

Aus dem Institut für Psychologie der Christian-Albrechts-Universität zu Kiel.
Direktoren: Prof. Dr. H. Wegener und Prof. Dr. D. Wendt.

Ein Beitrag zur Messung des Spannungsbogens von jugendlichen Delinquenten

Burkhard Hasenpusch und Wilfried Hommers

Ziel der Untersuchung war, erste Informationen über die Verwendungsfähigkeit einer im amerikanischen Raum entwickelten Auswahlliste zur Messung des Spannungsbogens zu gewinnen. Dazu wurden mit 234 Jugendstrafgefangenen verschiedenen Alters die Verfahren Auswahlliste und Verhaltensprobe zur Erhebung des Spannungsbogens durchgeführt. Außerdem wurden Daten zum hypothetischen Sparverhalten der Vpn erhoben. Durch klassische und probabilistische Itemanalyse ergaben sich verschiedene Skalen aus der Auswahlliste, die alle in keiner Beziehung zum Alter und dem hypothetischen Sparverhalten standen. Verschiedene Nebenbefunde werden im Hinblick auf ihre Konsequenzen für die Verwendungsfähigkeit der Auswahlliste und auf Erklärungsmöglichkeiten diskutiert.

1. Einleitung und Problemstellung

Unter Spannungsbogen wird die Bereitschaft verstanden, die Befriedigung eines Bedürfnisses zugunsten späterer, aber dafür höher bewerteter Befriedigung des gleichen oder eines anderen Bedürfnisses aufzuschieben. Die Bedeutung des Spannungsbogens für die Sozialisation des Menschen wird allgemein hervorgehoben.

Aufgrund verschiedener theoretischer Ansätze über den Spannungsbogen wurden unterschiedliche Methoden zur Messung seiner Ausprägung vorgeschlagen. Zwei Gruppen lassen sich u. E. unterscheiden:

a) Messung motorischer und kognitiver Impulskontrolle

Diese Methoden gründen sich auf die von Freud (1942, II/III S. 604 ff.) angenommene Impuls-Unterdrückung während des Aufschubens von Bedürfnisbefriedigung. Ihre Brauchbarkeit ist von der Gültigkeit

tigkeit der psychoanalytischen Annahmen abhängig. Die Höhe der Impulskontrolle wurde in diesem Zusammenhang z. B. bestimmt durch die Häufigkeit von Bewegungsantworten in Formdeuteversuchen, die Güte der Leistung beim langsamen Nachziehen von Linien und die Länge von bewegungslos durchgehaltenen Wartezeiten.

b) Ermittlung des Bedürfnisaufschubs in konkreten und hypothetischen Situationen

Hier wird das Verhalten in Situationen, die der Definition des Spannungsbogens unmittelbar entsprechen, zum Gegenstand der Beobachtung oder Beurteilung erhoben. Als Verfahren dienen:

- Verhaltensproben, in denen eine Entscheidung über einen vom Versuchsleiter künstlich ermöglichten Bedürfnisaufschub verlangt wird. Die Versuchsperson wird — oft als Belohnung für ihre vorherige Mitarbeit — vor die Wahl zwischen einer geringeren, sofort erhältlichen Belohnung und einer wertvolleren, aber erst später erhältlichen Belohnung gestellt;
- Auswahllisten, die die Entscheidungssituation der Verhaltensprobe wiederholt enthalten, wobei die Vpn aber darüber informiert wurde, daß die Belohnung nur in einem Falle tatsächlich zur Auszahlung kommt;
- Messung der für eine in Aussicht stehende Belohnung freiwillig durchgehaltenen Wartezeit;
- Beurteilung des Verhaltens von Pbn, wenn dieses geeignet erscheint, eine valide Schätzung des Spannungsbogens des Pbn zu fundieren;
- Fragebögen, in denen hypothetische Entscheidungen, ähnlich wie bei der Verhaltensprobe, verlangt werden oder Einstellungen zum Aufschub von Bedürfnisbefriedigung erhoben werden.

Die Motivation zu der vorliegenden Untersuchung bestand darin, zu prüfen, ob die Auswahlliste und die Verhaltensprobe geeignete Methoden sind, den Spannungsbogen von jugendlichen (Alter 14 bis 17 Jahre), heranwachsenden (Alter 18 bis 20 Jahre) und erwachsenen (Alter über 21 Jahre) Jugendstrafgefangenen zu erfassen. Dabei wird erster Aufschluß darüber erwartet, ob diese Verfahren bei der Begutachtung nach § 105 des JGG sinnvoll einsetzbar sind. Dies wäre dann der Fall, wenn entweder bei den genannten Altersgruppen oder zwischen irgendwelchen Altersstufen deutliche, sprunghafte Veränderungen der Häufigkeit aufgeschobener Wahlen auftreten würden. Solche sprunghaften Veränderungen könnten u. U. dann zur Diagnose verfrühter, verspäteter oder normaler Jugendlichkeit herangezogen werden. Es war in dieser Hinsicht also not-

wendig, die Beziehungen zwischen diesen Maßen für den Spannungsbogen und der Variablen Alter zu untersuchen. Wichtig erschien weiterhin, auch die testtheoretischen Eigenschaften der Meßinstrumente zu bestimmen und stichhaltige Validitätskriterien dieser Maße zu finden. Insbesondere sollte bei den testtheoretischen Fragen geprüft werden, ob, vom Standpunkt des 2-kategorialen Meßmodells von R a s c h aus gesehen, die Items der Auswahlliste zu einem Meßwert aufsummiert werden können.

Als ein Kriterium der Validität dieser Maße kann das Sparverhalten von Vpn angesehen werden. In verschiedenen amerikanischen Untersuchungen wurde, motiviert durch die Schwierigkeit der Erhebung des Sparverhaltens, die Beziehung zwischen Auswahllistenverhalten und hypothetischem Sparverhalten untersucht. M i s c h e l (1961), M i s c h e l (1962), S h y b u t (1968) und S t u m p h a u z e r (1972) fanden positive Beziehungen. In der Untersuchung S t u m p h a u z e r s wurde die Auswahlliste zur Messung der Effektivität eines Trainings zur Erhöhung des Spannungsbogens verwandt. Da die Untersuchung von S t u m p h a u z e r an einer vergleichbaren Stichprobe amerikanischer heranwachsender Delinquenten durchgeführt wurde, erschien es sinnvoll, zu Vergleichszwecken auch die Beziehung zwischen hypothetischem Sparverhalten und den Maßen des Spannungsbogens bei deutschen Delinquenten zu prüfen.

2. Durchführung der Untersuchung

Die von S t u m p h a u z e r (1972) entwickelten Auswahllisten schienen für diese Untersuchung geeignet, da sie an einer Stichprobe von heranwachsenden Strafgefangenen entwickelt worden waren. Die Items dieser Auswahllisten beinhalten allerdings ausschließlich Entscheidungen zwischen verschiedenen großen gleichartigen Verstärkern, so daß mit ihnen nur die Bereitschaft erfaßt werden kann, die sofortige Befriedigung eines Bedürfnisses zugunsten späterer Befriedigung desselben Bedürfnisses zurückzustellen. Verfahren, die eine Wahl zwischen Befriedigungen verschiedenartiger Bedürfnisse erlauben, sind uns nicht bekannt.

Die vier Auswahllisten S t u m p h a u z e r s erhielten 56 Items, von denen 28 kleinere Geldbeträge zum Entscheidungsgegenstand hatten (Geld-Items) und die restlichen 28 Items verschiedene weitere materielle Verstärker wie Zigaretten, Süßigkeiten oder Lesestoff zur Auswahl stellten (Sach-Items). Bei der Übersetzung der Items wurden die Dollarbeträge im Verhältnis 1:2 in DM-Beträge umgewandelt. Einige Sach-Items mußten verändert werden, da sie in Deutschland nicht bekannt oder ihr Besitz in der Anstalt verboten waren. In der Tabelle 1 werden vier Beispiele aufgeführt.

Tabelle 1
Vier Beispiel-Items

Item-Nr.	Bezeichnung	Text des Beispiels	Anzahl
15	Geld-Item	50 Pf heute oder 80 Pf in 2 Wochen	28
19	Sach-Item mit klarer Mengenangabe	3 Schachteln Zigaretten heute oder 5 Schachteln Zigaretten in 3 Wochen	13
40	Sach-Item mit unklarer Mengenangabe	1 große Dose Kabe in 3 Wochen, 1 kleine Dose heute	8
50	Sach-Item mit Zeitschriftenwahl	1 Jerry Cotton heute oder 2 Jerry Cotton in 4 Wochen	7

Die Testanweisung wurde in Anlehnung an Stumphauer (1972, S. 13) formuliert. Sie enthielt eine kurze Darstellung der Entscheidungssituation und ein Probe-Item sowie die Anleitung, wie irrtümliche Wahlen korrigiert werden konnten. In ihr wurde außerdem darauf hingewiesen, daß die Ergebnisse der Untersuchung nicht in die Personalakte kämen und daß es wichtig sei, immer die tatsächlich bevorzugte Alternative anzugeben, da die Belohnung an eines der Items gebunden sei. Schließlich wurde die Versuchsperson gebeten, ihre Entscheidungen möglichst spontan zu treffen. Nachdem die 56 Items der Auswahlliste von den Vpn bearbeitet worden waren, sollten diese angeben, wieviel Prozent (in den Abstufungen 100%, 75%, 50%, 25%, 0%) von 0,50 DM, 4,— DM und 40,— DM sie sparen würden, wenn sie diese Geldbeträge zur Verfügung hätten.

Die Stichprobe bestand aus 234 Insassen der Jugendstrafanstalt Herford, das Alter der Versuchspersonen lag zwischen 15 und 25 Jahren (Mittelwert = 19.5, Standardabweichung = 1.7).

Die Vorgabe der Auswahlliste erfolgte in Gruppen von 20 bis 25 Vpn. Die in der Testanweisung versprochene Belohnung für die Mitarbeit wurde zu einer neuen, nun aber endgültigen Verhaltensprobe ausgebaut. Die Gefangenen entschieden sich gegenüber den Alternativen des Items 15 noch einmal und erhielten dann je nach ihrer Wahl die Geldbeträge in Form von Briefmarken¹⁾ bei Abgabe des Fragebogens oder im Laufe der folgenden Woche ausgehändigt.

1) Geldbeträge konnten innerhalb der Anstalt nicht ausgezahlt werden.

3. Ergebnisse

3.1. Maß- und Test-Eigenschaften der Auswahlliste

Die Rohwerte (RW) der 234 Vpn wurden durch die Häufigkeit der Wahl der aufgeschobenen Alternative bestimmt. Die Verteilung der RW wich von der Gleichverteilung in Richtung auf eine U-Verteilung ab ($\chi^2 = 41.3$ bei $df = 6$). Angestrebt war eine Homogenitätsprüfung der 56 Items nach dem zweikategorialen Modell von Rasch. Wegen Überschreitung der Speicherkapazität des elektronischen Rechners²⁾ wurde eine klassische Itemanalyse zur Reduktion der Itemzahl durchgeführt. Wir wählten für die Bestimmung der Itemkennwerte die Vpn aus, die innerhalb des ersten und dritten Quartals der RW-Verteilung lagen, da wegen der U-Verteilung eine Verfälschung der Iteminterkorrelationen zu erwarten war. 33 Items wurden gemäß Lienert (1969, S. 141 f.) nach der Höhe der Selektionskennwerte ausgewählt. Auf diese Art kamen nur Gelditems oder Sachitems, die klare Mengenangaben hatten, in die reduzierte Liste, nicht dagegen Items mit unklaren Mengen oder Wahlen über Zeitschriften. Die Rohwerte dieser Liste (RW 33) waren ebenfalls U-förmig verteilt ($\chi^2 = 161.5$ bei $df = 7$ gegenüber Gleichverteilung). Die Homogenitätsprüfung dieser 33 Items nach dem zweikategorialen Rasch-Modell ergab unter Benutzung des Algorithmus von Fischer und Scheiblichner (1970) eine hochsignifikante Abweichung der 33 Items vom Modell (Modelltest nach Scoretrennung: $\chi^2 = 128.0$ bei $df = 31$ und einem Informationsbetrag von $I = 831.8$). Auch die Entfernung von Items, die eine signifikante Differenz der Itemparameter aufwiesen, ergab wieder signifikante Modellabweichungen.¹⁾

Wir prüften daraufhin getrennt die Geld-Items und Sach-Items auf Homogenität. Für die Geld-Items ergab sich beim Modelltest eine nicht-signifikante Abweichung vom Modell ($\chi^2 = 31.5$ bei $df = 27$ und Informationsbetrag $I = 926.8$). Bei Entfernung der drei Items mit signifikanter Differenz der Itemparameter ergab sich keine signifikante Modellabweichung ($\chi^2 = 17.1$ bei $df = 24$) und kein erneutes Auftreten von signifikanten Itemparameterdifferenzen. Für die Sach-Items führte der Modelltest nach Scoretrennung ebenfalls unmittelbar zu einer nichtsignifikanten Gesamtabweichung ($\chi^2 = 24.6$ bei $df = 27$ und Informationsbetrag $I = 905.9$). Keine Differenz der Itemparameter war signifikant.

Wir schlossen aus diesen Ergebnissen, daß es sich bei dem Auswahlverhalten um ein zweidimensional gesteuertes Verhalten handelt,

1) Auch verschiedene andere zufällig oder systematisch erstellte Mischungen aus Geld- und Sachitems zeigten dieselben hochsignifikanten Modellabweichungen.

2) Alle Rechnungen wurden auf der Rechenanlage X8 und PDP10 des Rechenzentrums der Universität Kiel durchgeführt.

und daß zwei Rohwerte aus der Summe der mit Aufschieben beantworteten Geld-Items (RWGE) bzw. Sach-Items (RWSA) mit Berechtigung gebildet werden könnten. Die RWGE- und RWSA-Werte waren ebenfalls U-förmig verteilt. Wir bestimmten die Split-Half-Reliabilitäten der bisher erwähnten vier Skalen (Tabelle 2) und die Skalen-Interkorrelationen (Tabelle 3) sowohl für die selektierte Stichprobe von $N_1 = 102$ und die Gesamtstichprobe von $N = 234$.

Tabelle 2

Split-Half-Reliabilitäten der vier Rohwerte (Korrelation der ersten Hälfte mit der zweiten Hälfte).

	$N_1 = 102$	$N = 234$
RW	.51	.92
RW 33	.68	.92
RWSA	.67	.89
RWGE	.73	.90

Die Split-Half-Reliabilität der RWGE scheint am wenigsten durch die Einschränkung des Stichprobenumfangs nach dem ursprünglichen Rohwert (RW) beeinträchtigt zu werden. Die Häufigkeitsverteilungen der vier Skalen nach Einschränkung der Stichproben aufgrund von RW wichen in keinem Fall von der Normalverteilung signifikant ab.

Tabelle 3

Interkorrelationen der Skalen RW, RW 33, RWGE und RWSA bei $N_1 = 102$ und $N = 234$ (in Klammern)

RW 33	.84 (.97)		
RWSA	.45 (.90)	-.08 (.78)	
RWGE	.93 (.98)	-.35 (.67)	
	RW	RW 33	RWSA

Bei den Skalen-Interkorrelationen zeigte sich, daß die Interkorrelation zwischen RWSA und einer beliebigen anderen Skala sehr stark von der Einschränkung des Stichprobenumfangs abhängt.

3.2. Beziehung zwischen Verhaltensprobe und Auswahllistenrohwert

Aus der Tabelle 4 geht hervor, daß ein sehr signifikanter Zusammenhang zwischen den verschiedenen Skalenwerten und dem Verhalten in der Verhaltensprobe bestand. Diese Beziehung scheint aber nur in den Extrembereichen prognostisch brauchbar zu sein. So kann man nur bei hinreichend stabilem Wählen der aufgeschobenen Bedürfnisbefriedigung auf das Aufschieben in der Verhaltensprobe schließen und umgekehrt. Außerdem scheint der Zusammenhang zwischen RWSA und der Verhaltensprobe bei Vergleich der Prüfstatistiken deutlich schwächer zu sein.

Die Korrelation zwischen der Beantwortung des Items 15 in der Auswahlliste und dem Verhalten in der Verhaltensprobenwiederholung betrug $r_{\text{phi}} = .37$ (bei $N = 234$).

Tabelle 4

Häufigkeitsverteilungen der Skalen RW, RW33, RWSA und RWGE

RW	0	8	16	24	32	40	48	
	$\frac{0}{7}$	$\frac{8}{15}$	$\frac{16}{23}$	$\frac{24}{31}$	$\frac{32}{39}$	$\frac{40}{47}$	$\frac{48}{56}$	
Verhaltensprobe = 0 (schiebt nicht auf)	46	19	17	8	6	8	4	108
								$\chi^2 = 45.5$
Verhaltensprobe = 1 (schiebt auf)	16	16	18	14	15	11	36	126
Gesamt	62	35	35	22	21	19	40	234
								$\chi^2 = 41.3$
RW33	0	4	8	12	16	20	24	28
	$\frac{0}{3}$	$\frac{4}{7}$	$\frac{8}{11}$	$\frac{12}{15}$	$\frac{16}{19}$	$\frac{20}{23}$	$\frac{24}{27}$	$\frac{28}{33}$
Verhaltensprobe = 0 (schiebt nicht auf)	54	17	10	6	4	7	6	4
								108
Verhaltensprobe = 1 (schiebt auf)	25	8	12	6	17	6	10	42
Gesamt	79	25	22	12	21	13	16	46
								234
								$\chi^2 = 161.5$

RWSA	0	4	8	12	16	20	24		
	$\frac{0}{3}$	$\frac{4}{7}$	$\frac{8}{11}$	$\frac{12}{15}$	$\frac{16}{19}$	$\frac{20}{23}$	$\frac{24}{28}$		
Verhaltens- probe = 0 (schiebt nicht auf)	34	26	16	7	10	7	10	108	$\chi^2 = 25.2$
Verhaltens- probe = 1 (schiebt auf)	18	17	17	15	8	18	33	126	
Gesamt	51	43	33	22	18	25	43	234	$\chi^2 = 27.8$

RWGE	0	4	8	12	16	20	24		
	$\frac{0}{3}$	$\frac{4}{7}$	$\frac{8}{11}$	$\frac{12}{15}$	$\frac{16}{19}$	$\frac{20}{23}$	$\frac{24}{25}$		
Verhaltens- probe = 0 (schiebt nicht auf)	62	16	7	8	7	4	4	108	$\chi^2 = 47.2$
Verhaltens- probe = 1 (schiebt auf)	32	10	16	6	9	15	38	126	
Gesamt	94	26	23	14	16	19	42	234	$\chi^2 = 143.6$

3.3. Beziehung zwischen Alter und Spannungsbogenmaßen

Parametrische Prüfverfahren erschienen aufgrund der U-Verteilung der Rohwerte zur statistischen Prüfung des Zusammenhangs ungeeignet. H-Tests nach Kruskal-Wallis mit den Altersklassen 15 bis 17, 18, 19, 20, 21 und 22 bis 24 Jahre und den Werten RW, RW33, RWSA, RWGE jeweils als abhängigen Variablen verliefen ohne positives Ergebnis. Es deutete sich auch keine Tendenz in den Daten an. Damit bestand zwischen den Variablen Alter und dem Auswahllistenverhalten kein Zusammenhang bei dieser Stichprobe.

Die Verhaltensprobe und die schon genannten Altersklassen definieren eine Kontingenztafel. Die Abweichung der beobachteten Häufigkeiten von den aufgrund der Randsummen zu erwartenden Häufigkeiten ist bei $df = 5$ mit $\chi^2 = 6.99$ nicht signifikant. Es ließ sich hier aber eine Tendenz für die Aussage ablesen: Wenn eine Vp über 21 Jahre alt ist, nimmt sie eher die sofortige Belohnung und umgekehrt. Tabelle 5 zeigt die entsprechende Vierfeldertafel.

Tabelle 5

Abhängigkeit von Alter und Wahl in der Verhaltensprobe ($\chi^2 = 5.94$ bei $df = 1$)

		Verhaltensprobe		
		schiebt nicht auf	schiebt auf	
Alter	unter 21	72	101	173
	über 21	31	20	51
		103	121	234

3.4. Beziehung zwischen Hypothetischem Sparverhalten und Spannungsbogenmaßen

Es wurde für jeden hypothetischen Sparbetrag der H-Test nach Kruskal-Wallis mit den Auswahllistenrohwerten als abhängige Variable durchgeführt. Außerdem wurden 3 Kontingenztafeln erstellt mit den 3 hypothetischen Sparbeträgen und jeweils der Verhaltensprobe als Variablen. Alle Prüfstatistiken erreichten keine signifikanten Werte. Damit liegt bei dieser Stichprobe von Vpn abweichend von amerikanischen Befunden keine Beziehung zwischen hypothetischem Sparverhalten und Spannungsbogenmaßen vor.

4. Diskussion

4.1. Geldskala und Sachskala in der Auswahlliste

In der Itemanalyse zeigte sich ein Gegensatz zwischen den aufgrund klassischer Itemanalyse und probabilistischer Testtheorie erstellten reduzierten Itemlisten. Erstere führte zu einer Mischung aus Geld- und Sach-Items und schied systematisch zwei Arten von Sach-Items aus. Die probabilistische Itemanalyse ließ dagegen den Schluß auf vom Iteminhalt her zwei deutlich unterscheidbare Skalen zu.

Während die aufgrund klassischer Analyse gewonnene RW33-Skala insgesamt keine neue Information im Vergleich zum Gesamtrohwert brachte, scheint in der Trennung von Sach- und Geldskala eine meßtechnische Differenzierung vorzuliegen, die weitere Beachtung verdient. In einigen anderen Befunden deutet sich an, daß die Sachskala tatsächlich nicht mit der Geld-Skala identisch ist. Zunächst weicht die Verteilung der Skalenwerte der Sachskala weniger stark von der Gleichverteilung ab. Dann ist der Zusammenhang zur Verhaltensprobe bei der Sachskala deut-

lich weniger stark ausgeprägt. Aufgrund der U-Förmigkeit der Verteilungen, die durch schmale Bandbreite des Meßinstruments Auswahlliste verursacht sein kann, sind zwar alle Korrelationen überhöht, aber bei Einschränkung des Stichprobenumfangs reagiert die Sachskala in ihren Interkorrelationen am empfindlichsten. Obwohl diese mögliche Differenzierung zwischen Sachskala und Geldskala für die Frage der Verwendung der Auswahlliste bei der forensischen Diagnostik wegen der fehlenden Altersveränderung keine Bedeutung haben kann, könnte diese Differenzierung von Wert sein bei Messungen des Spannungsbogens im Rahmen von empirischen Untersuchungen wie bei St u m p h a u z e r (1972).

Gründe für das Vorhandensein zweier Skalen könnten in den Verarbeitungsprozessen gesucht werden, die von den Itemeigenschaften zur Antwort des Ankreuzens führen. Dabei muß man davon ausgehen, daß zwei Gruppen von Vpn anscheinend eine globale Vorentscheidung getroffen haben und durchweg aufschieben oder sofortige Auszahlung anstreben. Nur eine dritte Gruppe von Vpn scheint abzuwägen, ob es sich lohnt aufzuschieben. Nimmt man an, daß die subjektiven Bewertungen der Sachmengendifferenzen interindividuell mehr variieren als die subjektiven Bewertungen der Geldbetragsdifferenzen, dann müßte sich bei der Sachskala ein unsystematischeres Verhalten dieser Vpn-Gruppe zeigen, was einerseits zur größeren Annäherung der Verteilung der Rohwerte der Sachskala an die Gleichverteilung führen müßte, andererseits aber auch die Korrelation der Sachskala mit anderen Skalen in starkem Maße davon abhängig machen, ob die Extremgruppen der U-Verteilung der Rohwerte mit in die Korrelation eingehen.

Weiterhin müßten subjektive Bewertungen u. E. bei unklaren Mengenangaben und bei Zeitschriftenauswahl am meisten schwanken. Es wäre deshalb zu erwarten, daß die Lösungswahrscheinlichkeiten dieser Items von den anderen Items systematisch abweichen. In einer Überprüfung der Unterschiede der Zentralen Tendenz zwischen den Lösungswahrscheinlichkeiten der Geld- und Sach-Items mit klaren Mengen einerseits und den restlichen Sachitems andererseits wichen die Lösungswahrscheinlichkeiten der eingeschränkten Stichprobe auch hochsignifikant ab ($p < 0.1\%$ bzw. $z = 4.05$ bei $N_1 = 15$ und $N = 56$ mit Wilcoxon-White-Test für unabhängige Stichproben). Dabei waren die Lösungswahrscheinlichkeiten der Itemgruppen „Unklare Mengen“ und „Zeitschriften“ die größeren und näherten sich der Gleichwahrscheinlichkeit $p = .5$. Angesichts dieser Tatsache ist das Ausscheiden dieser Items bei der klassischen Itemanalyse als bedeutsam aufzufassen, denn es wäre zu erwarten gewesen, daß gerade Items mit mittlerer Schwierigkeit in der Itemstichprobe bleiben. Möglicherweise wählt die klassische Analyse trotz der Kritik von Seiten der probabilistischen Theorie doch sinnvoll Items aus.

Weiterhin konnten zwischen den Sachitems mit klaren Mengen und den Gelditems bei dieser reduzierten Stichprobe keine Unterschiede in der zentralen Tendenz der Lösungswahrscheinlichkeiten gefunden werden. Zieht man hier in Betracht, daß es sich bei diesen Sachwerten zum größten Teil um Zigaretten oder Rauchwaren handelte, kann man diesen Umstand dahingehend interpretieren, daß innerhalb der Strafanstalt außer der Geldwährung eine zweite Währung, die Tabakwährung, herrschte und deswegen die Gruppe der abwägenden Vpn auch bei diesen Items kalkulieren kann, ob sich der Aufschub lohnt. Wiederum sei darauf hingewiesen, daß trotz fehlender theoretischer Fundierung die klassische Itemanalyse sinnvolle Ergebnisse zu geben imstande war, vorausgesetzt man berücksichtigte die korrelationsstatistischen Erfordernisse der Stichprobeneinschränkung wegen der U-Verteilung der Daten.

4.2. Konsequenzen anderer Ergebnisse

Das Ausbleiben eines erfolgreichen Nachweises einer Beziehung zwischen Alter und Auswahllistenverhalten kann inhaltlich so interpretiert werden, daß sich mögliche Altersveränderungen des Spannungsbogens bei dieser inhaftierten Vpn-Gruppe nicht mit dem vorliegenden Auswahllistenverfahren erfassen lassen. Mit dem anderen Verfahren zur Erfassung des Spannungsbogens, der Verhaltensprobe, zeigte sich eine schwache Tendenz zur Abnahme der Aufschubbereitschaft bei über 21jährigen Vpn. Einerseits weist dieses Ergebnis darauf hin, daß sich Altersabhängigkeiten des Spannungsbogens aufdecken lassen, andererseits geht der Befund aber nicht in die für die forensische Diagnostik gewünschte Richtung. Eine mögliche Erklärung wäre, daß die vermutlich längere Haftzeit dieser Vpn ihnen den Entscheidungsgrundsatz nahelegt: „Der Spatz in der Hand ist besser als die Taube auf dem Dach.“

Möglicherweise liegt in der mangelnden formalen Übereinstimmung der Auswahllistensituation und des hypothetischen Sparens der Grund dafür, daß keine Beziehung zwischen den Verhaltensweisen in beiden Situationen gefunden wurde. Sparen bedeutet Aufschub eines Bedürfnisses, ohne festzulegen, ob das Geld später für eine Bedürfnisbefriedigung des gleichen Bedürfnisses oder eines anderen Bedürfnisses ausgegeben wird. Dagegen hält die Auswahlliste zumindest in den Sachitems die Art des Verstärkers konstant. Allerdings zeigte sich bei den Gelditems auch keine Beziehung zum hypothetischen Sparverhalten. Zieht man noch die Befunde *Stump-ha u z e r s* in Betracht, der gerade bei ähnlichen Sparbeträgen zu den Auswahllistenbeträgen eine positive Beziehung fand, bleibt offen, ob sich in unserem negativen Befund nur der kulturelle Unterschied bemerkbar macht oder ob dieses Ergebnis einen Zweifel an dem Versuch fundiert,

auf diese Art die Validität des Auswahllistenverfahrens und der Verhaltensprobe zu bestimmen.

Die Entscheidung der Frage nach der Verwendungsfähigkeit der beiden Verfahren im Zusammenhang mit forensischen Begutachtungen dürfte sich noch nicht empfehlen. Zunächst sollte in zukünftigen Untersuchungen die Stichprobe der Jugendlichen in der Altersbandbreite und auf nichtinhaftierte Vpn ausgedehnt werden. Weiterhin ist in diesem Zusammenhang daran zu denken, daß das Meßinstrument Auswahlliste bezüglich der Items umgestaltet werden könnte. Es sollten auch Items benutzt werden, die einen Befriedigungsaufschub zugunsten anderer Verstärker abverlangen. Dies setzt verfahrenstechnisch aber voraus, daß eine Skalierung der subjektiven Werte der Verstärker vor dem eigentlichen Auswahllistenverfahren vorgenommen werden kann, so daß interindividuell gleiche Bewertungen in das Wahlverhalten eingehen.

Insgesamt weisen Befunde und Überlegungen u. E. auf einen Mangel der Auswahlliste hin, die den Spannungsbogen ohne Bestimmung der subjektiven Bewertungen erfassen soll. Wir folgern daraus, daß eine hinreichende Erfassung des Spannungsbogens in zwei Schritten erfolgen sollte. Im ersten Schritt wäre die subjektive Bewertung von verschiedenen Verstärkern festzustellen. Im zweiten Schritt könnte dann eine Auswahlliste mit individuell zugeschnittenen Items, die in ihrem subjektiven Wertanstieg bei Aufschub interindividuell als gleich beurteilt wurden, den Vpn vorgelegt werden. Ob dieses Verfahren dann von dem Durchführungsaufwand bzw. den kognitiven Anforderungen an die Vp noch praktisch verwendbar wäre, muß abgewartet werden.

Summary

The present paper studies the value of using an american check list for measuring desired delay of gratification. The desired delay of gratification of 234 young delinquents was measured with the help of this check list and by studying their behaviour. Besides, the saving behaviour of subjects in a hypothetical situation was recorded. Classical and probabilistic item analyses brought forth different scales, all of which showed no relation to the age or to the hypothetical saving behaviour of the subjects. Various secondary findings were discussed especially with respect to their consequences on the value of using the list of items and on the possibilities of interpretation.

R é s u m é

Les auteurs appellent „arc de tension“ (Spannungsbogen) la disposition à renoncer à la satisfaction immédiate d'un besoin au profit d'une satisfaction ultérieure jugée plus désirable. Ils se proposent de réunir un premier faisceau d'informations sur l'utilité d'un questionnaire américain (Stumphauer, 1972) visant à mesurer cet arc de tension. A cet effet ils s'adressent à 234 jeunes délinquants et leur demandent de répondre au questionnaire qu'ils comparent avec leur tendance plus ou moins marquée à faire des économies; cette tendance est estimée à partir d'une situation proroquée et partiellement hypothétique.

Les analyses d'items, classiques et probabilistes selon Rasch, permettent la construction de plusieurs échelles. Aucune de ces échelles n'est cependant en relation ni avec l'âge des sujets ni avec leur tendance à l'économie. Les auteurs discutent quelques observations liminaires, puis tentent une explication de leurs résultats et font le point quant — à l'utilité de l'adaptation allemande du questionnaire de Stumphauer.

L i t e r a t u r

- Fischer, G. H. und Scheiblechner, H.: Algorithmen und Programme für das probabilistische Testmodell von Rasch. Psych. Beitr. 12, S. 23—51 (1970).
- Freud, S.: Gesammelte Werke Band II/III. London 1942.
- Lienert, G. A.: Testaufbau und Testanalyse. 3. Auflage. Weinheim 1969.
- Mischel, W.: Preference for delayed reinforcement: An experimental study of a cultural observation. J. Abn. Soc. Psych. 56, S. 57—61 (1958).
- Mischel, W.: Preference for delayed reinforcement and social responsibility. J. Abn. Soc. Psych. 62, S. 1—7 (1961).
- Shybut, J.: Delay of gratification and severity of psychological disturbance among hospitalized psychiatric patients. J. Cons. Clin. Psych. 32, S. 462—468 (1968).
- Stumphauer, J. S.: Increased delay of gratification in young prison inmates through imitation of high-delay peer-models. J. Pers. Soc. Psych. 21, S. 10—17 (1972).

Anschrift der Verfasser:

Dipl.-Psych. Burkhard Hasenpusch
328 Besserer Street, K1N 6B5
Ottawa/Canada

Dr. phil. Wilfried Hommers
23 Kiel
Institut für Psychologie
Neue Universität
Olshausenstraße