftspraxis haben wir das Verfahren weiter entwickelt: in Richtung auf Bedarfe und Wünsche wichtiger Testnutzer. Die ursprünglichen Skalen (Neurotizismus, Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit) haben wir dazu in weiteren Forschungsschritten zu fünf in der Praxis auch von psychologischen Laien als relevant empfundenen Merkmalen transformiert bzw. die Skalenbeschreibungen semantisch von psychologischen Fachtermini weitgehend befreit, so dass der SELF-BF in seiner aktuell eingesetzten Form die Merkmale "Extraversion", "Gestaltungswille", "Stressresistenz", "Gewissenhaftigkeit" und "Offenheit" misst und entsprechende Ausprägungsgrade zurückmeldet.*