Gulba, Astrid & Westhoff, Karl

Strukturierte Auswertung von Personalauswahlinterviews: Prüfung ausgewählter psychometrischer Merkmale

Diagnostische Methoden wie das in der Personalauswahl sehr häufig eingesetzte Tiefeninterview sollen sowohl aus wissenschaftlicher als auch praktischer Sicht bestimmbaren Qualitätsansprüchen genügen, bevor sie als Informationsquelle für Entscheidungen genutzt werden. Die psychometrischen Gütekriterien wie Objektivität, Reliabilität und Validität, aber auch Nebengütekriterien wie Akzeptanz bei den Anwendern stellen dabei Kriterien dar, die auch für das Personalauswahlinterview gelten. Im Zentrum der wissenschaftlichen Untersuchung der Gütekriterien des Interviews standen bisher ihre Bestimmung und Maßnahmen zur Verbesserung im Hinblick auf die Durchführung von Interviews. Mittlerweile konnte hinreichend nachgewiesen werden, dass höhere psychometrische Kennwerte erzielt werden, wenn Interviews strukturiert durchgeführt werden (z.B. Conway, Jako & Goodman, 1995, Pulakos & Schmidt, 1995; Wiesner & Cronshaw, 1988). Ein zweiter wichtiger Aspekt des Interviews – die Auswertung – ist bisher in geringerem Umfang Gegenstand solcher Untersuchungen gewesen. Speziell dazu wird bislang vorgeschlagen, die Antworten der Bewerber auf einzelne Interviewfragen anhand von Ankerantworten zu beurteilen, die im Vorfeld entweder mit Fachleuten (Latham & Saari 1984) oder anhand tatsächlich geführter Bewerberinterviews entwickelt wurden (Janz, 1982; Janz, Hellervik & Gilmore, 1986).

Ergänzend schlagen wir nun eine strukturierte Vorgehensweise vor, bei der unabhängig von der Interviewfrage oder zu identifizierenden Ankerantworten zunächst einmal die Gesamtheit aller Äußerungen eines Bewerbers berücksichtigt wird. Gleichzeitig wird eine deutliche Trennung von Informationserhebung und Datenauswertung realisiert (Schuler, 1992), indem alle zuvor im Interview gemachten Äußerungen ausgewertet werden, ohne dass das konkrete Interview bereits auf Auswertungsaspekte abgestellt wird. Aufbauend auf Westhoff und Kluck (1998) und Jetter (1996) gehen wir wie folgt vor (Gulba & Westhoff, eingereicht): Der Interviewer prüft anhand eines a-priori bestimmten verhaltensorientierten Anforderungsprofils jede Äußerung eines Bewerbers hinsichtlich ihrer Relevanz für die vakante Position. Bei vorliegender Relevanz wird die Äußerung einer der Arbeitsanforderungen zugeordnet und einzeln bewertet (positiv-negativ). Es folgt die Relevanzprüfung für die nächste Aussage und so fort. Nach einer solchen Auswertung des kompletten Tiefeninterviews liegen die relevanten Äußerungen für jede der Arbeitsanforderungen in übersichtlicher Form vor, und die Informationen aus den sonstigen Informationsquellen (Bewerbungsunterlagen, Zeugnisse, standardisierte Testverfahren, Assessment Center usw.) können nun zur Beurteilung des Bewerbers hinzugezogen werden.

Wir führten zwei Untersuchungen durch, die der Prüfung der Objektivität, Reliabilität, Validität und Akzeptanz dieser Vorgehensweise bei der Auswertung von Interviews dienten. In einer ersten Untersuchung widmeten wir uns der Frage, ob eine in der vorgeschlagenen Weise strukturierte Auswertung von Personalauswahlinterviews zu einer angemesseneren Beurteilung von Bewerbern führt als eine unstrukturierte Auswertung (Objektivität, Validität und Akzeptanz des Instruments). Die zweite Untersuchung hatte zum Ziel, insbesondere die Reliabilität der strukturierten Vorgehensweise zu bestimmen. Gleichzeitig konnte eine Replikation der in der ersten Untersuchung gefundenen Validitäts- und Objektivitätswerte erfolgen.

Methode

Entwicklung des Materials

Zuerst entwickelten wir ein Anforderungsprofil für eine Position als Außendienstmitarbeiter im Finanzdienstleistungsbereich (vgl. Wottawa, 1997), da in unseren Untersuchungen eine solche Position zu besetzen sein sollte. Um den Teilnehmern die strukturierte Auswertung in einer angemessenen Zeit zu ermöglichen, konzentrierten wir uns auf folgende relevante Anforderungen: Berufser-

fahrung/Praxiswissen, Konzentration, Arbeitsstil, Erwartungen an die neue Tätigkeit, berufliche Ziele, körperliche Belastbarkeit, soziale Unterstützung und Kritikfähigkeit. Wir konstruierten anhand von konkret vorliegendem Material drei Transkripte von Personalauswahlinterviews, in denen ein Bewerber, der bislang im Vertrieb anderer Produkte (Metallwaren, Tupperware, Autos) tätig war, im Rahmen des Bewerbungsverfahrens für die „ausgeschriebene" Position befragt wird. Um die Zuordnung der Informationen zu den Arbeitsanforderungen nicht zu suggerieren, enthielten die Interviews keine Fragen, aus denen die interessierende Arbeitsanforderung hervorgeht. Stattdessen fragte der Interviewer z.B. danach, wo der Bewerber im Moment arbeitet, wie er zu dieser Tätigkeit gekommen ist und wie ein typischer Tagesablauf aussieht. Da uns der Umgang der Untersuchungsteilnehmer mit einzelnen Informationen aus den Interviews interessierte, haben wir die Transkripte in einzelne Informationseinheiten unterteilt. Dabei handelte es sich jeweils um eine Beschreibung, Interpretation oder Bewertung verschiedenster, meist beruflicher Aspekte, die den Bewerber oder dessen soziales Umfeld betreffen. Um prüfen zu können, wie zutreffend die Untersuchungsteilnehmer relevante Informationen aus den Personalauswahlinterviews herausfiltern, enthielten die Transkripte nicht zu allen zu beurteilenden Arbeitsanforderungen auch Informationen. Um die Validität der strukturierten Auswertung prüfen zu können, entwickelten wir eine Musterlösung.

Teilnehmer

An der Untersuchung zum Vergleich der strukturierten versus nicht strukturierten Auswertung nahmen 43 Personalverantwortliche aus dem Finanzdienstleistungsbereich teil (13 Frauen, 28 Männer, 2 ohne Angabe, Alter 26-57 Jahre), die jeweils zwei der Transkripte entweder strukturiert (N=17) oder nicht strukturiert (N=26) auswerteten. An der zweiten Untersuchung nahmen 37 Studierende der Psychologie der Technischen Universität Dresden teil (30 Frauen, 7 Männer, Alter 21-42 Jahre), die durch ihre Diagnostikausbildung in der strukturierten Auswertung von Gesprächen geübt waren. Die Studierenden werteten jeweils eines der Transkripte zweimal im zeitlichen Abstand von zwei Wochen strukturiert aus. Erfahrungen im Umgang mit dem Anforderungsprofil für die zu besetzenden Stelle hatten diese Studierenden ebenso wie die Personalverantwortlichen nicht. In beiden Untersuchungen teilten wir die Interviewtranskripte zufällig zu.

Ergebnisse

Entsprechend des gemeinsamen Ziels beider Untersuchungen (die Bestimmung psychometrischer Gütekriterien der strukturierten Auswertung) werden im folgenden die Ergebnisse zu einzelnen Gütekriterien zusammengefasst, wobei auf die entsprechenden Resultate beider Untersuchungen Bezug genommen wird.

Objektivität: In beiden Untersuchungen wurde zunächst analysiert, wie gut die Teilnehmer, die mit dem Auswertungsinstrument arbeiteten, erkennen, dass für ausgewählte Arbeitsanforderungen nicht genügend Informationen vorliegen. Die Personalverantwortlichen in Untersuchung 1, die die Transkripte strukturiert auswerteten, erkannten bei beiden Transkripten für drei von vier Arbeitsanforderungen häufiger, dass für diese Arbeitsanforderungen keine Informationen vorliegen, als diejenigen ohne strukturiertes Vorgehen (in drei Fällen wird dieser Unterschied signifikant mit $p<.05$).

Zwischen 67% - 95% der Studierenden in Untersuchung 2 erkannten für das Transkript mit dem Metallwarenverkäufer korrekt, dass für die Arbeitsanforderungen keine relevanten Informationen enthalten waren. Für das Transkript mit dem Autoverkäufer erkannten dies zwischen 57% - 100% der Studierenden korrekt (Angaben bezogen auf beide Messzeitpunkte). Eine Ausnahme bildet die Arbeitsanforderung "Kostenbewusstsein", für die im letztgenannten Fall nur 31% bzw. 50 % der Studierenden das Fehlen der Informationen korrekt erkannten. Eine Analyse der hier fälschlicherweise zugeordneten Informationen zeigt, dass die Teilnehmer entgegen der Beschreibung im Anforderungsprofil solche Aussagen des Bewerbers zu "Kostenbewusstsein" zugeordnet haben, in denen der Bewerber von Geld spricht – obwohl nichts darüber ausgesagt wird, wie mit Geld umgegangen wird.

Die Personalverantwortlichen beurteilen die Bewerber hinsichtlich der verschiedenen Arbeitsanforderungen übereinstimmender, wenn sie die Transkripte strukturiert auswerten. Kendalls W liegt für

Transkript 1 (Metallwarenverkäufer) bei .41** (strukturiert) bzw. .22** (unstrukturiert); für Transkript 2 (Tupperwareberaterin) ist W= .41** (strukturiert) bzw. .29** (unstrukturiert). In der zweiten Untersuchung, in der ausschließlich strukturiert ausgewertet wurde, liegt die Beurteilungsübereinstimmung für Transkript 1 (Metallwarenverkäufer) bei .55** (Messzeitpunkt 1) bzw. .60** (Messzeitpunkt 2), für Transkript 3 (Autoverkäufer) bei .66** (Messzeitpunkt 1) bzw. .53** (Messzeitpunkt 2). Alle Übereinstimmungsmaße sind signifikant, $p<.01$.

Validität: Tabelle 1 gibt den Prozentsatz der korrekten Zuordnungen im Vergleich zur der verwendeten Musterlösung für beide Untersuchungen wieder. Die Werte liegen zwischen 33% und 49%.

Tabelle 1: Prozentsatz der korrekten Zuordnungen.

	Personalverantwortliche (Untersuchung 1)	Studierende (Untersuchung 2)	
		Messzeitpunkt 1	Messzeitpunkt 2
Transkript 1 (Metallwaren) 106 Informationseinheiten	43%	47%	49%
Transkript 2 (Autos) 107 Informationseinheiten		36%	40%
Transkript 3 (Tupperware) 137 Informationseinheiten	33%		

Um zu prüfen, ob die Beurteilungen der Bewerber hinsichtlich der Arbeitsanforderungen bei den Personalverantwortlichen mit strukturierter Auswertung korrekter sind als bei denen ohne strukturiertes Vorgehen, haben wir die mittleren Abweichungen in der Beurteilung von der Musterlösung berechnet. Die Überlegenheit der strukturierten Auswertung zeigt sich nur für Transkript 1 (Metallwaren) mit $p<.05$.; für Transkript 2 (Tupperware) lassen sich keine Unterschiede in der Beurteilung zwischen den Gruppen feststellen.

Reliabilität: Die Bestimmung der Reliabilität erfolgte in Untersuchung 2 (Studierende). Korreliert man die Zahl der korrekten Zuordnungen zu beiden Messzeitpunkten, so ergibt sich für Transkript 1 (Metallwarenverkäufer) $r=.79$** ($p<0.01$). Für Transkript 2 (Autoverkäufer) liegt der entsprechende Koeffizient bei $r=.65$** ($p<0.01$). Der Anteil der korrekt zugeordneten Informationseinheiten bleibt über die Zeit stabil. Darüber hinaus bleibt die abschließende Entscheidung, den Bewerber aufgrund der vorliegenden Informationen einzustellen oder nicht, über beide Messzeitpunkte stabil: Phi=1.0** (Metallwarenverkäufer) und Phi = .68** (Autoverkäufer), $p<0,01$. Drei Untersuchungsteilnehmer stellen den Bewerber zu keinem der beiden Zeitpunkte ein, eine Person ändert ihre Entscheidung und stellt ihn zum zweiten Messzeitpunkt nicht mehr ein. Die anderen befürworten eine Einstellung zu beiden Messzeitpunkten.

Akzeptanz der strukturierten Auswertung: 64% (Metallwaren) bzw. 58% (Tupperware) der Personalverantwortlichen gaben an, dass die strukturierte Auswertung eine Hilfe bei der Bearbeitung der Transkripte darstellte. Als Hilfe bei der Beurteilung der Bewerber nahmen 79% (Metallwaren) bzw. 69% (Tupperware) der Personalverantwortlichen die strukturierte Vorgehensweise bei der Auswertung wahr.

Diskussion

Die Überprüfung der psychometrischen Gütekriterien einer strukturierten Auswertung zu einem Tiefeninterview in der Personalauswahl stellte ein erstes Ziel unserer Studien dar. Hier zeigte sich, dass die Untersuchungsteilnehmer sehr übereinstimmend und korrekt erkennen, dass für einzelne Arbeitsanforderungen keine Informationen vorliegen. Die Beurteilung der Bewerber erfolgt recht übereinstimmend und tendenziell auch zutreffender als ohne strukturierte Auswertung. Auch die

dargestellten Aspekte der Reliabilität sind zufriedenstellend: Wer beim ersten Messzeitpunkt viele Informationen korrekt zuordnet und sich für eine Einstellung des Bewerbers entscheidet, tut dies auch zum zweiten Zeitpunkt.

Da das Anforderungsprofil verhaltensorientiert formuliert war, vermuteten wir, dass die Zuordnung der Informationen bei strukturiertem Vorgehen leichter fallen müsste. Durch die Betrachtung der einzelnen Schritte im Auswertungsprozess konnten wir feststellen, dass dies in beiden Untersuchungen nicht der Fall war. Möglicherweise war es den Teilnehmern nicht leicht möglich, die Definitionen der Arbeitsanforderungen für die Zuordnung zu nutzen. Beispielsweise hat die Erwähnung des Schlüsselwortes „Geld“ bei den Untersuchungsteilnehmern ausgereicht, um eine Zuordnung zu Kostenbewusstsein vorzunehmen, obwohl in der Definition klar beschrieben war, dass hierfür Informationen über den Umgang mit Geld relevant waren. In weiteren Untersuchungen wäre zu untersuchen, welchen Einfluss die Benennung der Arbeitsanforderungen auf diesen Teilaspekt der strukturierten Auswertung hat, und durch welche (Trainings-) Maßnahmen die Zuordnung von Informationen erleichtert werden kann.

Insgesamt ergänzt die von uns vorgeschlagene strukturierte Auswertung die existierenden Strukturierungsvorschläge des Personalauswahlinterviews (Conway, Jako & Goodman, 1995, Pulakos & Schmidt, 1995). Eine Fokussierung auf sämtliche Aussagen eines Bewerbers unabhängig von den gestellten Fragen (Latham & Saari 1984; Janz, 1982) stellt auch eine geeignete Vorgehensweise dar, um die Qualität der Beurteilung von Bewerbern zu verbessern. Eine klare Trennung von Datenerhebung, Strukturierung der Informationen und anschließende Interpretation erweist sich erneut als vorteilhaft (Schuler, 1992). Für die Anwendung in der Praxis ist die Nützlichkeit einer strukturierten Auswertung – über die psychometrische Güte hinaus – entscheidend. Die gute Akzeptanz der Vorgehensweise durch die Personalverantwortlichen sowie deren positive Beurteilung in der mündlichen Nachbefragung macht deutlich, dass des weiteren die Entwicklung eines praxistauglichen Instruments wünschenswert ist, das die strukturierte Auswertung von Tiefeninterviews in der Personalauswahl in der Praxis ermöglicht.

Literatur

Conway, J. M., Jako, R. A. & Goodman, D. F. (1995). A Meta-Analysis of Interrater and Internal Consistency Reliability of Selection Interviews. *Journal of Applied Psychology, 80 (5),* 565-579.

Gulba, A. & Westhoff, K. (eingereicht im: *International Journal of Selection and Assessment*). Structured Analysis of Personnel Selection Interviews: Psychometric Properties.

Janz, T. (1982). Initial comparisons of patterned behavior description interview versus unstructured interviews. *Journal of Applied Psychology, 67,* 577-580.

Janz, T., Hellervik, L. & Gilmore, D. C. (1986). *Behavior description interviewing.* Boston: Allyn & Bacon.

Jetter, W. (1996). *Effiziente Personalauswahl. Durch strukturierte Einstellungsgespräche die richtigen Mitarbeiter finden.* Stuttgart: Schaeffer-Poeschel.

Latham, G. P. & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. *Journal of Applied Psychology, 69,* 569-573.

Pulakos, E. & Schmitt, N. (1995). Experience-Based and Situational Interview Questions: Studies of Validity. *Personnel Psychology, 48,* 289-308.

Schuler, H. (1992). Das Multimodale Einstellungsinterview. *Diagnostica, 38,* 281-300.

Westhoff, K. & Kluck, M. L. (1998). *Psychologische Gutachten schreiben und beurteilen.* Berlin: Springer.

Wiesner, W. H. & Cronshaw, S. F. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61,* 275-290.

Wottawa, H. (1997). *Anwendungsfelder psychologischer Diagnostik.* Göttingen: Hogrefe.

Erscheint im Frühjahr 2003 im Processing-Band des 9. Dresdner Symposiums "Psychologie der Arbeit" (27.02.-01.03.2003) im Rahmen des 175-jährigen Jubiläums der Technischen Universität Dresden.