

Instrumente der Arbeits- und Organisationspsychologie

START-P: Testbatterie für Berufseinsteiger – Persönlichkeit

Meike Wenzel

Zusammenfassung. START-P (Beauducel & Kersting, 2010) ist ein psychometrisch überprüfter Persönlichkeitsfragebogen, der orientiert an der DIN 33430 für die Zielgruppe *Berufseinsteiger* entwickelt wurde. Das Verfahren dient der standardisierten Selbsteinschätzung der Persönlichkeit und kann im Kontext berufsbezogener Entscheidungen eingesetzt werden. Es werden die Module *Verhaltens-Input* und *Verhaltens-Output* unterschieden, die als Gesamtversion oder aber als Einzelmodul *Verhaltens-Output* vorgegeben werden können. Die Autoren legen vielfältige Befunde zur Konstrukt- und Kriteriumsvalidität vor, die den Einsatz von START-P in der Persönlichkeitsdiagnostik untermauern. Aufgrund der zielgruppenspezifischen Formulierung der Items und der Normierung des Verfahrens an 14- bis 32-Jährigen kann der Einsatz von START-P bei Berufseinsteigern empfohlen werden.

Schlüsselwörter: Berufseignungsdiagnostik, Persönlichkeit, Persönlichkeitsdiagnostik, Verhaltens-Input, Verhaltens-Output

START-P: Test battery for job starters – personality

Abstract. START-P (Beauducel & Kersting, 2010) is a psychometrically validated personality questionnaire, which was designed for job starters in consideration of DIN 33430. The instrument measures personality by self-assessment in a standardized way. START-P can be used in the context of occupational decisions. The instrument consists of the modules *behavior-input* and *behavior-output*; the behavior-output module can be used on its own or in combination with the behavior-input module. The authors present a variety of research results addressing the construct validity and the predictive validity, which show that START-P can be useful in measuring personality traits. The items are formulated with a focus on the target group. The instrument was standardized on the basis of people aged 14 to 32 years. Thus, the use of START-P with job starters can be recommended.

Key words: occupational diagnostic, personality, personality measurement, behavior-input, behavior-output

Beauducel, A. & Kersting, M. (2010). START-P – Testbatterie für Berufseinsteiger – Persönlichkeit. Göttingen: Hogrefe.

Ziele des Verfahrens

Wenngleich in der Vergangenheit Persönlichkeitsfragebogen in der Berufseignungsdiagnostik nur zögernd eingesetzt wurden, nehmen die Vorbehalte gegenüber Verfahren zur Persönlichkeitsdiagnostik seit Entwicklung berufsbezogener Verfahren ab (Manual S. 7–8). Derzeit existiert bereits eine Vielzahl von Verfahren zur Persönlichkeitsdiagnostik, die speziell für den Einsatz im Berufskontext entwickelt wurden, beispielsweise das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (Hossiep & Paschen, 2003) oder, im Bereich

der onlinebasierten Verfahren, die Internettestplattform PERLS (Kirbach & Montel, 2005).

Mit dem START-P Inventar („P“ steht für Persönlichkeit) von Beauducel und Kersting (2010) liegt ein weiteres Instrument zur Persönlichkeitsdiagnostik vor. Das Verfahren ist als ein Modul im Rahmen einer umfassenden Testbatterie für Berufseinsteiger, Studienanfänger, Jugendliche und junge Erwachsene konzipiert. START-P dient der standardisierten Selbsteinschätzung der Persönlichkeit und wird im Kontext berufsbezogener Entscheidungen eingesetzt. Im Vergleich zu anderen Persönlichkeitsinventaren fokussiert das START-P Inventar jüngere Erwachsene und berücksichtigt die Befunde zum Zusammenhang zwischen Persönlichkeit und Ausbildungs-/Berufserfolg. Ferner liegen spezielle Normen für den Bewerbungskontext vor.

Theoretischer Hintergrund

START-P besteht aus zwei Modulen, dem Modul *Verhaltens-Input* sowie dem Modul *Verhaltens-Output*, und unterscheidet Input- und Output-bezogene Persönlichkeitseigenschaften. Dieser Zweiteilung liegt die Annahme Allports (1961, zitiert nach Beauducel & Kersting, 2010) zugrunde, nach der Eigenschaften einen funktionalen Zusammenhang zwischen verschiedenen Reizen herstellen sowie äquivalente Reaktionen herleiten. Auf der Input-Seite erfolgt somit die Bewertung von Reizen, auf der Output-Seite werden die Reaktionstendenzen fokussiert. Demnach beinhaltet das Modul *Verhaltens-Input* Items, die sich auf die Bewertung von Reizen beziehen, das Modul *Verhaltens-Output* hingegen umfasst Items zur Erfassung der Reaktionstendenz.

Das Modul *Verhaltens-Input* basiert auf der Verstärkungs-Sensitivitätstheorie von J. A. Gray (1991, zitiert nach Beauducel & Kersting, 2010). In der Theorie werden zwei biopsychologische Systeme unterschieden, die die Reaktionen auf konditionierte Stimuli beeinflussen. Das Verhaltens-Annäherungs-System bestimmt die Sensitivität gegenüber Belohnungsreizen, reagiert aber auch auf den Wegfall von Strafreizen. Das Verhaltens-Hemm-System nimmt Einfluss auf die Sensitivität gegenüber Bestrafungsreizen, reagiert aber auch auf den Wegfall von Belohnungsreizen. Ferner nehmen die Autoren Bezug zu den Quellen der Verstärkung. Sie gehen davon aus, dass interindividuelle Unterschiede in Bezug auf die Sensitivität von Reizen bestehen, die von Personen unterschiedlicher Hierarchiestufen ausgehen. Aus den Modellannahmen wird ein Facetten-Design abgeleitet, das zwischen den Verstärkungsarten Belohnung/Aufhebung von Bestrafung sowie Bestrafung/Aufhebung von Belohnung und der Quelle der Verstärkung Ausbilder/Vorgesetzter sowie Team unterscheidet.

Die Grundlage für das Modul *Verhaltens-Output* bilden sowohl das Fünf-Faktoren-Modell der Persönlichkeit sowie metaanalytische Befunde zum Zusammenhang zwischen Persönlichkeitseigenschaften und beruflicher Eignung. Die Autoren verweisen u. a. auf die Ergebnisse von Barrick und Mount (1991), die einen Einfluss von Persönlichkeitsmerkmalen auf den Berufserfolg aufzeigen. Kersting (2005) stellt drei Persönlichkeitsdimensionen heraus, die für die berufliche Eignung relevant sind. Die genannten Dimensionen Gewissenhaftigkeit, emotionale Stabilität sowie Extraversion werden mit sechs Skalen im START-P abgebildet.

Testkonstruktion

Die Konstruktion des START-P erfolgte in drei Schritten. Zunächst wurden auf den theoretischen Modellannahmen 160 Items, die sich jeweils zur Hälfte auf die beiden Module aufteilen, generiert. Anschließend sind die 160 Items in verschiedenen Datenerhebungen eingesetzt worden. Auf dieser Datenbasis aufbauend wurden Faktorenanalysen, Trennschärfe-Analysen sowie Reliabili-

tätsanalysen durchgeführt. Zusätzlich erfolgte die Analyse der Facettenstruktur des Moduls *Verhaltens-Input* mithilfe konfirmatorischer Faktorenanalysen. Die Ergebnisse führten zur Umformulierung von Items sowie zur Reduzierung der Itemanzahl auf 152 Items. Im letzten Schritt wurden die 152 Items bei insgesamt 2.063 Personen eingesetzt und anschließend weitere faktorenanalytische Auswertungen und Trennschärfe- und Itemanalysen durchgeführt. Die nun vorliegende Endversion umfasst 86 Items.

Die Trennschärfen aller im START-P erfassten Items liegen zwischen .11 und .72. Dabei zeigen die Items des Moduls *Verhaltens-Input* geringere Trennschärfen im Vergleich zum Modul *Verhaltens-Output*. Dies wird auf die Facettenstruktur zurückgeführt, bei der jedes Item für zwei Faktoren relevant ist.

Aufbau des Verfahrens

START-P besteht insgesamt aus 86 Items. Diese werden auf einer sechsstufigen Antwortskala von *trifft überhaupt nicht zu* bis *trifft vollständig zu* bewertet.

Das Modul *Verhaltens-Input* umfasst 38 Items, die die emotionalen und kognitiven Reaktionen auf unterschiedliche situative Bedingungen oder Reize erfassen und in ein Facetten-Design eingebettet sind. Dieses Modul beinhaltet die folgenden Grunddimensionen:

- *Belohnungssensitivität*, operationalisiert als Grad an Sensitivität gegenüber Belohnungsreizen
- *Bestrafungssensitivität*, operationalisiert als Grad an Sensitivität gegenüber Bestrafungsreizen
- *Sensitivität gegenüber Vorgesetzten und Ausbildern*, operationalisiert als erhöhte Sensitivität auf Belohnungs- und Bestrafungsreize durch diese Personengruppen
- *Sensitivität gegenüber dem Team*, operationalisiert als erhöhte Sensitivität auf Belohnungs- und Bestrafungsreize durch das Team bzw. die Arbeitsgruppe

In das Facetten-Design eingebundene Beispielitems zeigt Tabelle 1.

Das Modul *Verhaltens-Output* umfasst 48 Items. Es werden sechs Persönlichkeitseigenschaften erfasst:

- *Emotionale Stabilität*, operationalisiert als Stabilität des Verhaltens bei Belastungen in der Schule, im Studium oder im Beruf
- *Empathie*, operationalisiert als Gespür für die Bedürfnisse und Kompetenzen von Personen
- *Kontaktorientierung*, operationalisiert als ausbildungs- und berufsbezogene Kontaktorientierung
- *Durchsetzungsvermögen*, operationalisiert als Durchsetzungsvermögen und Durchsetzungsneigung eigener Entscheidungen und Ideen im Ausbildungs- und Berufskontext

Tabelle 1. Beispielitems im Facetten-Design der Verstärkungsarten und Verstärkungsquellen (Manual, S. 18)

		Verstärkungsart	
		Belohnung/Aufhebung von Bestrafung	Bestrafung/Aufhebung von Belohnung
Quelle der Verstärkung	Ausbilder/	Die Vorstellung, dass mich meine	Wenn ich mir vorstelle, dass mich meine
	Vorgesetzte	Ausbilder/Vorgesetzte mögen, tut mir sehr gut.	Ausbilder/Vorgesetzte kritisieren, berührt mich das sehr.
	Team	Wenn man in einer Arbeitsgruppe gut über mich redet, bin ich meist richtig begeistert.	Ich bin sehr unglücklich, wenn ich in meinem Team abgelehnt werde.

Tabelle 2. Beispielitems für die Dimensionen des Moduls *Verhaltens-Output*

Dimension	Beispielitem
Durchsetzungsvermögen	Wenn eine Sache heftig diskutiert wird, vertrete ich meinen Standpunkt auch gegen Widerstände.
Emotionale Stabilität	Auch wenn ich viel zu tun habe, bleibe ich ganz locker.
Empathie	Ich kann mich gut gedanklich in andere hineinversetzen.
Kontaktorientierung	Es macht mir Spaß, mit mir unbekannten Menschen ein Gespräch zu führen.
Leistungsmotivation	Ich investiere nicht mehr Mühe in eine Arbeit als unbedingt notwendig.
Sorgfalt	Ich plane meine Arbeit stets genau.

- *Sorgfalt*, operationalisiert als Sorgfalt und Genauigkeit bei der Arbeit
- *Leistungsmotivation*, operationalisiert als Intensität, mit der Leistungen in der Ausbildung oder dem Beruf verfolgt werden

Beispielitems zu den Dimensionen des Moduls *Verhaltens-Output* sind in Tabelle 2 zusammengestellt.

START-P ist in zwei Versionen verfügbar. Version A („A“ steht für Ausbildung/Arbeit) ist für ausbildungsheterogene Gruppen geeignet und verwendet die berufsbezogenen Begriffe „Ausbilder/innen“ und „Vorgesetzte“. Version S („S“ steht für Schule/Ausbildung) ist an die Zielgruppe Schüler und Studenten adressiert und verwendet die Begriffe „Lehrer/innen“ und „Ausbilder/innen“.

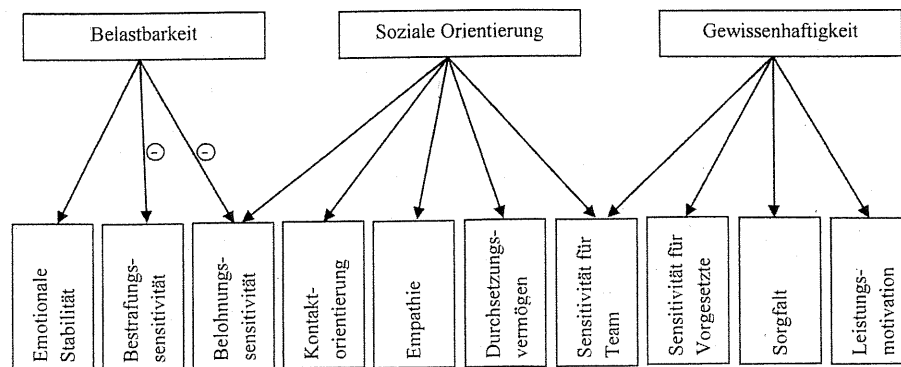
Die im START-P erfassten Dimensionen sind in ein hierarchisches Modell eingebettet. Die zehn Primär- bzw.

Grunddimensionen werden in drei Globaldimensionen (Belastbarkeit, Soziale Orientierung, Gewissenhaftigkeit) zusammengefasst (vgl. Abbildung 1).

Durchführung

START-P kann sowohl als Gruppentestung mit Testleiter als auch allein in Selbstadministration durchgeführt werden. Das Verfahren wird ohne Zeitvorgaben bearbeitet, die maximale Bearbeitungsdauer wird auf 40 Minuten geschätzt.

Das Modul *Verhaltens-Input* kann nur in Kombination mit dem Modul *Verhaltens-Output* verwendet werden. Dies begründen die Autoren mit der komplexen Itemformulierung, die bei separater Durchführung die Teilnehmerinnen und Teilnehmer stark belasten und die Akzeptanz des Verfahrens gefährden können. Das Modul *Verhaltens-*



⊖ steht für negative Ladungen

Abbildung 1. Grund- und Globaldimensionen des START-P (in Anlehnung an das Manual, S. 24).

Output hingegen kann separat vorgegeben werden. Dieses liegt auch in einer Kurzversion mit 24 Items vor. Die Bearbeitungsdauer für das Modul *Verhaltens-Output* wird auf maximal 20 Minuten, für das Kurz-Modul *Verhaltens-Output* auf 10 Minuten geschätzt.

START-P liegt sowohl als Paper-Pencil- als auch als PC- und Internet-Version vor.

Auswertung

In der Paper-Pencil-Version wird die Auswertung über das Auswertungsprogramm vorgenommen. Bei PC- oder Internettestungen erfolgt die Auswertung automatisch. Die Auswertung gibt grundsätzlich Faktorwerte, teilweise auch Skalenwerte aus. Bei Einsatz beider Module werden sowohl die Faktorwerte für die zehn Primärdimensionen als auch für die Globaldimensionen ausgegeben. Im Rahmen der Auswertung kann zwischen verschiedenen Normgruppen gewählt werden. Es stehen die Gesamtnormen sowie alters-, geschlechts- und bildungsspezifische (Gymnasiasten und Nicht-Gymnasiasten) Normen zur Verfügung. Ferner werden gesonderte Normen für den Bewerbungskontext angeboten. Neben den Normgruppen kann auch zwischen verschiedenen Skalenumrechnungen gewählt werden. Zur Verfügung stehen u. a. Prozentränge, Z-Werte und Stanine-Werte. Voreingestellt sind Standardwerte mit Mittelwert 100 und Standardabweichung 10.

Im Testreport (Norminterpretation) wird das individuelle Testprofil mit unterschiedlichen Textbausteinen für unterdurchschnittliche (Prozentrang 0–25), durchschnittliche (Prozentrang 25–75) sowie überdurchschnittliche (Prozentrang 75–100) Testwerte erläutert. Darüber hinaus werden die Testergebnisse im Profilblatt graphisch dargestellt. In der Profilansicht können mehrere Testteilnehmer miteinander verglichen werden.

Gütekriterien

Objektivität

START-P ist ein standardisiertes Testverfahren. Bei Einhaltung der Instruktion bei der Durchführung als Paper-Pencil-Testung ist die Durchführungsobjektivität anzunehmen, bei der PC- und Internet-Durchführung ist die Durchführungsobjektivität gegeben. Die Auswertung erfolgt bei der Paper-Pencil-Version über das PC-Auswertungsprogramm, bei der Computer- bzw. Internet-Version automatisch, sodass die Auswertungsobjektivität gegeben ist.

Reliabilität

Für das Modul *Verhaltens-Output* wurden zur Analyse der Reliabilitäten Cronbachs Alpha-Werte berechnet. Diese

liegen für die einzelnen Skalen zwischen $\alpha = .74$ und $\alpha = .88$. Wegen des Facetten-Ansatzes wurden die Reliabilitäten für das Modul *Verhaltens-Input* über die Faktorwerte ermittelt. Die Reliabilitäten der Faktorwerte für das Modul *Verhaltens-Input* liegen zwischen .70 und .87. Bei Betrachtung der Reliabilitäten der Faktorwerte für das Modul *Verhaltens-Output* ergeben sich Werte zwischen .90 und .96. Die Reliabilitäten für die Globaldimensionen liegen zwischen .94 und .97. Ferner wurden Retest-Reliabilitäten ermittelt. Die Retest-Reliabilitäten der Skalen des Moduls *Verhaltens-Output* liegen zwischen .77 und .87. Die Retest-Reliabilitäten der Faktorwerte beider Module liegen zwischen .66 und .86. Die niedrigsten Retest-Reliabilitäten ergeben sich für die Primärdimensionen *Belohnungssensitivität* und *Sensibilität für Team*.

Konstruktvalidität

Die im START-P angenommenen Dimensionen wurden im Rahmen von Dimensionsanalysen überprüft. Zur Bewertung der Modellgüte werden die bei Backhaus, Erichson, Plinke und Weiber (2006) beschriebenen Richtwerte für die Fit-Indizes CFI (≥ 0.9) und RMSEA (≤ 0.05) sowie der bei Bühner (2006) erläuterte Richtwert für SRMR ($\leq .11$) herangezogen.

Das Modell zum Modul *Verhaltens-Output* entspricht den Erwartungen der Autoren und weist auf eine adäquate Modellgüte hin [Modellfit: $\chi^2 = 8794.15$; $df = 1065$; $p < .001$; RMSEA = 0.059; SRMR = .076; CFI = .94; $N = 2.063$]. Die teilweise hohen Faktoreninterkorrelationen verdeutlichen, dass die Faktoren zu globaleren Persönlichkeitseigenschaften zusammengefasst werden könnten.

Das Modell zum Modul *Verhaltens-Input* überprüft das Facetten-Konzept. Hier wurden die Items zunächst zu Aggregaten zusammengefasst, die die Facetten Verstärkungsart und Verstärkungsquelle miteinander kombinieren. Das Modell zeigt einen ausreichenden Modellfit [$\chi^2 = 129.11$; $df = 10$; $p < .001$; RMSEA = 0.076; SRMR = .028; CFI = .99; $N = 2.063$].

Das Gesamtmodell beinhaltet zehn Primärfaktoren, die zu drei Sekundärfaktoren zusammengefasst werden. Die Ergebnisse zeigen eine akzeptable Anpassung an das Modell [$\chi^2 = 12057.12$; $df = 1464$; $p < .001$; RMSEA = 0.059; SRMR = .083; CFI = .94; $N = 2.063$]. Die Faktorladungen der Primärfaktoren auf den Sekundärfaktoren entsprechen der erwarteten Struktur.

Ferner überprüften die Autoren die konvergente bzw. diskriminante Validität des START-P. Dazu wurden Korrelationen mit den folgenden Inventaren ermittelt:

- NEO-FFI (Borkenau & Ostendorf, 1993, zitiert nach Beauducel & Kersting, 2010) – Hier zeigen sich u. a. positive Korrelationen zwischen *Kontaktorientierung* (START-P) und *Extraversion* (NEO-FFI) ($r = .58$),

- Empathie* (START-P) und *Verträglichkeit* (NEO-FFI) ($r = .35$), *Sorgfalt* (START-P) und *Gewissenhaftigkeit* (NEO-FFI) ($r = .74$), *Bestrafungssensitivität* (START-P) und *Neurotizismus* (NEO-FFI) (.48) sowie negative Korrelationen zwischen *emotionaler Stabilität* (START-P) und *Neurotizismus* (NEO-FFI) ($r = -.59$). Die Ergebnisse sprechen insgesamt für die konvergente Validität.
- BFI (Rammstedt & John, 2005) – Es zeigen sich u. a. Zusammenhänge zwischen *Kontaktorientierung* (START-P) und *Extraversion* (BFI/BFI-10) ($r = .77 / r = .78$), *Sorgfalt* (START-P) und *Gewissenhaftigkeit* (BFI/BFI-10) ($r = .76 / r = .52$) sowie *Leistungsmotivation* (START-P) und *Extraversion* (BFI-10) ($r = .69$). Diese Ergebnisse können als Hinweise für die konvergente Validität bewertet werden.
 - WIT-2 – Hier zitieren die Autoren die Studie von Busen (2007, zitiert nach Beauducel & Kersting, 2010) und berichten von niedrigen Zusammenhängen, maximal $r = -.14$ für die START-P Skala *Sorgfalt* mit der Skala *sprachliches Denken* des WIT-2. Dies ist erwartungskonform und wird als Hinweis für die diskriminante Validität herangezogen.
 - *Soziale Intelligenz* (Amelang, Schwarz & Wegemund, 1989) und *verbale Intelligenz* (MWT-B; Lehrl, 1976) – Erwartungskonform ergeben die Berechnungen eine Korrelation von .52 zwischen der START-P Skala *Empathie* und der *sozialen Intelligenz*. Die Korrelation mit dem MWT-B beträgt .04.
 - BIS/BAS-Skalen (Strobel, Beauducel, Debener & Brocke, 2001) – Als Indiz für die konvergente Validität wird u. a. die Korrelation zwischen der START-P Skala *Bestrafungssensitivität* mit der *Verhaltenshemmung* (BIS) ($r = .53$) herangezogen. Insgesamt sind die eher geringen Korrelationen der START-P Faktorwerte mit den BIS/BAS-Skalen ein Indiz für die diskriminante Validität der START-P Dimensionen.
- Zusammenfassend kann festgehalten werden, dass die Korrelationen den Erwartungen entsprechen und die konvergente bzw. diskriminante Validität des Verfahrens als gegeben bewertet werden kann.
- ### Kriteriumsvalidität
- Die Kriteriumsvalidität ist anhand der folgenden Außenkriterien ermittelt worden:
- Zufriedenheit in der Schule – Mittels einer schrittweisen Regressionsanalyse ist das Kriterium mit den zehn Grunddimensionen als Prädiktoren vorhergesagt worden. Die multiple Korrelation ist eher gering. Die Analysen zeigen ferner, dass eine höhere *Sorgfalt* eher mit Unzufriedenheit in der Schule einhergeht.
 - Anzahl der Ehrenämter – Die Regressionsanalysen zeigen, dass ein positiver Zusammenhang zwischen *Kontaktorientierung* und der Anzahl an Ehrenämtern besteht. Eine höhere *Leistungsmotivation* hingegen geht mit einer geringeren Zahl an Ehrenämtern einher.
 - Schulnoten – Die Autoren berechneten Korrelationen zwischen den Schulnoten und den zehn Grunddimensionen sowie den drei Globaldimensionen des START-P ($N = 1.719$). Die Berechnungen werden getrennt für Nicht-Gymnasiasten ($N = 1.310$) und Gymnasiasten ($N = 409$) durchgeführt. Die Korrelationen sind überwiegend gering negativ. Die negative Richtung ist logisch positiv, da kleine Schulnoten eine gute Leistung repräsentieren. Die Korrelationen liegen insgesamt im Spektrum von $r = -.23$ (Kunstnote Nicht-Gymnasiasten und *Sorgfalt*) bis zu $r = .15$ (Deutschnote Gymnasiasten und *emotionale Stabilität*).
 - Umfeldbezogener Studienerfolg – Der umfeldbezogene Studienerfolg wurde operationalisiert über die Dimensionen *Hilfsbereitschaft*, *Unkompliziertheit*, *Eigeninitiative* und *Gewissenhaftigkeit* des University Citizenship Behavior (UCB) Fragebogen (Gehring & Kersting, 2007, zitiert nach Beauducel & Kersting, 2010). Die Korrelationen der START-P Skalen *Empathie*, *Sorgfalt* und *Kontaktorientierung* liegen zwischen $r = .02$ (*Kontaktorientierung* und *Unkompliziertheit*) und $r = .69$ (START-P *Sorgfalt* und *Gewissenhaftigkeit*).
 - Aufgabenbezogene Leistung – Die aufgabenbezogene Leistung wird erfasst über die Abschlussnoten im Abitur, der Vordiplomnote sowie der voraussichtlichen Diplomnote. Zu beachten ist die umgekehrte Polung der Notenskala (vgl. oben). Die Korrelationen der START-P Skalen *Empathie*, *Sorgfalt* und *Kontaktorientierung* mit den Einzelkriterien der aufgabenbezogenen Leistung liegen zwischen $r = -.02$ (*Kontaktorientierung* und Abiturnote) und $r = -.27$ (*Sorgfalt* und voraussichtliche Diplomnote).
 - Selbst- und Fremdbeurteilungen – Zur Überprüfung der Selbst-Fremd-Übereinstimmung wurden die START-P Skalen *Kontaktorientierung*, *Sorgfalt* und *Empathie* ausgewählt. Die Fremdbeurteiler sollten die zu beurteilende Person auf den Skalen bewerten (Fremdbeurteilung II). Ferner sollten die Fremdbeurteiler einschätzen, wie sich die Person selbst eingeschätzt hat (Fremdbeurteilung I). Als Zusammenhangsmaß wurde die Intraclass-Korrelation genutzt. Für die Skalen *Kontaktorientierung* und *Sorgfalt* wurden für beide Arten der Fremdeinschätzung mit der Selbsteinschätzung Korrelationen zwischen $r = .54$ und $r = .58$ ermittelt. Die Korrelationen für die Skala *Empathie* fielen geringer aus ($r = .26$ mit Fremdbeurteilung I, $r = .12$ mit Fremdbeurteilung II).
- ### Soziale Validität
- Bei Berücksichtigung der verfahrensspezifischen Empfehlungen der Autoren ist die soziale Validität anzuneh-

men. Die Autoren empfehlen, die Teilnehmerinnen und Teilnehmer zu Beginn der Testdurchführung über die Ziele, den Ablauf, die Dauer sowie den Zweck der Untersuchung zu informieren. Ferner sollen die Teilnehmerinnen und Teilnehmer über den Verbleib der Unterlagen und die Freiwilligkeit der Teilnahme aufgeklärt werden.

Zugänglichkeit des Verfahrens

START-P ist im Hogrefe-Verlag erschienen und kann online über <http://www.testzentrale.de> oder postalisch über Testzentrale, Postfach 3751, 37027 Göttingen bezogen werden. Das Verfahren kostet 158,-EUR. Darin enthalten sind das Manual, zwei Exemplare des Fragebogens Gesamtform A, zwei Exemplare des Fragebogens Gesamtform S, vier Exemplare des Fragebogens zum Modul *Verhaltens-Output* (Fragebogen Modul V), vier Exemplare des Kurz-Fragebogens zum Modul *Verhaltens-Output* (Fragebogen Kurzmodul V), das Auswertungsprogramm auf CD inklusive 20 Auswertungen sowie die Box.

Anwendungsmöglichkeiten

START-P kann bei Einstellungsentscheidungen eingesetzt werden.

Zusammenfassende Bewertung

Mit dem START-P haben die Autoren ein Inventar zur Persönlichkeitsdiagnostik speziell für die Zielgruppe *Berufseinsteiger* entwickelt. Die Testkonstruktion und Validierung erfolgte nach wissenschaftlichen Standards und der DIN 33430.

Das Manual zum START-P gibt nicht nur einen Überblick über den theoretischen Hintergrund, die Testkonstruktion und Gütekriterien des Verfahrens, sondern beinhaltet auch für den Testanwender nützliche Informationen zur Qualitätssicherung in der psychologischen Eignungsdiagnostik. Die Testanwender erhalten zusätzlich Hinweise zur Interpretation der Testergebnisse sowie zur Einbindung dieser in den Auswahlprozess (beispielsweise Informationen zur anforderungsorientierten Integration von Verfahrensergebnissen).

Zur Kriteriumsvalidität legen die Autoren bereits viele Befunde vor. Die meisten Befunde beziehen sich jedoch auf die Skalen des Moduls *Verhaltens-Output*. Ferner liegen derzeit keine Befunde mit *berufsnahen Außenkriterien*, wie beispielsweise IHK-Abschlussnoten oder Bewährung im Betrieb, vor. Wünschenswert wären daher weitere Analysen zur Kriteriumsvalidität.

Wenngleich die Autoren mit dem Verfahren primär die Zielgruppe *Berufseinsteiger* fokussieren, könnten, vor allem unter Berücksichtigung des Einsatzes in der Praxis, die Einsatzmöglichkeiten des START-P stärker expliziert werden. Dies erleichtert es dem Testanwender, zu entscheiden, ob das Verfahren für die eigene Zielsetzung interessant ist.

Zusammenfassend ist START-P ein Persönlichkeitsfragebogen, der sich durch seinen Bezug zum Berufskontext und den Fokus auf junge Menschen von allgemeinen Persönlichkeitsverfahren abgrenzt.

Literatur

- Amelang, M., Schwarz, G. & Wegemund, A. (1989). Soziale Intelligenz als Trait-Konstrukt und Test-Konzept bei der Analyse von Verhaltenshäufigkeiten. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 10, 37–57.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Barrick, M. R. & Mount, M. K. (1991). The Big 5 personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Beauducel, A. & Kersting, M. (2010). *START-P – Testbatterie für Berufseinsteiger – Persönlichkeit*. Göttingen: Hogrefe.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Hossiep, R. & Paschen, M. (2003). *Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP)* (2. Aufl.). Göttingen: Hogrefe.
- Kersting, M. (2005). Arbeits- und Organisationspsychologie. In H. Weber & Th. Rammsayer (Hrsg.), *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie* (S. 535–545). Göttingen: Hogrefe.
- Kirbach, C. & Montel, C. (2005). Die Internettestplattform PERLS. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Aufl., S. 879–892). Lengerich: Pabst Science Publishers.
- Lehrl, S. (1976). *Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)*. Erlangen: Straube.
- 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.
- Strobel, A., Beauducel, A., Debener, S. & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 216–227.

Dr. Meike Wenzel

eligo GmbH
Universitätsstraße 142
44799 Bochum
E-Mail: meike.wenzel@eligo.de