

Frauen unterschätzen sich: Selbstbeurteilungs-Biases in einer simulierten Bewerbungssituation

Women Underevaluate Themselves: Self-Evaluation-Biases in a Simulated Job Interview

Monika Sieverding

Freie Universität Berlin

Zusammenfassung: In einer im Labor simulierten Bewerbungssituation wurden leistungsbezogene Selbstbeurteilungen in Abhängigkeit von Geschlecht und Instrumentalität im Selbstkonzept untersucht. 74 Männer und Frauen absolvierten individuell einen schriftlichen Leistungstest, einen Vortrag zur Selbstdarstellung der beruflichen Qualifikation sowie ein standardisiertes Bewerbungsgespräch. Vortrag und Bewerbungsgespräch wurden videographiert. Frauen schätzten sich in allen Phasen des Versuchs als deutlich weniger erfolgreich ein. Im Vergleich zur erreichten Punktezahl im Leistungstest und zur Fremdbeurteilung eines Videoausschnittes aus dem Bewerbungsgespräch lag bei Frauen eine eindeutige Selbstunterschätzung vor. Bei den Selbstbeurteilungen der Männer zeigte sich im Vergleich zum Leistungstest eine Selbstüberschätzung, im Vergleich zur Fremdbeurteilung eine realistische Selbsteinschätzung. Die Instrumentalität im Selbstkonzept war mit der Selbsteinschätzung als «erfolgreich» assoziiert, allerdings nur bei den männlichen Probanden.

Schlüsselwörter: Bewerbungsgespräch, Selbstbeurteilung, Geschlechtsunterschiede, Instrumentalität

Abstract: Achievement-related self-evaluations were assessed in a simulated job interview situation. Each of the 74 men and women who participated in the laboratory study was asked to complete a written test, a verbal self-presentation of his or her professional qualifications, and a personal interview. The verbal self-presentation and the personal interview were videotaped. Female participants rated themselves as less successful in all phases of the interview situation. Compared to the actual performance in the achievement test and to external assessments of the performance in the personal interview, women showed a significant underevaluation. Males overevaluated themselves compared to the performance in the written test, but not compared to the external assessments of the personal interview. Instrumentality in self-concept was associated with the self-evaluations of success, but this was true only for the male participants.

Keywords: Job interview, self-evaluation, gender differences, instrumentality

Dass Frauen so viel seltener als Männer Karriere machen, ist auf ein komplexes Zusammenwirken sozialer und psychologischer Barrieren zurückzuführen (Abele, 2000; Rappensberger & Spiess, 1997; Sieverding, 1990). Eine Ursache wird darin gesehen, dass es in Beurteilungsprozessen von Frauen zu Vorurteilen und Potenzialfehlschätzungen (insbesondere zu Unterschätzungen) kommt, da Führungsqualifikationen (wie Selbstsicherheit) eher mit dem männlichen als mit dem weiblichen Geschlechtsstereotyp assoziiert werden (Eagly, in press; Friedel-Howe, 1995; Kirchler, Wagner & Buchleitner, 1996; Sczesny, in diesem Heft). Fremdbeurteilungs-Biases und Vorurteile bei der Beurteilung karriererelevanter Qualifikationen von Frauen sind mehrfach empirisch nachgewiesen worden (Eagly, in diesem Heft; Weneras & Wold, 1997; Meta-Analyse von Davison & Burke, 2000).

Die beruflichen Karrierechancen von Frauen werden jedoch nicht nur durch äußere Faktoren wie offene Diskriminierung oder Fremdbeurteilungs-Biases behindert, sondern auch durch psychologische Charakteristika der Frauen selbst, wie z. B. ein typisch feminines Geschlechtsrollen-Selbstkonzept oder eine starke Identifikation mit der Mutter-Kind-Ideologie (wobei solche «inneren» Barrieren natürlich nicht unabhängig von «äußeren» Barrieren entstehen, s. Sieverding, 1990). Kay Deaux, die in mehreren Studien zunächst Fremdbeurteilungs-Biases in der Beurteilung weiblicher Leistung untersucht hatte (z. B. Deaux & Emswiler, 1974), formulierte schon Ende der 70er Jahre die Vermutung, dass (zu) ungünstige Selbstbeurteilungen von Frauen ihre Karrierechancen beeinträchtigen können (Deaux, 1979).

In einer Fülle von Studien wurden Selbstbeurteilungen im Geschlechtervergleich untersucht, mit dem weitgehend konsistenten Ergebnis, dass Frauen in Leistungssituationen ihre Leistungen unterschätzen (Beyer, 1998; Lenney, 1977), nach den Ergebnissen einiger Studien verstärkt in Leistungsbereichen, die als «männlich» stereotypisiert sind (Beyer, 1990; Lenney, 1981). Eine Reihe von Studien analysierte Selbstbeurteilungen bei SchülerInnen und Studierenden. So verglichen Hannover und Bettge (1993) bei SchülerInnen die Schulleistung (in Zensuren) mit der subjektiven Selbsteinschätzung der eigenen Leistungsfähigkeit. Sie fanden insbesondere im «Filterfach» Mathematik, dass Mädchen ihre Leistungen unterschätzten,

während Jungen sie eher überschätzten. Tiedemann und Faber (1995) untersuchten Attributions-tendenzen bei GrundschülerInnen im Geschlechtervergleich. Auch hier zeigte sich, dass bei gleichen Leistungen Mädchen deutlich unzufriedener waren als Jungen und ihren Erfolg stärker external attribuierten (s. auch Rustemeyer & Jubel, 1996). Heatherington und Mitarbeiterinnen (Heatherington, Daubman, Bates, Ahn, Brown & Preston, 1993) ließen Collegestudierende ihr Abschneiden (GPAs = grade point averages) im ersten Semester vorhersagen und verglichen die Selbsteinschätzungen mit den tatsächlich erhaltenen Punkten. Während in den tatsächlich erreichten Punkten keine Geschlechtsunterschiede auftraten, zeigte sich bei den Vorhersagen ein signifikanter Geschlechtereffekt: Frauen sagten sich signifikant niedrigere Punktwerte voraus als Männer, allerdings waren die absoluten Diskrepanzen zwischen den Geschlechtern eher klein (Frauen: $M = 2.94$, Männer: $M = 3.04$), ähnlich wie in einer vergleichbaren Studie von Beyer (1999). Insgesamt zeigen diese Studien Geschlechtsunterschiede in den Selbsteinschätzungen, die entweder auf eine Selbstunterschätzung bei den Frauen, auf eine Selbstüberschätzung bei den Männern (oder beides) zurückzuführen sind. Jedoch sind die Unterschiede in den Selbsteinschätzungen im universitären Setting – wenn sie gefunden werden – absolut gesehen recht gering. Inwiefern ihnen eine verhaltensbezogene Relevanz gerade im Hinblick auf die weitere berufliche Entwicklung zukommt, kann nicht beurteilt werden. Es gibt jedoch eine vielversprechende, kürzlich veröffentlichte Arbeit, die verhaltensrelevante Konsequenzen nahe legt. Ehrlinger und Dunning (2003) untersuchten die Selbsteinschätzungen der wissenschaftlichen Fähigkeit von Collegestudierenden, wobei sie zwischen allgemeinen Selbsteinschätzungen («chronic self-views») der wissenschaftlichen Fähigkeit und der Selbsteinschätzung der konkreten Leistung in einem Test unterschieden. Sie konnten zeigen, dass Frauen eine negativere allgemeine Selbsteinschätzung ihrer wissenschaftlichen Fähigkeit besaßen, und dass sie auch ihre konkrete Leistung in einem Wissenschaftsquiz negativer beurteilten als Männer, obwohl sich in der objektiven Leistung im Test keine Geschlechtsunterschiede zeigten. Frauen zeigten im Anschluss an den Test weniger Interesse, an einem Wissenschaftswettbewerb teilzunehmen, in dem interes-

sante Preise zu gewinnen waren. Dabei hing das Desinteresse der Frauen stärker mit ihren Selbsteinschätzungen der Leistungen im Test als mit ihren tatsächlichen Testleistungen zusammen.

Aufschlussreich sind weiterhin Studien, die Selbsteinschätzungen im beruflichen Setting untersuchten. Deaux (1979) erfasste die Selbstbeurteilungen von weiblichen und männlichen Managern in vergleichbaren Positionen und fand eine Reihe von Geschlechtsunterschieden. Männliche Manager beurteilten nicht nur ihre Gesamtleistung als besser, sie schätzten auch ihre Fähigkeiten und ihre Intelligenz höher ein als weibliche Manager. Männliche Manager attribuierten darüber hinaus ihren Erfolg mit höherer Wahrscheinlichkeit auf ihre Fähigkeiten als weibliche Manager. In der Fremdbeurteilung durch die (männlichen) Supervisoren fanden sich jedoch nur geringfügige Unterschiede; insbesondere wurden weibliche und männliche Manager in Hinsicht auf ihre allgemeinen Leistungen, Fähigkeiten und Motivation nicht als unterschiedlich beurteilt. Auch in einer neueren Studie zeigten sich noch Geschlechtsunterschiede in der Selbstbeurteilung von Managern: Managerinnen attribuierten ihren Erfolg weniger stark auf ihre Fähigkeiten als Manager (Rosenthal, Guest & Peccei, 1996).

Lindeman, Sundvik und Rouhiainen (1995) verglichen die Selbstbeurteilungen von männlichen und weiblichen Bankkaufleuten mit ihrer objektiven Leistung. Als Indikator der objektiven Leistung wurde die Verkaufsleistung in einem Dreimonatszeitraum herangezogen. Die Akkuratheit der Selbsteinschätzung wurde definiert als die Differenz zwischen der (jeweils standardisierten) Verkaufsleistung und Selbsteinschätzung. Sie fanden, dass Männer sich selbst positiver als Frauen beurteilten; zwei Drittel der Männer wurden als «Selbstüberschätzer» diagnostiziert, während Frauen sich gleichmäßig auf die Gruppen der «Selbstunterschätzer», «akkurate Selbsteinschätzer» und «Selbstüberschätzer» verteilten.

Inwieweit in Vorstellungsgesprächen geschlechtsabhängige Selbstbeurteilungs-Biases eine Rolle spielen, ist bisher – meines Wissens – noch nicht empirisch überprüft worden. Dabei stellt ein Vorstellungsgespräch in der Regel eine notwendig zu überwindende «Hürde» auf dem Weg zur gewünschten Stelle oder zur beruflichen Position dar. Der Eindruck, den ein Bewerber oder eine Bewerberin in dieser zeitlich sehr begrenzten Situation macht, gibt häufig den entscheidenden Ausschlag

über Anstellung oder Nicht-Anstellung und damit über wichtige Karriereoptionen. In einem solchen Job-Interview, in dem man in der Regel unmittelbar mit den Mitbewerbern der engeren Wahl verglichen wird, ist es besonders wichtig, die eigenen Stärken zu kennen und entsprechend nach außen darstellen zu können. Assertive Selbstdarstellung, die Fähigkeit, «sich gut verkaufen zu können» (erfasst mit einer Self-Monitoring-Kurzskala) erwies sich in der Studie von Diemand und Schuler (1991) als wesentlicher Prädiktor für das erfolgreiche Abschneiden in einem Assessment-Center. Auch in der Studie von Gallois und MitarbeiterInnen (Gallois, Callan & McKenzie Palmer, 1992) zeigten Personalexperthen eine Vorliebe für selbstbewusste Bewerber. Um jedoch selbstbewusst auftreten zu können, ist eine realistische Einschätzung der eigenen Fähigkeiten notwendig. In der vorliegenden Arbeit geht es darum, potenzielle Selbstbeurteilungs-Biases bei Frauen und Männern in einer Bewerbungssituation zu untersuchen.

Dabei soll kontrolliert werden, ob mögliche Geschlechtsunterschiede in der Selbstbeurteilung auf Unterschiede im Geschlechtsrollen-Selbstkonzept zurückzuführen sind. Seit dem klassischen Artikel von Unger (1979), in der sie die Hypothese aufgestellt hatte, dass viele «sex differences» in Wirklichkeit «gender differences» seien, wurde in einer Vielzahl von Studien untersucht, welche Bedeutung dem Geschlechtsrollen-Selbstkonzept zur Erklärung von Geschlechtsunterschieden beikommt (s. Alfermann, 1996; Sieverding & Alfermann, 1992; zur Entstehung des Geschlechtsrollen-Selbstkonzeptes, s. Hannover, 1997). Dabei hat sich die Instrumentalität im Selbstkonzept als ein aussagekräftiger Prädiktor von beruflicher Karriere(motivation) herausgestellt, während expressive Eigenschaften für das berufliche Fortkommen keine entscheidende Rolle spielen (s. Abele, in diesem Heft, 2000; Abele, André & Schute, 1999; Sieverding, 1990). Entsprechend wird erwartet, dass die Instrumentalität im Selbstkonzept auch die Selbstbeurteilung in einer Bewerbungssituation als «erfolgreich» (positiv) beeinflusst.

Folgende Fragen und Hypothesen sollen überprüft werden:

1. Zeigen sich in einer Bewerbungssituation Geschlechtsunterschiede in der Selbstbeurteilung? Es wird erwartet, dass Frauen sich als weniger erfolgreich beurteilen als Männer.

2. Welcher Zusammenhang besteht zu Indikatoren der «objektiven» Leistung in der Bewerbungssituation? Es wird erwartet, dass weibliche Bewerber bei den objektiven Indikatoren von Leistung nicht schlechter abschneiden als männliche Bewerber.
3. Wie ist das Verhältnis zwischen Selbst- und Fremdbeurteilung? Durch den Vergleich von Selbst- und Fremdeinschätzung der Leistung soll überprüft werden, inwiefern bei Frauen eine Selbstunterschätzung und/oder bei Männern eine Selbstüberschätzung vorliegt.
4. Wie ist die Rolle des Geschlechtsrollen-Selbstkonzeptes für die Selbsteinschätzung der Leistung? Hier wird erwartet, dass die Instrumentalität im Selbstkonzept positiv mit der Selbstbeurteilung der Leistung zusammenhängt.

Methode

Überblick

Im Labor wurde eine realitätsnahe Bewerbungssituation simuliert, in der Männer und Frauen individuell einen schriftlichen Leistungstest, eine Selbstdarstellung der beruflichen Qualifikation sowie ein standardisiertes Bewerbungsgespräch absolvierten. Den Probanden wurde gesagt, dass ihre Leistungen später von Personalexperten beurteilt würden. Der Vortrag zur beruflichen Selbstdarstellung und das Bewerbungsgespräch wurden videographiert.

Stichprobe

Die Stichprobe ($N = 74$) bestand aus 37 Männern und 37 Frauen aus verschiedenen Studienfächern, für die eine berufliche Bewerbung in absehbarer Zeit aktuell war. Die meisten Probanden befanden sich im Hauptstudium, kurz vor Abschluss des Studiums, einige hatten ihr Studium bereits abgeschlossen (4 Männer, 7 Frauen). Die Probanden waren zwischen 22 und 34 Jahre alt; das Durchschnittsalter (arithmetischer Mittelwert) lag bei 26 Jahren ($M = 26.4$, $SD = 2.7$). Sie erhielten für die Teilnahme eine Aufwandsentschädigung von DM 30.–.

Untersuchungsablauf

Die Probanden wurden über Aushänge in verschiedenen Instituten der Freien Universität Berlin geworben. Die Testung wurde nach telefonischem Vorgespräch und Terminverabredung mit jeder Person individuell durchgeführt und dauerte ca. 1 3/4 Stunden. Zu Beginn der jeweiligen Untersuchung wurden die Probanden nach ihrem konkreten beruflichen Ziel gefragt; sie sollten angeben, was für sie eine attraktive berufliche Stelle wäre, auf die sie sich gern bewerben würden. Sie wurden anschließend gebeten, sich so zu verhalten, als würden sie sich auf diese Stelle bewerben. Den Probanden wurde mitgeteilt, dass ihre Leistungen später von externen Personalexperten beurteilt würden. Eine Versuchsleiterin bediente die Geräte. Die Instruktionen erfolgten über ein Tonbandgerät.

Nach dem Ausfüllen verschiedener Fragebögen erfolgten die Aufgaben: Ein schriftlicher Leistungstest unter Zeitdruck (10 Minuten), die innere Vorbereitung einer beruflichen Selbstdarstellung (5 Minuten), ein freier Vortrag zur beruflichen Selbstdarstellung (5 Minuten) sowie ein Bewerbungsgespräch (10 Minuten). Die Probanden füllten nach jeder Versuchsphase einen kurzen Fragebogen aus, in dem sie ihre Befindlichkeit in der vorangegangenen Phase dokumentieren sollten. Nach Beendigung der Aufgabenphasen wurden die Probanden in einer schriftlichen Nachbefragung gebeten, sich zum Erleben der verschiedenen Abschnitte der Untersuchung zu äußern.

Um die Versuchsbedingungen des Bewerbungsgesprächs zu standardisieren und soziale Interaktionseffekte in der sozialen Eindrucksbildung (s. Schuler, 1996) auszuschalten, wurden die Fragen des Bewerbungsgesprächs von einem männlichen Interviewer über Tonkassette eingespielt. Die Probanden hatten jeweils eine Minute Zeit, darauf zu antworten. Es wurden zehn Fragen gestellt, wie sie typisch für Bewerbungsgespräche, insbesondere für Stressinterviews, sind (z. B. «Was qualifiziert Sie eigentlich für diese Stelle?», «Wozu sind Sie sich zu schade?»). Bei dem Vortrag zur beruflichen Selbstdarstellung sowie bei dem Bewerbungsgespräch sollten die Probanden sich vorstellen, dass sie aufgrund einer schriftlichen Bewerbung auf die von ihnen gewünschte Stelle zu einem Vorstellungsgespräch eingeladen wurden und dass sie nun der Bewerbungskommission gegenüber sitzen.

würden. Ihr Vortrag zur beruflichen Selbstdarstellung sowie ihre Antworten im Bewerbungsinterview wurden mit der Videokamera aufgezeichnet.

Untersuchungsmaterial

Selbsteinschätzungsdaten

Die *Instrumentalität im Selbstkonzept* wurde zu Beginn der Untersuchung anhand des Personal Attributes Questionnaire (PAQ) von Spence und Helmreich (1978) in der deutschen Version (Runge, Frey, Gollwitzer, Helmreich & Spence, 1981) erfasst. Die Instrumentalitäts-Skala besteht aus sieben Items (unabhängig, aktiv, fälle leicht Entscheidungen, gebe nicht leicht auf, selbstsicher, mich überlegen fühlend, halte Druck gut stand). Die Probanden sollten auf einer 5-stufigen Skala von 0 («trifft gar nicht zu») bis 4 («trifft voll zu») einschätzen, inwieweit jede einzelne Eigenschaft auf sie normalerweise zutrifft.

Der *Kurzfragebogen zur subjektiven Befindlichkeit* (erweitert nach Höfer, Wallbott & Scherer (1984; s. a. Wallbott & Scherer, 1991) wurde zu Untersuchungsbeginn und Untersuchungsende sowie nach jeder Aufgabenphase ausgefüllt. Er besteht aus 13 Items («aktiv», «gelassen», «erregt», «müde», «entspannt», «gedrückt», «ausgeglichen», «reizbar», «gestresst», «nervös», «ängstlich», «erfolgreich», «körperlich wohl»). Die Probanden sollten für jedes Item auf einer 10-stufigen Skala (von 0 «überhaupt nicht» bis 9 «sehr stark») einschätzen, wie sie sich in der vorangegangenen Untersuchungsphase gefühlt hatten. Hier interessiert ausschließlich die Selbsteinschätzung als «erfolgreich» in den verschiedenen Aufgabenphasen.

Fremdeinschätzungsdaten

Ein Videoausschnitt aus dem Bewerbungsinterview wurde von zwei Raterinnen, einer Diplompsychologin/Betriebswirtin und einer Ärztin, die die Fragestellung und Hypothesen der Untersuchung nicht kannten, beurteilt. Als Verhaltensausschnitt zur Beurteilung wurde die Beantwortung der Frage «Welche Schwächen haben Sie?» ausge-

wählt.¹ Die Raterinnen schätzten zunächst die nonverbale Expressivität des Gesichtes ein², anschließend wurden die Antworten der Probanden nach stimmlichen und inhaltlichen Kriterien ausgewertet. So sollten die Raterinnen unter anderem beurteilen, wie souverän die jeweilige Versuchsperson mit der Frage umgegangen war. Hierzu konnten Einschätzungen von 0 («überhaupt nicht souverän») bis 9 («sehr souverän») abgegeben werden. Als «sehr souverän» wurde eine Antwort eingeschätzt, wenn folgende Bedingungen erfüllt waren: a) Pb nennt eigene beruflich relevante und behebbare Schwäche(n), b) nennt positive Gesichtspunkte der Schwäche, relativiert Schwäche (Schwäche als Stärke), nennt Möglichkeiten, wie er/sie die Schwäche in den Griff bekommen kann, wie die Schwäche zu beheben ist, c) es gelingt Pb, die Frage in der Verfügung stehenden Zeit angemessen zu beantworten. Die Übereinstimmung zwischen den Raterinnen betrug nach einem intensiven Rater-Training bei dieser Einschätzung $r = .95$ (Rangkorrelation).

Abschließend wurde für jede Versuchsperson eine allgemeine, eher intuitive Einschätzung abgegeben («Pb wirkt . . .») und zwar mittels des gleichen Kurzfragebogens, mit dem sich die Probanden nach jeder Aufgabenphase selbst einschätzten (s. o.: Kurzfragebogen zur subjektiven Befindlichkeit). Hier interessiert ebenfalls ausschließlich die Fremdeinschätzung als «erfolgreich». Die Interrater-Übereinstimmung betrug bei dieser Einschätzung $r = .90$ (Rangkorrelation).

Leistungstest

Der schriftliche Leistungstest bestand aus drei Teilen: einem Rechentest, bei dem laut Instruktion die Rechenzeichen transformiert werden mussten (so sollte beispielsweise bei einem Pluszeichen multipliziert und nicht addiert werden), einem Wortauswahl-Test, bei dem aus vier Wörtern jeweils das ausgesucht werden musste, welches nicht zu den anderen passt, sowie dem Würfel-Test aus dem IST-70 (Amthauer, 1970) zur Messung des räumlichen Vorstellungsvermögens. Alle Aufgaben mussten unter starkem Zeitdruck bearbeitet werden. So waren für die 20 Aufgaben des Würfel-

1 Das Training der Raterinnen sowie die Durchführung der Ratings – beide Raterinnen beurteilten alle Probanden – war bereits in der hier beschriebenen Konstellation sehr zeit- und kostenaufwendig, weshalb es nicht möglich war, mehrere Sequenzen beurteilen zu lassen.

2 Dieser Aspekt der Studie wird hier nicht weiter berichtet (s. dazu Sieverding, 2000).

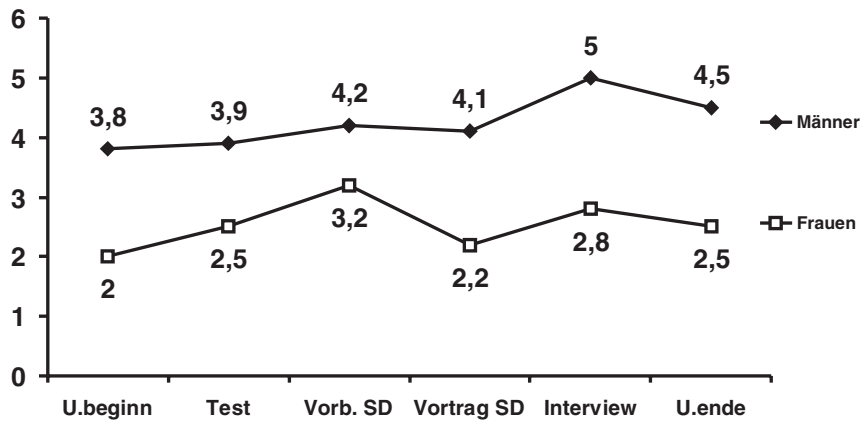


Abbildung 1. Selbsteinschätzung als erfolgreich in den verschiedenen Aufgabenphasen in Abhängigkeit vom Geschlecht; arithmetische Mittelwerte; mögliche Werte von 0 («überhaupt nicht erfolgreich») bis 9 («sehr erfolgreich»); *SD* = Selbstdarstellung.

Tests nur 3 Minuten Zeit, während im Original dafür 9 Minuten Zeit vorgegeben sind.

eine öffentliche (d. h. videographierte) Leistung abverlangt wurde, nämlich im Vortrag zur beruflichen Selbstdarstellung und im Bewerbungsinterview.

Ergebnisse

Selbsteinschätzung als «erfolgreich»

Zur Überprüfung der ersten Hypothese, dass Frauen sich schlechter einschätzen als Männer, wurde die Selbsteinschätzung als «erfolgreich» in einer Varianzanalyse mit dem Faktor (biologisches) Geschlecht (männlich, weiblich) und einem 6stufigen Messwiederholungsfaktor für die experimentelle Phase (Untersuchungsbeginn, Leistungstest, Antizipation des Selbstdarstellungsvortrages, Halten des Selbstdarstellungsvortrages, Bewerbungsinterview und Untersuchungsende) analysiert. Es resultierte ein hoch signifikanter Haupteffekt des Geschlechts; $F(1,72) = 20.66, p < .001$. Frauen schätzten sich in allen Phasen der Untersuchung als deutlich weniger erfolgreich ein als Männer, der Unterschied betrug mindestens 1, höchstens 2.2 Skaleneinheiten auf der 10-Punkte-Skala (s. Abb. 1). Die Varianzaufklärung der Selbsteinschätzung als «erfolgreich» durch den Faktor biologisches Geschlecht ist beachtlich: Die Effektstärke beträgt $\eta^2 = .29$. Bemerkenswert ist, dass der Geschlechtsunterschied schon zu Beginn der Untersuchung auftrat, das heißt, bevor sie überhaupt mit der ersten Aufgabe begonnen hatten, fühlten Männer sich schon deutlich erfolgreicher. Wie in Abbildung 1 ebenfalls zu erkennen ist, ist die Diskrepanz zwischen den Geschlechtern besonders groß in den Aufgabenphasen, in denen

Die Ergebnisse im Leistungstest

Die Mittelwerte der männlichen und weiblichen Probanden in den drei Untertests des schriftlichen Leistungstests wurden mittels t-Tests miteinander verglichen. Hypothesenunterstützend ergaben sich keine signifikanten Geschlechtsunterschiede in diesen objektiven Maßen der Leistung. Bei dem Rechentest hat eine Reihe von Personen die Instruktion überlesen und die Rechenzeichen nicht transformiert, sondern die Aufgaben genauso gerechnet, wie abgebildet (was deutlich einfacher war), deshalb gibt es bei dem ersten Test eine hohe Missing-Rate. Doppelt so viele Männer ($n = 12$) wie Frauen ($n = 6$) machten diesen Fehler, möglicherweise ein Hinweis auf Unterschiede in der Sorgfalt bei der Bearbeitung des Tests. Einige Personen haben den Fehler während der Bearbeitung noch bemerkt und mit den Berechnungen noch einmal von vorn angefangen (weshalb dann die Anzahl der richtig gelösten Aufgaben entsprechend geringer war). Von den Personen, die den Test entsprechend der Instruktion richtig bearbeitet haben ($n = 56$) lösten die Männer ($n = 25$) im Durchschnitt $M = 19.6$ Aufgaben richtig ($SD = 5.8$), die Frauen ($n = 31$) hatten einen Mittelwert von $M = 18.4$ ($SD = 5.6$), $t(54) = .79$, n. s. Wegen der hohen Quote von Probanden, die den Rechentest nicht richtig bearbeitet haben (bei den Männern jeder Dritte), werden die in diesem Test er-

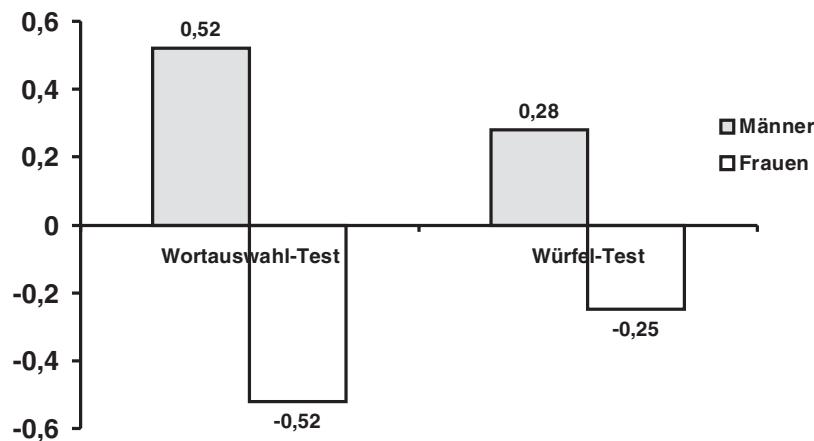


Abbildung 2. Mittlere Diskrepanzen zwischen Selbsteinschätzung als «erfolgreich» im Leistungstest und dem tatsächlichen Abschneiden in den beiden Tests (z-standardisierte Werte) in Abhängigkeit vom Geschlecht; positive Diskrepanzwerte indizieren eine Selbstüberschätzung, negative Diskrepanzwerte eine Selbstunterschätzung.

reichten Scores für die weiteren Analysen nicht berücksichtigt. Bei dem zweiten Test, dem Wortauswahl-Test, schnitten Frauen tendenziell etwas besser ab ($M = 9.4, SD = 2.9$) als Männer ($M = 8.4, SD = 2.3$), $t(72) = -1.65, p = .10$. Bei dem dritten Test, dem Würfel-Test, haben die Männer im Durchschnitt $M = 7.0 (SD = 2.5)$ Aufgaben richtig gelöst, die Frauen $M = 6.6 (SD = 2.3)$, der Unterschied ist nicht signifikant, $t(71) = .76, n. s.$.

Selbsteinschätzung im Vergleich zur objektiven Leistung im Leistungstest

Um den Zusammenhang zwischen der Selbsteinschätzung und der objektiven Leistung zu ermitteln, wurden zunächst Korrelationen zwischen der Selbsteinschätzung als «erfolgreich» und dem Abschneiden in den Leistungstests getrennt für Männer und Frauen berechnet. Bei den Männern waren

Tabelle 1. Korrelationen zwischen den Leistungstests-Scores und der Selbsteinschätzung als «erfolgreich» in Abhängigkeit vom Geschlecht.

| | Wortauswahl-Test | Würfel-Test |
|----------------------------------|------------------|-------------|
| <i>Selbsteinschätzung Männer</i> | | |
| nach dem Leistungstest | .40* | .36* |
| Gesamteinschätzung ^a | .44** | .41* |
| | (n = 37) | (n = 36) |
| <i>Selbsteinschätzung Frauen</i> | | |
| nach dem Leistungstest | .03 | .22 |
| Gesamteinschätzung ^a | .23 | .23 |
| | (n = 37) | (n = 37) |

Anmerkungen. ^aGesamteinschätzung gemittelt über die 6 Untersuchungsphasen; Produkt-Moment-Korrelationen: * $p < .05$, ** $p < .01$.

die objektiven Leistungen im Wortauswahl-Test und Würfel-Test positiv mit der Selbsteinschätzung als «erfolgreich» korreliert, und zwar sowohl mit der Selbsteinschätzung unmittelbar nach den Leistungstests als auch mit der Gesamteinschätzung, die über alle 6 Untersuchungsphasen gemittelt wurde (s. Tab. 1). Bei den Frauen sind die entsprechenden Korrelationen nicht signifikant: Das heißt, die Selbsteinschätzung als (nicht) erfolgreich erfolgte vollkommen unabhängig von ihrer tatsächlichen Leistung in den Leistungstests.

Im nächsten Schritt wurden sowohl die Selbstbeurteilungen als auch die Punktwerte im zweiten und dritten Leistungstest standardisiert (z-Standardisierung mit $M = 0$ und $SD = 1$). Die Differenz zwischen subjektivem Erfolg im Leistungstest und tatsächlichem Abschneiden kann als ein Indikator für Selbstunterschätzung bzw. Selbstüberschätzung angesehen werden. Positive Diskrepanzwerte indizieren eine Überschätzung, negative Diskrepanzwerte eine Unterschätzung. Anhand von t -Tests für abhängige Stichproben wurde getrennt für Männer und Frauen ermittelt, inwiefern die Diskrepanzwerte zwischen (standardisierter) subjektiver Einschätzung und (standardisierten) Testpunktwerten sich signifikant von 0 unterscheiden. Im Wortauswahl-Test zeigt sich eine signifikante Abweichung, und zwar in der weiblichen Teilstichprobe, $t(36) = -2.26, p < .05$, wie auch in der männlichen Teilstichprobe, $t(36) = 3.13, p < .01$. Im Vergleich zur tatsächlichen Leistung im Wortauswahl-Test haben sich die Frauen (im Durchschnitt) unterschätzt, die Männer überschätzt (s. Abb. 2). Die Diskrepanzwerte beim Würfel-Test liegen in der gleichen Richtung, unterscheiden sich jedoch in beiden Sub-Stichproben nicht signi-

fikant von 0 (Frauen: $t(36) = -1.27, p = .21$; Männer: $t(35) = 1.50, p = .14$).

Geschlecht, Instrumentalität und Leistungstest-Scores als Prädiktoren der Selbstbeurteilung als «erfolgreich»

Inwieweit ist das Geschlecht noch ein signifikanter Prädiktor der Selbstbeurteilung als «erfolgreich», wenn die Leistungstest-Scores sowie die Instrumentalität im Selbstkonzept berücksichtigt werden? Zur Beantwortung dieser Frage wurde zunächst ein Gesamtscore der sechs Selbsteinschätzungen als «erfolgreich» gebildet (Summe der Einzeleinschätzungen geteilt durch 6). Anschließend wurde in einer hierarchischen multiplen Regressionsanalyse (Methode: forced entry) die Bedeutung der Prädiktoren «biologisches Geschlecht» (Eingabe im ersten Schritt), Scores des zweiten und dritten Leistungstests (Eingabe im zweiten Schritt), Instrumentalität im Selbstkonzept (Eingabe im dritten Schritt) sowie deren Interaktionen (Eingabe im vierten Schritt) zur Vorhersage der Gesamt-Selbstbeurteilung als «erfolgreich» analysiert. Das biologische Geschlecht, die Leistungstest-Scores und die Instrumentalität im Selbstkonzept erwiesen sich dabei als unabhängige Prädiktoren der Selbsteinschätzung als «erfolgreich» (s. Tab. 2). Männliches Geschlecht, höhere Scores in den Leistungstests sowie eine hohe Instrumentalität sagen eine hohe Selbstbeurteilung als «erfolgreich» voraus. Die These, dass die Instrumentalität im Selbstkonzept die Selbstbeurteilung positiv beeinflusst, wird somit unterstützt. Jedoch zeigt sich, dass das biologische Geschlecht

auch dann noch ein signifikanter Prädiktor bleibt, wenn die Leistungstest-Scores und die Instrumentalität im Selbstkonzept berücksichtigt werden. Wie in Tabelle 2 zu erkennen ist, sind im vierten Schritt der Regressionsanalyse zwei Interaktionen tendenziell signifikant: Die Interaktion zwischen Geschlecht und Instrumentalität sowie die Interaktion zwischen Wortauswahl-Test und Instrumentalität.

Um die Interaktion zwischen Geschlecht und Instrumentalität näher zu untersuchen, wurde zunächst der Zusammenhang zwischen der Instrumentalität im Selbstkonzept und der Selbstbeurteilung getrennt für beide Geschlechter in Korrelationsanalysen ermittelt. Es zeigt sich bei den Männern ein signifikanter Zusammenhang zwischen Instrumentalität und Selbstbeurteilung als «erfolgreich», und zwar sowohl für die Gesamtbeurteilung ($r = .56, p < .001$) als auch für die Einzelbeurteilungen in den verschiedenen Aufgabenphasen (s. Abb. 3). Bei den Frauen besteht dagegen kein signifikanter Zusammenhang zwischen der Instrumentalität im Selbstkonzept und der Gesamtbeurteilung als «erfolgreich» ($r = .14, n. s.$) sowie den Einzelbeurteilungen.

Anschließend wurden die Prädiktoren Leistungstest-Scores und Instrumentalität für die Selbstbeurteilung als «erfolgreich» in getrennten Regressionsanalysen (forced entry) für Männer und Frauen überprüft. Hier zeigen sich besonders große Unterschiede. Bei den Männern lässt sich im ersten Schritt bereits ein bedeutsamer Varianzanteil in der Gesamt-Selbsteinschätzung als «erfolgreich» durch das Abschneiden in den Leistungstests vorhersagen ($R^2 = .29$). Durch die Hinzunahme der Instrumentalität im Selbstkonzept im

Tabelle 2. Hierarchische multiple Regression zur Vorhersage der Gesamt-Selbstbeurteilung^a als «erfolgreich».

| Schritt Prädiktor | R^2_{cum} | $Beta^b$ | t | p |
|------------------------------------|-------------|----------|---------|-------|
| 1. (biol.) Geschlecht ^c | .22 | -.32 | -2.98** | < .01 |
| 2. Leistungstests | .35 | | | |
| Wortauswahl-Test | | .19 | 1.79 | .08 |
| Würfel-Test | | .22 | 2.29* | < .05 |
| 3. Instrumentalität | .43 | .35 | 3.26** | < .01 |
| 4. Interaktionen | .47 | | | |
| Geschlecht × Instru. | | -.23 | -1.91 | .06 |
| WA-Test × Geschl. | | | < 1 | |
| Würfel-Test × Geschl. | | | < 1 | |
| WA-Test × Instru. | | .22 | 1.84 | .07 |
| Würfel-Test × Instru. | | | < 1 | |

Anmerkungen. Methode: enter; ^ader Gesamtscore wurde aus den 6 Einzelbeurteilungen ermittelt. ^bAngegeben ist das Beta-Gewicht, das im letzten Analyseschritt ermittelt wurde. ^cKodierung Geschlecht: 1 = männlich, 2 = weiblich.

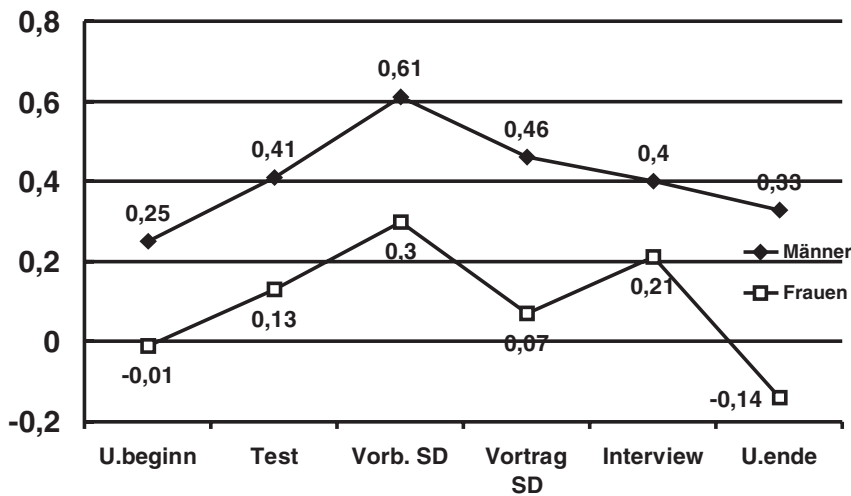


Abbildung 3. Produkt-Moment-Korrelationen zwischen der Instrumentalität im Selbstkonzept und der Selbsteinschätzung als «erfolgreich» in den verschiedenen Untersuchungsphasen. Bei den Männern sind mit Ausnahme der ersten Korrelation alle Korrelationen mindestens auf dem .05 % – Niveau signifikant, bei den Frauen sind alle Korrelationen nicht signifikant.

zweiten Schritt der Regressionsanalyse kann die aufgeklärte Varianz um weitere 18 % auf $R^2 = .47$ erhöht werden. Bei Männern lässt sich somit ihre Selbstbeurteilung als «erfolgreich» recht gut aufgrund ihrer objektiven Leistung in den Leistungstests sowie ihrer Instrumentalität im Selbstkonzept vorhersagen. Bei den Frauen dagegen erweist sich keiner der Prädiktoren als signifikant (s. Tab. 3). Die Selbstbeurteilung der Frauen als «erfolgreich» lässt sich weder durch ihr Abschneiden in den Leistungstests noch durch ihre Instrumentalität im Selbstkonzept vorhersagen.

Vergleich Selbst- und Fremdbeurteilung

Zunächst wurde überprüft, ob die Probanden von den Raterinnen in Abhängigkeit von ihrem biologischen Geschlecht unterschiedlich beurteilt wurden. Mittels t -Tests wurden dazu die Beurteilungen der männlichen und weiblichen Bewerber durch die Raterinnen zu der Frage: «Wie souverän geht Pb mit der Frage um?» und zu der Fremdeinschätzung als «erfolgreich» miteinander verglichen. Bei der Einschätzung, wie souverän die BewerberInnen mit der Frage umgegangen waren, erhielten Männer absolut betrachtet etwas bessere Werte ($M = 5.4$, $SD = 2.1$) als Frauen ($M = 4.9$, $SD = 1.4$), der Unterschied ist jedoch nicht signifikant; $t(72) = 1.2$, $p = .23$. Bei dem allgemeinen Eindruck, den die Raterinnen von den Probanden hatten («Pb wirkt ... erfolgreich»), zeigte sich jedoch ein signifikanter Unterschied in Abhängigkeit vom Geschlecht der Stimulusperson. Frauen

Tabelle 3. Hierarchische multiple Regressionsanalysen zur Vorhersage der Gesamt-Selbstbeurteilung^a als «erfolgreich» getrennt für Männer und Frauen.

| Schritt Prädiktor | R^2_{cum} | $Beta^b$ | t | p |
|---------------------|-------------|----------|------|-------|
| <i>Männer</i> | | | | |
| 1. Leistungstests | .29 | | | |
| Wortauswahl-Test | | .25 | 1.80 | .08 |
| Würfel-Test | | .27 | 1.96 | .06 |
| 2. Instrumentalität | .47 | .44 | 3.23 | < .01 |
| <i>Frauen</i> | | | | |
| 1. Leistungstests | .09 (n. s.) | | | |
| 2. Instrumentalität | .12 (n. s.) | | | |

Anmerkungen. Methode: enter; ^ader Gesamtscore wurde aus den 6 Einzelbeurteilungen ermittelt. ^bAngegeben ist das Beta-Gewicht, das im letzten Analyseschritt ermittelt wurde; ** $p < .01$; * $p < .05$.

wurden als weniger erfolgreich ($M = 3.7$, $SD = 1.5$) eingeschätzt als Männer ($M = 4.7$, $SD = 2.2$); $t(72) = 2.32$, $p < .05$.

Im nächsten Schritt wurde anhand einer hierarchischen Regressionsanalyse die Bedeutung vom biologischen Geschlecht der BewerberInnen, deren Abschneiden in den Leistungstests, deren Instrumentalität im Selbstkonzept sowie deren Selbstbeurteilung als «erfolgreich» für die Fremdeinschätzung analysiert. Dazu wurde im ersten Schritt das biologische Geschlecht eingegeben, im zweiten Schritt die Leistungstest-Scores, im dritten Schritt die Instrumentalität im Selbstkonzept, und im vierten die (Gesamt-) Selbstbeurteilung der BewerberInnen als «erfolgreich». In Tabelle 4 ist das Ergebnis dieser Regressionsanalyse dargestellt. Es zeigt sich, dass das biologische Ge-

schlecht zwar ein signifikanter Prädiktor im ersten Schritt ist und einen Varianzanteil von 7 % der Fremdeinschätzung als «erfolgreich» aufklären kann. Die Leistungstest-Scores sowie die Instrumentalität der BewerberInnen tragen dagegen nicht signifikant zur Varianzaufklärung bei. Der wichtigste Prädiktor ist die Selbsteinschätzung als «erfolgreich». Wenn diese Variable im 4. Schritt eingeführt wird, ist das biologische Geschlecht kein signifikanter Prädiktor mehr. Personen, die sich über die verschiedenen Phasen der Bewerbungssituation als erfolgreicher einschätzten, wurden auch von den Raterinnen als erfolgreicher eingeschätzt.

Um zu überprüfen, ob die Diskrepanz zwischen den Selbsteinschätzungen als «erfolgreich» im Bewerbungsinterview eher auf eine Selbstunterschätzung der Frauen oder auf eine Selbstüberschätzung der Männer zurückzuführen ist, wurde die Differenz aus der Selbst- und Fremdbeurteilung

als «erfolgreich» im Bewerbungsinterview gebildet. (Da die Antworten auf der gleichen Skala gegeben wurden, war hier keine z-Standardisierung notwendig.) Ein positiver Score besagt, dass die Selbstbeurteilung besser war als die Fremdbeurteilung (Selbstüberschätzung), ein negativer Score indiziert dagegen eine Selbstunterschätzung. Die Bewerber hatten eine mittlere Differenz von $M = .26$ ($SD = 2.4$), die Bewerberinnen eine mittlere Differenz von $M = -.93$ ($SD = 2.1$). Mittels t -Tests für abhängige Stichproben wurde dann für die männliche und weibliche Sub-Stichprobe getrennt berechnet, inwiefern sich die Diskrepanzen zwischen Selbst- und Fremdeinschätzung signifikant von 0 unterscheidet. Bei Männern war der entsprechende Test nicht signifikant; $t(36) = .66$, $p = .51$, bei Frauen unterschied sich die Diskrepanz jedoch signifikant von 0; $t(36) = -2.69$, $p < .05$. Bei der Selbstbeurteilung des Abschneidens im Bewerbungsinterview liegt somit (im Durchschnitt) keine Selbstüberschätzung der Männer, wohl aber eine Selbstunterschätzung der Frauen vor (s. dazu auch Abb. 4).

Tabelle 4. Hierarchische multiple Regression zur Vorhersage der Fremdeinschätzung als «erfolgreich» im Bewerbungsinterview.

| Schritt Prädiktor | R^2_{cum} | Beta ^a | t | p |
|---|-------------|-------------------|------|------|
| 1. Geschlecht ^b | .07 | | | |
| 2. Leistungstests | .09 | | | |
| 3. Instrumentalität | .10 | | | |
| 4. Selbstbeurteilung als «erfolgreich» ^c | .19 | .40 | 2.73 | <.01 |

Anmerkungen. Methode: enter; ^aAngegeben ist das Beta-Gewicht, das im letzten Analyseschritt ermittelt wurden. ^bKodierung Geschlecht: 1 = männlich, 2 = weiblich; ^cGesamtbeurteilung, gemittelt über 6 Untersuchungszeitpunkte.

Diskussion

Frauen unterschätzen sich – dieser Befund, der schon in verschiedenen Studien im schulischen und universitären Bereich sowie auch im beruflichen Setting festgestellt wurde (Beyer, 1990, 1998, 1999; Deaux, 1979; Ehrlinger & Dunning, 2003; Hannover & Bettge, 1993; Heatherington et

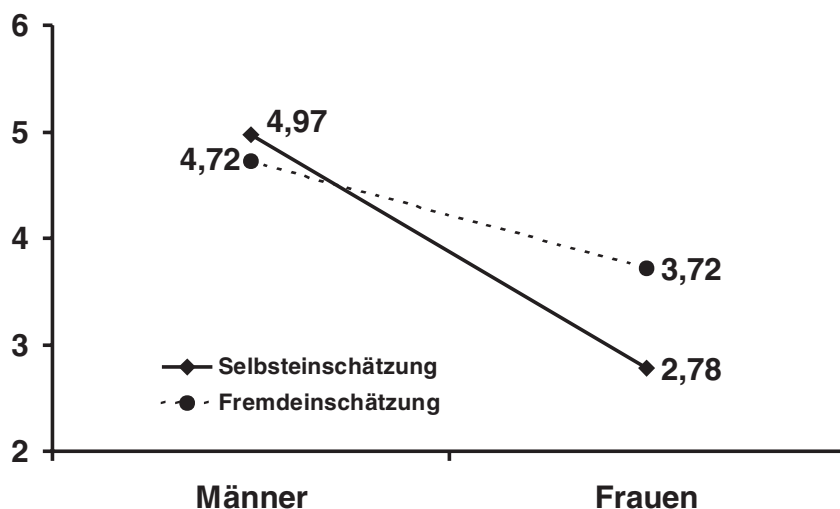


Abbildung 4. Selbsteinschätzung und Fremdeinschätzung (durch zwei Raterinnen) als «erfolgreich» im Bewerbungsinterview; arithmetische Mittelwerte; mögliche Werte von 0 («überhaupt nicht erfolgreich») bis 9 («sehr erfolgreich»).

al., 1993; Lindeman et al., 1995), wird durch die Ergebnisse der hier vorgestellten Untersuchung eindeutig unterstützt. In einer realitätsnahen Bewerbungssituation schätzten weibliche Probanden sich vom Untersuchungsanfang bis zum Untersuchungsende und in allen Aufgabenphasen als deutlich weniger erfolgreich ein als die männlichen Probanden, wobei der Geschlechtsunterschied gravierend war (die Effektstärke für das biologische Geschlecht beträgt $\eta^2 = .29$). Dabei waren die Frauen in objektiven Indikatoren (dem Leistungstest) genauso erfolgreich wie die Männer, auch dies in Einklang mit anderen Studien (s. z. B. Ehrlinger & Dunning, 2003). Bei den weiblichen Probanden wurde eine eindeutige Selbstunterschätzung festgestellt, und zwar nicht nur im Vergleich zur objektiven Leistung im Leistungstest, sondern auch im Vergleich zu der Fremdbeurteilung durch zwei Raterinnen.

In der Fremdeinschätzung wurden die weiblichen Probanden ebenfalls als weniger erfolgreich eingeschätzt als die männlichen, hier waren die Geschlechtsunterschiede jedoch deutlich geringer als in der Selbsteinschätzung. Zudem fand sich der Geschlechtsunterschied in der Fremdeinschätzung nur in dem allgemeinen, eher intuitiven Urteil über die BewerberInnen («Pb wirkt . . . erfolgreich»). Bei der Fremdeinschätzung, wie souverän ein Bewerber bzw. eine Bewerberin mit der Frage «Welche Schwächen haben Sie?» umgegangen war, gab es keinen signifikanten Unterschied in Abhängigkeit vom Geschlecht der beurteilten Person. Zur Einschätzung der Souveränität waren den Raterinnen klare Kriterien vorgegeben, aufgrund derer sie ihr Urteil fällen sollten. Es ist nicht auszuschließen, dass bei dem eher intuitiven Urteil Geschlechtsstereotype die Fremdbeurteilung beeinflusst haben, während die Wirkung von Geschlechtsstereotypen auf die Fremdbeurteilung bei der Anwesenheit von klaren Kriterien minimiert wurden.

Während bei Männern die Selbsteinschätzung der eigenen Leistung mit ihrer objektiven Leistung in den Tests korreliert war (die entsprechenden Korrelationen sind denen in anderen Studien vergleichbar, s. dazu den Überblick bei Ehrlinger & Dunning, 2003), war bei den Frauen die Selbsteinschätzung als «erfolgreich» im Leistungstest überhaupt nicht mit ihren tatsächlichen Leistungen korreliert. Auch die Instrumentalität im Selbstkonzept war nur für die Männer, nicht aber für die Frauen, ein signifikanter Prädiktor für die Selbst-

einschätzung als «erfolgreich». Dieses Ergebnis ist besonders überraschend, da man aufgrund der vorliegenden Befunde zur Bedeutung von Instrumentalität für Karriereambitionen von Frauen (Abele, in diesem Heft, 2000; Abele et al., 1999; Sieverding, 1990) annehmen würde, dass die Instrumentalität gerade für die Selbsteinschätzung von Frauen eine entscheidende Rolle spielen würde.

Wenn es nicht die tatsächliche Leistung ist, und auch nicht eine relevante personelle Variable, nämlich die Instrumentalität im Selbstkonzept, die die Selbsteinschätzung der Frauen beeinflusst haben, was ist es dann? Nach der Theorie und den empirischen Befunden von Ehrlinger und Dunning (2003) beeinflussen allgemeine Selbstkonzepte («chronic self-views») die Selbsteinschätzungen von konkreten Leistungen. Sie konnten in einer Studie zeigen, dass die – mehrere Wochen vor der eigentlichen Untersuchung erfassten – Selbstkonzepte über wissenschaftliche Fähigkeiten die Selbstbeurteilungen des Abschneidens in einem Wissensquiz beeinflussten. Dass Frauen ihr Abschneiden in dem Wissensquiz schlechter einschätzten als Männer, obwohl ihre objektiven Leistungen genauso gut waren, wurde durch ein allgemein negativeres Selbstkonzept über die eigenen wissenschaftlichen Fähigkeiten beeinflusst. In der vorliegenden Studie wurden solche allgemeinen Selbstkonzepte nicht erfasst, dieses müsste optimalerweise einige Wochen vor der eigentlichen Untersuchung geschehen. Es ist möglich, dass Frauen ein negativeres allgemeines Selbstkonzept bezüglich ihrer Fähigkeit zur Selbstpräsentation in einem Job-Interview besitzen, und dass ein solches negativeres Selbstbild die konkreten Selbsteinschätzungen der Bewerberinnen in der Bewerbungssituation beeinflusst hat. Auf jeden Fall war den Frauen die Selbstpräsentation deutlich unangenehmer als den Männern, dies ging aus den schriftlichen Nachbefragungen eindeutig hervor. So beschrieben einige Frauen es als geradezu «unwürdig», ihre Fähigkeiten anpreisen zu müssen. Dass die Instrumentalität und auch die objektiven Leistungen im Test die Selbsteinschätzungen der Frauen überhaupt nicht tangierten, eröffnet noch eine weitere Interpretationsmöglichkeit: Möglicherweise bestehen für Männer und Frauen in einer solchen Bewerbungssituation unterschiedliche Display-Rules (Ekman & Friesen, 1969), Darbietungs- oder Darstellungsregeln, in

denen festgelegt ist, wer wem gegenüber wann welches Gefühl zeigen bzw. nicht zeigen sollte (s. auch Sieverding, 2000). So wäre es denkbar, dass die weiblichen Teilnehmenden dieser Studie glaubten, nicht «zu angeberisch» auftreten zu dürfen und stattdessen mehr Bescheidenheit an den Tag legen zu müssen, um von den Beurteilern nicht zu negativ eingeschätzt zu werden (siehe dazu z. B. Wosinska, Dabul, Whetstone-Dion & Cialdini, 1996). Das Wirken einer solchen geschlechtsabhängigen Darstellungsregel könnte erklären, warum die Instrumentalität, aber auch die objektive Leistung der Frauen keine nachweisbare Rolle für ihren selbsteingeschätzten Erfolg spielte.

Obwohl also innerhalb der Gruppe der Frauen keine psychologische Variable identifiziert werden konnte, die erklären könnte, warum die einen sich als erfolgreicher und die anderen sich als weniger erfolgreich einschätzten, ist der Unterschied zwischen den Geschlechtern auffallend groß. Die Diskrepanzen zwischen den mittleren Selbstbeurteilungen von Männern und Frauen waren in dieser Studie besonders hoch und lagen deutlich über den Diskrepanzen, die beispielsweise aus Studien berichtet wurden, in denen Studierende ihre Studienleistung vorhersagen sollten (Heatherington et al., 1993; Beyer, 1999).

Die Potenzialunterschätzungsthese von Friedel-Howe (1995) geht davon aus, dass die Manifestation von Führungspotenzial einer Frau erwartungswidriges Verhalten verkörpert und deshalb von Beurteilern tendenziell negiert bzw. unterschätzt wird (vgl. auch Eagly & Karau, 2002). Die Ergebnisse der vorliegenden Studie weisen darauf hin, dass Potenzialunterschätzung nicht nur bei der Beurteilung anderer Personen eine Rolle spielen kann, sondern auch bei der Beurteilung der eigenen Person. Wenn Frauen jedoch ihr eigenes Potenzial unterschätzen, ist es unwahrscheinlich, dass sie in einem Vorstellungsgespräch ihr Gegenüber überzeugen können. Tatsächlich wurden ja die Kandidatinnen als signifikant weniger erfolgreich beurteilt als die Kandidaten, und zwar von Beurteilerinnen. Wie Davison und Burke (2000) in ihrer Meta-Analyse aufgezeigt haben, neigen weibliche vergleichbar wie männliche Beurteiler zu einer ungünstigeren Beurteilung von Frauen in simulierten Bewerbungs- und Einstellungsstudien.

Eine weitere Erklärungsmöglichkeit für die bescheidene Selbstbeurteilung könnte sein, dass vie-

le Frauen an sich selbst ein unrealistisch hohes Anspruchsniveau anlegen, vor dem dann ihre Selbstbeurteilung – im Vergleich zu ihren männlichen Konkurrenten – zu schlecht ausfällt. Hennig und Jardim (1987) haben das Phänomen beschrieben, wonach Männer die Welt häufiger so sehen, wie sie ist und Frauen sie eher so sehen, wie sie sein sollte. Viele Frauen in einer Bewerbungssituation scheinen sich auch so zu sehen, wie sie sein sollten, und unterliegen möglicherweise einem stärkeren Druck angesichts dieser hohen Erwartungshaltung. Das negative psychologische Potenzial solcher Selbsteinschätzung ist aus einigen Aufsätzen in der Nachbefragung zu erkennen. Während Männer häufig im Nachhinein die Situation als «Herausforderung» oder als «Chance» beschrieben, bewerteten nicht wenige Frauen die Situation als «Bedrohung» oder «Verlust», ihr eigenes Verhalten als «Versagen».

Es ist zu vermuten, dass die Selbsteinschätzung (und Selbstdarstellung) von Frauen in Bewerbungssituationen besser ist (bzw. verbessert werden kann), wenn sie mehr realistische Rückmeldungen über ihre tatsächlichen Leistungen suchen und erhalten, und zwar, *bevor* sie in ein Vorstellungsgespräch gehen. In Anlehnung an die Studien von Ehrlinger und Dunning (2003) könnte man beispielsweise versuchen, das allgemeine Selbstkonzept von Bewerbern und Bewerberinnen vor der eigentlichen Untersuchung (positiv) zu beeinflussen, z. B. durch eine Bewusstmachung der bisherigen schulischen und beruflichen Erfolge. Möglicherweise ist jedoch auch der Selbstunterschätzungseffekt bei den Frauen in dieser Studie aufgrund der Versuchsbedingungen (Fehlen eines persönlichen Gesprächspartners, kein unmittelbares Feedback über die Leistung) künstlich vergrößert und tritt in realen Bewerbungssituationen in dieser Form nicht auf. Deshalb ist eine Replikation der Studie unter Normalbedingungen dringend erforderlich.

Danksagung

Die vorliegende Studie wurde durch die großzügige Förderung des Instituts für Medizinische Psychologie (Leitung: Prof. Dr. Hans Peter Rosemeier) der Freien Universität Berlin ermöglicht, für die ich mich besonders bedanke. Außerdem danke ich Dr. Bettina von Volkmann für die engagierte

und kompetente Versuchsleitung und Dr. Birgit Herbeck für die Durchführung des Raterinnentrainings.

Literatur

- Abele, A. E. (in diesem Heft). Geschlecht, geschlechtsbezogenes Selbstkonzept und Berufserfolg. Befunde aus einer prospektiven Längsschnittstudie mit Hochschulabsolventinnen und -absolventen. *Zeitschrift für Sozialpsychologie*.
- Abele, A. E. (2000). Gender gaps in early career development of university graduates. Why are women less successful than men? *European Bulletin of Social Psychology*, 12, 22–37.
- Abele, A. E., Andrä, M. S. & Schute, M. (1999). Wer hat nach dem Hochschulexamen schnell eine Stelle? Erste Ergebnisse der Erlanger Längsschnittstudie (BELA-E). *Zeitschrift für Arbeits-, Organisations- und Gesundheitspsychologie*, 43, 95–101.
- Alfermann, D. (1996). *Geschlechterrollen und geschlechtstypisches Verhalten*. Stuttgart: Kohlhammer.
- Amthauer, R. (1970). *IST 70. Intelligenz-Struktur-Test*. Göttingen: Hogrefe.
- Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, 59, 960–970.
- Beyer, S. (1998). Gender differences in self-perception and negative recall biases. *Sex Roles*, 38, 103–133.
- Beyer, S. (1999). Gender differences in the accuracy of grade expectancies and evaluations. *Sex Roles*, 41, 279–296.
- Davison, H. K. & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior*, 56, 225–248.
- Deaux, K. & Emswiller, T. (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of Personality and Social Psychology*, 29, 80–85.
- Deaux, K. (1979). Self-evaluations of male and female managers. *Sex Roles*, 5, 571–580.
- Diemand, A. & Schuler, H. (1991). Sozial erwünschtes Verhalten in eignungsdiagnostischen Situationen. In H. Schuler & U. Funke (Hrsg.), *Eignungsdiagnostik in Forschung und Praxis* (Beiträge zur Organisationspsychologie, Band 10, S. 242–248). Stuttgart: Verlag für Angewandte Psychologie.
- Eagly, A.-H. (in diesem Heft). The rise of female leaders. *Zeitschrift für Sozialpsychologie*.
- Eagly, A.-H. (in press). Few women at the top: How role incongruity produces prejudice and the glass ceiling. In D. van Knippenberg & M. A. Hogg (Eds.), *Identity, leadership, and power*. London: Sage.
- Eagly, A.-H. & Karau, S.-J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109, 573–598.
- Ehrlinger, J. & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology*, 84, 5–17.
- Ekman, P. & Friesen, W. V. (1969). The repertoire of non-verbal behavior: Categories, origins, usage and coding. *Semiotica*, 1, 49–98.
- Friedel-Howe, H. (1995). Probleme der Führungspotentialbeurteilung bei Frauen. In W. Sarges (Hrsg.), *Management-Diagnostik* (2. Aufl., S. 97–101). Göttingen: Hogrefe.
- Gallois, C., Callan, V. J. & McKenzie Palmer, J.-A. (1992). The influence of applicant communication style and interviewer characteristics on hiring decisions. *Journal of Applied Social Psychology*, 22, 1041–1060.
- Hannover, B. (1997). Zur Entwicklung des geschlechtsrollenbezogenen Selbstkonzepts: Der Einfluß «maskuliner» und «femininer» Tätigkeiten auf die Selbstbeschreibung mit instrumentellen und expressiven Personeigenschaften. *Zeitschrift für Sozialpsychologie*, 28, 60–75.
- Hannover, B. & Bettge, S. H. (1993). *Mädchen und Technik*. Göttingen: Hogrefe.
- Heatherington, L., Daubman, K. A., Bates, C., Ahn, A., Brown, H. & Preston, C. (1993). Two investigations of «female modesty» in achievement situations. *Sex Roles*, 29, 739–754.
- Hennig, M. & Jardim, A. (1987). *Frau und Karriere*. Reinbek: Rowohlt.
- Höfer, I., Wallbott, H. G. & Scherer, K. R. (1984). Messung multimodaler Streßindikatoren in Belastungssituationen: Person- und Situationsfaktoren. In H. W. Krohne (Hrsg.), *Angstbewältigung in Leistungssituationen* (S. 94–114). Weinheim: Edition Psychologie.
- Kirchler, E., Wagner, J. & Buchleitner, S. (1996). Der langsame Wechsel in Führungsetagen – Meinungen über Frauen und Männer als Führungspersonen. *Zeitschrift für Sozialpsychologie*, 27, 148–166.
- Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin*, 84, 1–13.
- Lenney, E. (1981). What's fine for the gander isn't always good for the goose: Sex differences in self-confidence as a function of ability area and comparison with others. *Sex Roles*, 7, 905–923.
- Lindeman, M., Sundvik, L. & Rouhiainen, P. (1995). Under- or overestimation of self? Person variables and self-assessment accuracy in work settings. *Journal of Social Behavior and Personality*, 10, 123–134.
- Rappensberger, G. & Spiess, E. (1997). Weibliche Karrieren. In L. v. Rosenstiel, M. Hockel & W. Molt (Hrsg.), *Handbuch für Angewandte Psychologie* (4. Erg. Lfg. 4., S. 1–22). Stuttgart: Poeschel.
- Rosenthal, P., Guest, D. & Peccei, R. (1996). Gender differences in managers' causal explanations for their

- work performance: A study in two organizations. *Journal of Occupational and Organizational Psychology*, 69, 145–151.
- Runge, T. E., Frey, D., Gollwitzer, P. M., Helmreich, R. L. & Spence, J. T. (1981). Masculine (instrumental) and feminine (expressive) traits. A comparison between students in the United States and West Germany. *Journal of Cross-Cultural Psychology*, 12, 142–162.
- Rustemeyer, R. & Jubel, A. (1996). Geschlechtsspezifische Unterschiede im Unterrichtsfach Mathematik hinsichtlich der Fähigkeitseinschätzung, Leistungserwartung, Attribution sowie im Lernaufwand und im Interesse. *Zeitschrift für Pädagogische Psychologie*, 10, 13–25.
- Schuler, H. (1996). *Psychologische Personalauswahl*. Göttingen: Verlag für Angewandte Psychologie.
- Szczesny, S. (in diesem Heft). Führungskompetenz: Selbst- und Fremdwahrnehmung weiblicher und männlicher Führungskräfte. *Zeitschrift für Sozialpsychologie*.
- Sieverding, M. (1990). *Psychologische Barrieren in der beruflichen Entwicklung von Frauen – Das Beispiel der Medizinerinnen*. Stuttgart: Enke.
- Sieverding, M. (2000). «Alle wahren Gefühle verbergen und mit fester Stimme und wohlformulierten Sätzen glänzen!» – Die Bedeutung von Selbstdarstellungsregeln im Bewerbungsgespräch. *Zeitschrift für Arbeits- und Organisationspsychologie*, 44, 152–156.
- Sieverding, M. & Alfermann, D. (1992). Instrumentelles (maskulines) und expressives (feminines) Selbstkonzept: Ihre Bedeutung für die Geschlechtsrollenforschung. *Zeitschrift für Sozialpsychologie*, 23, 6–15.
- Spence, J. T. & Helmreich, R. L. (1978). *Masculinity & femininity. Their psychological dimensions, correlates & antecedents*. Austin: University of Texas Press.
- Tiedemann, J. & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 61–71.
- Unger, R. K. (1979). Toward a redefinition of sex and gender. *American Psychologist*, 34, 1085–1094.
- Wallbott, H. G. & Scherer, K. R. (1991). Stress specificities: Differential effects of coping style, gender, and type of stressor on autonomic arousal, facial expression, and subjective feeling. *Journal of Personality and Social Psychology*, 61, 147–156.
- Wenneras, C. & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387, 341–343.
- Wosinska, W., Dabul, A.-J., Whetstone-Dion, R. & Cialdini, R.-B. (1996). Self-presentational responses to success in the organization: The costs and benefits of modesty. *Basic and Applied Social Psychology*, 18, 229–242.

Monika Sieverding

Freie Universität Berlin
 Institut für Psychologie
 Habelschwerdter Allee 45
 D-14195 Berlin
 Tel. +49 30 838-55094
 Fax +49 30 838-54122
 E-mail mosiever@zedat.fu-berlin.de