

Bei Leistungstests wird der Mehrfach-Wahl-Antwort-Typ bei der Aufgabenstellung häufig benutzt. Bei dieser Antwortform muß unter mehreren vorgegebenen Antwortmöglichkeiten die richtige herausgesucht werden. Die Beliebtheit dieser Multiple-Choice-Tests ist einerseits durch ihre zeitsparende Durchführung und Auswertung begründet, andererseits sind Multiple-Choice-Tests oft von höherer Objektivität als andere Formen der Leistungsmessung und zudem läßt sich die Qualität eines solchen Tests durch gewisse Meßgrößen erfassen, die bei Analyse nach Durchführung des Tests berechnet werden können.

V. Lehmann / H. Ensel

Korrektur von Multiple-Choice-Tests mit dem TI-59

Das hier vorgestellte Programm führt für einen Multiple-Choice-Test die vollständige Testauswertung durch. Diese umfaßt neben der Testkorrektur und dem Ausdruck der Ergebnisse auch die **Distraktorenanalyse** einschließlich Ausdruck gewisser statistisch relevanter Größen, die zur Beurteilung der Qualität des verwendeten Tests benötigt werden.

Die Kapazität des Programms ist durch die Speicherzahl des TI-59 beschränkt und darf 20 Testfragen mit höchstens fünf Alternativen, wobei genau eine Antwort richtig ist, nicht übersteigen. Die Anzahl der an dem Test teilnehmenden Personen ist auf 40 begrenzt.

Hauptproblem ist, je 20 Antworten von 40 Personen, also 800 Daten gleichzeitig im Rechner zu speichern und dann die Auswertung durchführen zu lassen. Die Antwortalternativen sind mit 1 bis 5 bezeichnet; 0 bedeutet, daß keine Antwort gegeben wurde. Die Speicherung der 800 Antworten wird dadurch möglich, daß die maximal 20 Antworten jeder Testperson zu zwei ‚Antwortzeilen‘ zusammengefaßt werden. Entsprechend der 10-stelligen Anzeigekapazität des Rechners können so zweimal 10 Antworten in zwei Registern gespeichert werden. Die ‚Lösungszeile‘ wird in gleicher Weise zusammengefaßt und eingegeben.

Das Programm ist durch einen Trick ganz erheblich beschleunigt worden, was an einem einfachen Zahlenbeispiel erläutert sei. Man bildet zur richtigen Lösungszeile das **Fünferkomplement**; ist beispielsweise $L = 321$ die richtige Lösungszeile, so geht man über zu $\bar{L} = 555 - 321 = 234$. Die Summe aus dem *Komplement* \bar{L} und der gegebenen Antwort A enthält an der Stelle, die einer richtigen Antwort entspricht, stets eine 5. Ist beispielsweise $A = 311$ die gegebene Antwortzeile, so ergibt sich $\bar{L} + A = 234 + 311 = 545$. Nur die zweite Antwort ist falsch, entsprechend ist nur die zweite Ziffer

Distraktoren: Die zu jeder Aufgabe vorgegebenen Falsch-Antworten. Distraction heißt Trennung, Zerstreuung. Hier bedeutet *Distraktorenanalyse* die Interpretation der Testergebnisse durch Ermittlung statistischer Kennwerte, welche z.B. über Schwierigkeit und Trennschärfe der Testaufgaben für eine bestimmte Schülerstichprobe eine Aussage machen.

Fünferkomplement: Komplement bedeutet Ergänzung. In der Mathematik ist z.B. das *Einerkomplement* die Ergänzung einer Ziffer auf den Wert 1, d.h. das Einerkomplement von 1 ist 0, das von 0 ist 1. Entsprechend ist das *Fünferkomplement* die Ergänzung auf den Wert 5.

der Summe $\bar{L} + A$ von 5 verschieden. Durch diese Umformung braucht im Korrekturteil immer nur gegen 5 getestet zu werden, und der Inhalt des Testregisters muß nicht ständig geändert werden. Die Folge ist eine erhebliche Zeitersparnis bei der Durchführung des Korrekturprogramms.

Das Gesamtprogramm besteht aus neun Einzelprogrammen, die auf mehrere **Magnetkarten** verteilt sind:

1. Eingabeteil mit Berechnung der Punktesummen (Karten 1 und 2)
2. Häufigkeitsverteilung der Punktesummen (Karte 3a)
3. Durchschnitt, **Standardabweichung** und **Reliabilitätskoeffizient** (Karte 3b)
4. Antworttabelle (Karte 4a)
5. Vorbereitungsprogramm für die Distraktorenanalyse und die **Schwierigkeitsindizes** (Karte 4b)
6. Distraktorenanalyse (Karte 5a)
7. Schwierigkeitsindizes (Karte 5b)
8. Vorbereitungsprogramm für die Trennschärfeindizes (Karte 6a)
9. Trennschärfeindizes (Karte 6b)

Aus Platzgründen wird nur das erste Programm (Eingabeteil und Berechnung der Punktesummen) mit **Programmablaufplan** und Auflistung des Programms vorgestellt. Die anderen Programme werden erläutert aber nicht aufgelistet*).

Das Programm decodiert die Antwortzeilen während der Testauswertung und speichert die den einzelnen Antworten zugeordneten Zwischenergebnisse als Nachkommateil im Schutzstellenbereich des Rechners. Dadurch genügen 90 Speicher, um die Rechenresultate für spätere statistische Auswertungen zu sichern, ohne daß die eingegebenen Daten für weitere Rechnungen verloren gehen. Der TI-59 rechnet intern mit 13 Stellen, obwohl nur 10 Stellen in der Anzeige sichtbar sind. Der nicht angezeigte Bereich wird als ‚Schutzstellenbereich‘ bezeichnet.

Die Anfangsverteilung des Rechners ist 479.59, das bedeutet, daß 480 Programmspeicherschritte und 60 Datenspeicher zu Verfügung stehen. Da vor Beginn der Eingabe der Antwortzeilen nicht alle 90 Datenspeicher benötigt werden, die für das Korrekturprogramm erforderlich sind, kann die zweite Hälfte

*) Interessierte Leser mögen sich mit dem erstgenannten der beiden Autoren in Verbindung setzen:

Volkhart Lehmann, Simmerner Str. 51, 5400 Koblenz

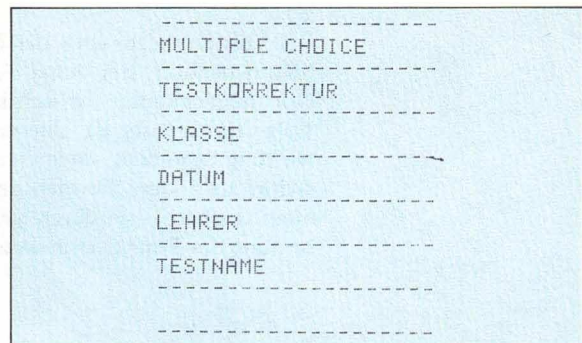


Bild 1 Ausgedruckte Überschriften

des Programmspeichers für das Ausdrucken der Überschriften (**Bild 1**) verwendet werden. Nach Abarbeiten des Programms springt der Rechner durch einen **Reset-Befehl** an den Anfang des Programms und schaltet automatisch auf die Verteilung 239.89 um. Die für das Ausdrucken der Überschriften erforderlichen Codezahlen (**Bild 2**) werden zu Beginn über die Magnetkarten eingelesen und sind in den Speichern R10 bis R37 enthalten.

Magnetkarte: Plastikstreifen mit magnetischer Beschichtung (wie beim „Tonband“) zur Speicherung von Programmen und Daten (zweimal 240 Programmschritte beim TI-59).

Standardabweichung: Bei Meßreihen von wiederholten Beobachtungen kann man in der Regel die „Streuung“ um einen *Mittelwert* angeben. Die Quadratwurzel aus dieser Streuung wird Standardabweichung genannt.

Reliabilität: (engl.: *Reliability*). Zuverlässigkeit bei Tests oder Meßreihen.

Koeffizient: Vorzahl oder Beizahl; eine Zahl, mit der eine unbekannte oder veränderliche Größe multipliziert wird.

Indizes: Mehrzahl von Index (Anzeiger).

Programmablaufplan: Zur grafischen Darstellung von Programmen und Datenbewegungen werden *Flußdiagramme* verwendet (s. dazu Beitrag *Schneider: Grafische Darstellung von Programmabläufen*). Man unterscheidet dabei *Programmablaufplan* und *Datenflußplan*.

Reset: Rücksetzen. Kommando zur Erzeugung eines definierten Anfangszustands in einem Datenverarbeitungssystem.

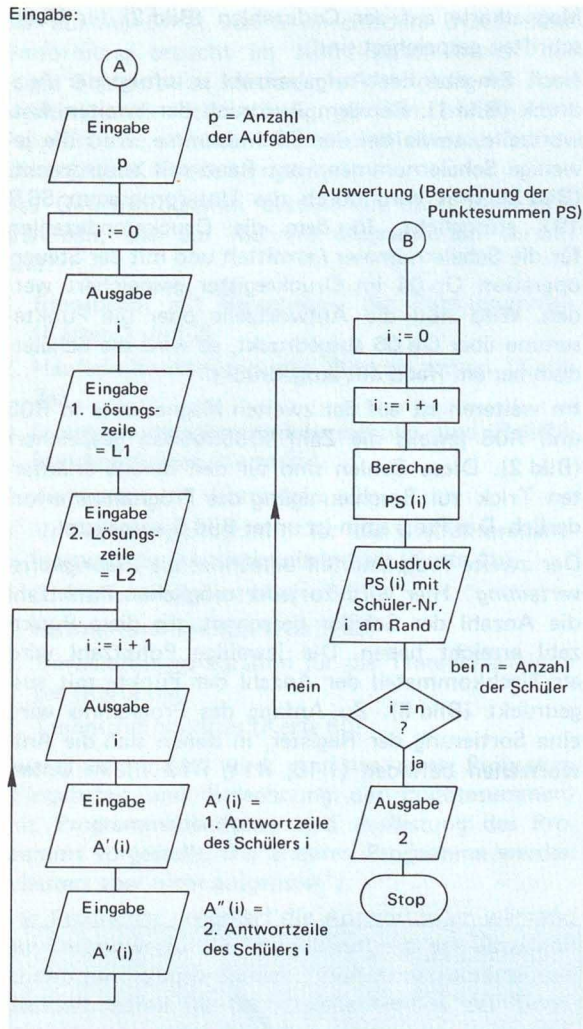


Bild 4 Programmablaufplan für die Berechnung der Punktesummen

Dieser Teil des Programms ist auch notwendig für die Berechnung der Trennschärfeindizes. Wenn man die Häufigkeitsverteilung nicht durchführen möchte, jedoch die Berechnung der Trennschärfeindizes, so ist es möglich, nur den Programmteil des Sortierens ohne Berechnung der Häufigkeitsverteilung durchzuführen.

Im dritten Teil des Programms lassen sich der *Durchschnitt* und die *Standardabweichung* der Häufigkeitsverteilung berechnen (Bild 7). Dies läßt sich mit dem in dem Rechner eingebautem Modul 1 sehr leicht realisieren, das ein Programm zur Berechnung solcher statistischen Werte hat. Damit läßt sich auch der sogenannte *Reliabilitätskoeffizient* berechnen. Er stellt ein Maß für die Zuverlässigkeit des Tests dar, d.h. er macht eine Aussage, inwieweit der Test reproduzierbar ist.

Lienert*) schlägt folgende Formel zur Berechnung des Reliabilitätskoeffizienten vor:

$$r = 1 - \frac{n - M}{m - s^2}$$

mit n = Anzahl der Aufgaben, M = Arithmetisches Mittel, m = Anzahl der Wahlmöglichkeiten pro Aufgabe, s^2 = Varianz.

Lienert empfiehlt diese Formel nur unter folgenden Voraussetzungen, die ein Test enthalten muß:

1. Die Aufgaben sind einander gleichgestellt,
2. Alle Schüler konnten prinzipiell alle Aufgaben in Angriff nehmen,
3. Die Ratewahrscheinlichkeit für alle Aufgaben ist gleich groß.

Problem 1 wird in unserem Test als Voraussetzung akzeptiert. Ob der zweite Punkt seine Gültigkeit hat, muß bei den jeweiligen Tests einzeln untersucht werden. Der dritte Punkt läßt sich mit Hilfe der Distraktorenanalyse untersuchen.

Der vierte Teil des Programms ermittelt eine „Antworttabelle“ (Bild 8). Es wird eine Tabelle aufgestellt, aus der man ablesen kann, wieviele Schüler die Alternative A, B, C, D oder E angekreuzt haben. Die jeweilige Aufgabennummer wird am Rand mit ausgedruckt. Für jeden Schüler wird bei der jeweiligen Aufgabe die Lösungsziffer aus der 10-stelligen Antwortzeile isoliert, und durch INV log in eine Platzziffer verwandelt, die in die jeweilige Tabellenzeile dazuaddiert wird.

*) Lienert: Testaufbau und Testanalyse (S. 228).

programm SBR 000 vorgenommen. Die Sortierung erfolgt nach der Anzahl der richtigen Antworten, die als Nachkommateil in R11, R13... gespeichert sind. Das Programm sortiert also die Speicherinhalte der erwähnten ungeraden Speicher entsprechend dem Nachkommateil und beim Austauschen tauscht es die jeweiligen vorherigen geraden Speicher mit aus. So befinden sich schließlich die Antwortzeilen des besten Schülers in R10 und R11 usw. ...

000	69	DP	069	69	DP	139	02	02	208	08	8	277	69	DP	346	29	CP
001	17	17	070	22	22	140	01	1	209	32	X:T	278	22	22	347	42	STD
002	01	1	071	29	CP	141	01	1	210	22	INV	279	69	DP	348	04	04
003	76	LBL	072	43	RCL	142	42	STD	211	77	GE	280	05	05	349	75	-
004	13	C	073	03	03	143	02	02	212	02	02	281	87	IFF	350	01	1
005	42	STD	074	59	INT	144	43	RCL	213	17	17	282	00	00	351	00	0
006	00	00	075	22	INV	145	02	02	214	85	+	283	03	03	352	95	=
007	65	x	076	67	EQ	146	75	-	215	02	2	284	03	03	353	77	GE
008	02	2	077	00	00	147	09	9	216	95	=	285	97	DSZ	354	03	03
009	85	+	078	51	51	148	95	=	217	42	STD	286	03	03	355	73	73
010	08	8	079	92	RTN	149	55	+	218	03	03	287	02	02	356	50	I×I
011	95	=	080	42	STD	150	02	2	219	69	DP	288	58	58	357	22	INV
012	42	STD	081	07	07	151	95	=	220	24	24	289	25	CLR	358	28	LDG
013	01	01	082	00	0	152	71	SBR	221	01	1	290	91	R/S	359	35	1/X
014	43	RCL	083	69	DP	153	01	01	222	00	0	291	44	SUM	360	65	x
015	00	00	084	04	04	154	93	93	223	00	0	292	08	08	361	43	RCL
016	91	R/S	085	69	DP	155	73	RC*	224	49	PRD	293	99	PRT	362	05	05
017	72	ST*	086	05	05	156	02	02	225	04	04	294	91	R/S	363	95	=
018	01	01	087	43	RCL	157	22	INV	226	43	RCL	295	44	SUM	364	59	INT
019	99	PRT	088	08	08	158	59	INT	227	04	04	296	09	09	365	42	STD
020	91	R/S	089	59	INT	159	55	+	228	44	SUM	297	99	PRT	366	05	05
021	69	DP	090	22	INV	160	01	1	229	03	03	298	86	STF	367	25	CLR
022	21	21	091	44	SUM	161	00	0	230	04	4	299	00	00	368	42	STD
023	72	ST*	092	05	05	162	55	+	231	22	INV	300	61	GTD	369	06	06
024	01	01	093	43	RCL	163	43	RCL	232	28	LOG	301	02	02	370	61	GTD
025	71	SBR	094	09	09	164	08	08	233	49	PRD	302	55	55	371	03	03
026	01	01	095	22	INV	165	22	INV	234	03	03	303	00	0	372	90	90
027	91	91	096	44	SUM	166	59	INT	235	43	RCL	304	42	STD	373	94	+/-
028	73	RC*	097	06	06	167	85	+	236	03	03	305	02	02	374	85	+
029	01	01	098	01	1	168	73	RC*	237	69	DP	306	03	3	375	01	1
030	69	DP	099	00	0	169	02	02	238	04	04	307	03	3	376	00	0
031	06	06	100	42	STD	170	22	INV	239	92	RTN	308	04	4	377	95	=
032	69	DP	101	01	01	171	59	INT	240	43	RCL	309	01	1	378	22	INV
033	20	20	102	43	RCL	172	65	x	241	10	10	310	03	3	379	28	LDG
034	69	DP	103	05	05	173	01	1	242	69	DP	311	01	1	380	35	1/X
035	21	21	104	85	+	174	00	0	243	01	01	312	02	2	381	65	x
036	98	ADV	105	73	RC*	175	00	0	244	69	DP	313	06	6	382	43	RCL
037	61	GTD	106	01	01	176	95	=	245	02	02	314	03	3	383	06	06
038	00	00	107	95	=	177	58	FIX	246	69	DP	315	07	7	384	95	=
039	14	14	108	71	SBR	178	03	03	247	03	03	316	69	DP	385	59	INT
040	76	LBL	109	00	00	179	69	DP	248	69	DP	317	01	01	386	86	STF
041	12	B	110	51	51	180	06	06	249	04	04	318	01	1	387	02	02
042	69	DP	111	69	DP	181	22	INV	250	69	DP	319	07	7	388	42	STD
043	30	30	112	21	21	182	58	FIX	251	05	05	320	03	3	389	06	06
044	43	RCL	113	43	RCL	183	02	2	252	69	DP	321	06	6	390	43	RCL
045	00	00	114	06	06	184	44	SUM	253	00	00	322	04	4	391	04	04
046	42	STD	115	85	+	185	02	02	254	92	RTN	323	01	1	392	55	+
047	04	04	116	73	RC*	186	97	DSZ	255	01	1	324	03	3	393	01	1
048	71	SBR	117	01	01	187	00	00	256	42	STD	325	00	0	394	00	0
049	00	00	118	95	=	188	01	01	257	03	03	326	03	3	395	00	0
050	80	80	119	71	SBR	189	44	44	258	71	SBR	327	00	0	396	95	=
051	55	+	120	00	00	190	91	R/S	259	02	02	328	69	DP	397	42	STD
052	01	1	121	51	51	191	43	RCL	260	40	40	329	02	02	398	08	08
053	00	0	122	43	RCL	192	00	00	261	73	RC*	330	01	1	399	22	INV
054	95	=	123	02	02	193	55	+	262	02	02	331	07	7	400	87	IFF
055	42	STD	124	55	+	194	01	1	263	69	DP	332	03	3	401	02	02
056	03	03	125	01	1	195	00	0	264	01	01	333	01	1	402	02	02
057	22	INV	126	00	0	196	75	-	265	69	DP	334	00	0	403	58	58
058	59	INT	127	00	0	197	59	INT	266	22	22	335	00	0	404	75	-
059	65	x	128	95	=	198	42	STD	267	73	RC*	336	00	0	405	93	.
060	01	1	129	74	SM*	199	04	04	268	02	02	337	00	0	406	01	1
061	00	0	130	01	01	200	95	=	269	69	DP	338	00	0	407	95	=
062	95	=	131	00	0	201	65	x	270	02	02	339	00	0	408	42	STD
063	32	X:T	132	42	STD	202	01	1	271	69	DP	340	69	DP	409	09	09
064	05	5	133	02	02	203	00	0	272	22	22	341	03	03	410	44	SUM
065	22	INV	134	69	DP	204	85	+	273	73	RC*	342	09	9	411	06	06
066	67	EQ	135	21	21	205	01	1	274	02	02	343	81	RST	412	61	GTD
067	00	00	136	97	DSZ	206	95	=	275	69	DP	344	76	LBL	413	02	02
068	71	71	137	04	04	207	32	X:T	276	03	03	345	11	R	414	58	58
			138	01	01												

Bild 5 Programmliste „Eingabeteil mit Berechnung der Punktesummen“

HAEUFIGKEITSV.
0.20
1.19
2.18
7.17
0.16
6.15
5.14
5.13
4.12
0.11
0.10
1.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00

↑ ↑
Punktzahl
Anzahl der Schüler

Bild 6 Zuordnung der Ergebnisse von 31 Schülern zu den erreichbaren Punktzahlen

A. -TABELLE		
A	B	C D E
29020000.		01
31000000.		02
900200200.		03
101260300.		04
3000000100.		05
2803000000.		06
201012700.		07
1614000100.		08
204250000.		09
2003050300.		10
615070300.		11
2301020500.		12
224040100.		13
507041500.		14
201111700.		15
2802010000.		16
100280200.		17
518080000.		18
722020000.		19
808011400.		20

↑
Aufgabennummer

Bild 8 Antworttabelle

DISTRAKTORENANALYSE	
0.01	01
0.	02
1.06	03
1.03	04
0.01	05
0.02	06
0.02	07
2.08	08
1.03	09
1.06	10
2.08	11
1.04	12
1.04	13
2.08	14
1.07	15
0.02	16
0.02	17
1.07	18
1.05	19
2.09	20

↑
Aufgabennummer

Bild 9 Distraktorenanalyse (vgl. Text)

DURCHSCHNITT	14.71
STANDARDABWEICHUNG	2.28
RELIABIL. KOEFF.	0.74

Bild 7 Auswertungen des mit Bild 6 abgedruckten Ergebnisses

Der fünfte Teil des Programms druckt keinen Wert aus. Er dient lediglich zur Vorbereitung der Distraktorenanalyse und der Berechnung der Schwierigkeitsindizes. Hier wird für jede Aufgabe ermittelt, wieviele Schüler die jeweilige Aufgabe richtig beantwortet haben. Die einzelnen Werte werden als Nachkommateil im Schutzstellenbereich des Rechners in R10, R12, R14 usw. eingespeichert.

Der sechste Teil dient der Ausführung der ‚Distraktorenanalyse‘ (Bild 9). Dieser Teil ermittelt, ob die Falschantworten (Distraktoren) eine genügende Attraktivität aufweisen. Dies ist notwendig, um eine Verschiebung der Ratewahrscheinlichkeit zugunsten der richtigen Lösung auszuschließen. Über die Frage,

nach welchen Kriterien brauchbare von unbrauchbaren Distraktoren zu trennen sind, schlagen Seelig und Kuschmann/Zwartz folgendes Verfahren vor:

Entfallen bei drei Distraktoren pro Item (Aufgabe) mehr als die Hälfte aller falschen Antworten auf einen Distraktor, so ist dieser Distraktor zu attraktiv. Entfallen weniger als 10% aller falschen Lösungen auf einen Distraktor, so ist der Distraktor zu unattraktiv. *)

Das Programm gibt für jede Aufgabe nebeneinander zwei Zahlen an, wobei im Intervall dieser beiden Zahlen ein Distraktor der jeweiligen Aufgabe den gestellten Ansprüchen genügt, d.h. der erste Wert gibt 10% der falschen Lösungen dieser Aufgabe, der zweite Wert 50% der falschen Lösungen an. Bei jedem Ausdruck wird hier die jeweilige Aufgabennummer am Rand mit ausgedruckt. Mit diesen Werten kann man mit Hilfe der Antworttabelle die Analyse durchführen. Wir haben es bei der Auflistung dieser Werte gelassen, da eine endgültige Analyse durch den Rechner zu zeitraubend wäre.

*) Kuschmann, W./Zwartz, H.: Objektive Leistungsmessung in der Schulpraxis. Neue Unterrichtspraxis 1975. Heft 3 S. 164.

Im siebten Programm wird für jede Aufgabe ein ‚*Schwierigkeitsindex*‘ berechnet, mit dessen Hilfe man nicht nur eine Aussage über das betreffende Item erhält, als vielmehr auch ein Maß bezüglich des Lernziels in dem betreffenden Lernstoff. *Lienert* schlägt hierzu folgende Formel vor:

$$P = 100 \frac{N_r - \frac{N_f}{m-1}}{N}$$

mit N_r = Zahl der Schüler, die die Aufgabe richtig beantworteten, N_f = Zahl der Schüler, die die Aufgabe falsch beantworteten, m = Anzahl der Wahlmöglichkeiten pro Item, N = Anzahl der Schüler.

Daraus ergibt sich, daß hohe Schwierigkeitsindizes für leichte und niedrigere Schwierigkeitsindizes für schwerere Aufgaben stehen. Neben dem jeweiligen Index wird die Aufgabennummer als Nachkommateil mit ausgedruckt (**Bild 10**).

Das achte und neunte Programm befassen sich mit der Ermittlung der ‚*Trennschärfe*‘. Dieser Index gibt für jede Aufgabe an, inwieweit diese die leistungsfähigen von den weniger leistungsfähigen Schülern trennt. So gibt ein hoher Trennschärfenindex an, daß die Aufgabe von den guten Schülern besser gelöst wurde als von den schlechteren Schülern, wobei ein negativer Index darauf hinweisen würde, daß die schlechteren Schüler die Aufgabe erfolgreicher angingen als die guten Schüler. *Gaude/Teschner**) schlagen folgende Formel vor:

$$T = \frac{N_{ro} - N_{ru}}{N} \times 100$$

mit N_{ro} = Richtige Lösungen für das betreffende Item in der oberen Extremgruppe (25% des am Test teilnehmenden Schüler), N_{ru} = Richtige Lösungen für das betreffende Item in der unteren Extremgruppe, N = Gesamtzahl der teilnehmenden Schüler.

Der Trennschärfeindex setzt die Sortierung der Antwortzeilen (vgl. Häufigkeitsverteilung) voraus. Sie wird in gleicher Form wie der Schwierigkeitsindex ausgedruckt (**Bild 11**).

*) *Gaude, P./Teschner, W.-P.*: Objektivierete Leistungsmessung in der Schule, Frankfurt a.M. 1973. 3. Auflage Seite 112.

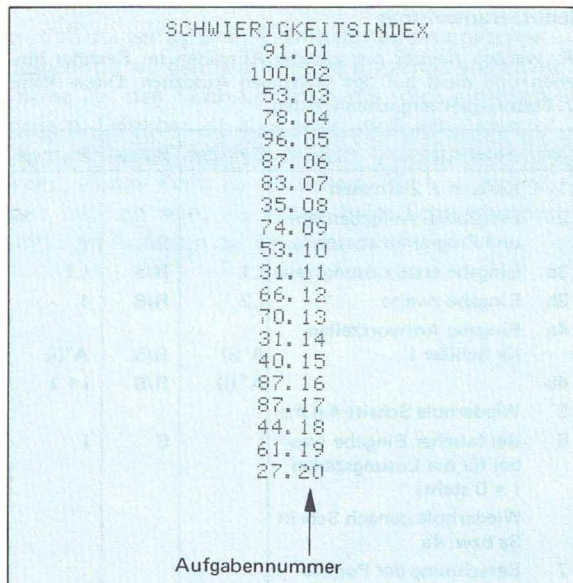


Bild 10 Schwierigkeitsindex (vgl. Text)

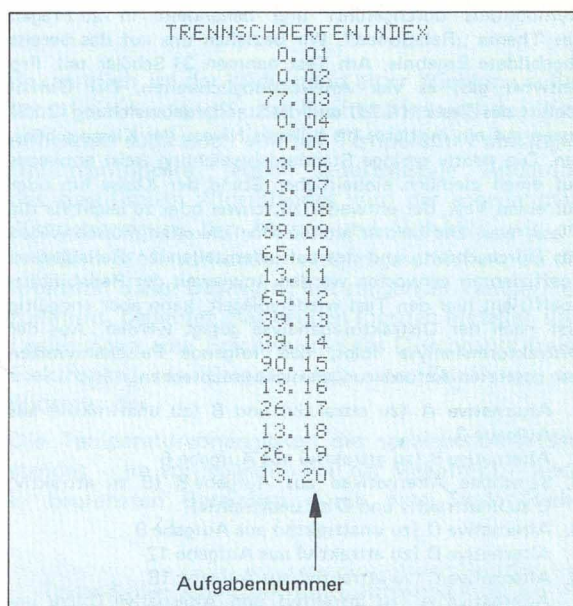


Bild 11 Trennschärfeindex (vgl. Text)

Benutzeranweisung

Wir werden hierbei nur auf die Ausgaben im Rechner eingehen und nicht auf den jeweiligen Ausdruck. Dieser kann im Testbeispiel verglichen werden.

Nr.	Anweisung	Eingabe	Befehl	Anzeige
1	Karte 1 u. 2 einlesen			
2	Eingabe d. Aufgabenzahl und Programm starten	p	A	0
3a	Eingabe erste Lösungszeile	L1	R/S	L1
3b	Eingabe zweite "	L2	R/S	1
4a	Eingabe Antwortzeilen für Schüler i	A'(i)	R/S	A'(i)
4b		A''(i)	R/S	i + 1
5	Wiederhole Schritt 4 n mal			
6	Bei falscher Eingabe (wobei für die Lösungszeilen i = 0 steht) Wiederhole danach Schritt 3a bzw. 4a	i	C	i
7	Berechnung der Punktesummen		B	2

Für die anderen Programme wird die jeweilige Karte eingelesen und über A das Programm gestartet. Bei der Häufigkeitsverteilung ist es möglich, nur einen Teil des Programmes durchlaufen zu lassen (vgl. Häufigkeitsverteilung). Außer dem ersten Programm können alle anderen Programme beliebig übersprungen werden.

Testbeispiel

Der hier vorgestellte Test wurde in einer neunten Klasse eines Gymnasiums durchgeführt und behandelte in 20 Fragen das Thema „Relationen“. Wir beziehen uns auf das bereits abgebildete Ergebnis. Am Test nahmen 31 Schüler teil. Pro Antwort gibt es vier Antwortmöglichkeiten. Der Durchschnitt des Tests (14,71) und die Standardabweichung (2,28) lassen auf ein mittleres bis höheres Niveau der Klasse schließen. Die relativ geringe Standardabweichung weist entweder auf einen ziemlich einheitlichen Stand der Klasse hin oder auf einen Test, der entweder zu schwer oder zu leicht für die Klasse war. Die beiden letzten Möglichkeiten müssen wegen des Durchschnitts und des zufriedenstellenden Reliabilitätskoeffizienten verworfen werden. Inwieweit der Reliabilitätskoeffizient hier den Test widerspiegelt, kann aber endgültig erst nach der Distraktorenanalyse gesagt werden. Aus der Distraktorenanalyse folgt, daß folgende Falschantworten den gesetzten Anforderungen nicht entsprechen:

1. Alternative A (zu attraktiv) und B (zu unattraktiv) aus Aufgabe 3
2. Alternative B (zu attraktiv) aus Aufgabe 6
3. Sämtliche Alternativen aus Aufgabe 8 (B zu attraktiv; C zu unattraktiv und D zu unattraktiv)
4. Alternative D (zu unattraktiv) aus Aufgabe 9
5. Alternative D (zu attraktiv) aus Aufgabe 12
6. Alternative C (zu attraktiv) aus Aufgabe 18
7. Alternative A (zu attraktiv) und Alternative D (zu unattraktiv) aus Aufgabe 19
8. Alternative C (zu unattraktiv) aus Aufgabe 20

Hier zeigt sich, daß relativ viele Distraktoren den gestellten Anforderungen nicht genügen. Besonders auffallend sind die Aufgaben 3, 8 u. 19, die bei erneuter Verwendung des Testes überarbeitet werden sollten, da auch, besonders bei der Aufgabe 8, sich ein überdurchschnittlicher Schwierigkeitsindex zeigt.

Die Schwierigkeitsindizes weisen auf sehr unterschiedliche Aufgabentypen hin; dies zeigen auch die Trennschärfeindizes. Dabei fällt besonders Aufgabe 2 auf, die mit einem Schwierigkeitsindex von 100 zweifellos zu leicht ist.

In Aufgabe 20 finden wir den geringsten Schwierigkeitsindex (27), der aber ohne weiteres noch vertretbar ist. Insgesamt spiegelt sich in den Schwierigkeitsindizes der respektable Durchschnitt wieder. Man kann hier also von einem durchschnittlich schwierigen Test sprechen.

Aus der Bearbeitung bisher geht hervor, daß man dem Reliabilitätskoeffizienten nur beschränkt seine Gültigkeit hier zusprechen kann, da aus der Distraktorenanalyse hervorgeht, daß die Ratewahrscheinlichkeit zugunsten der richtigen Alternative nicht bei jeder Aufgabe gleich hoch ist. Dennoch kann man sich ohne weiteres dazu entschließen, den Test nach Überarbeitung wieder zu benutzen, zumal die Analyse keine gravierenden Mängel aufweist.

Der Trennschärfeindex zeigt, daß vier von den 20 gestellten Aufgaben keine genügende Trennschärfe aufweisen. Dies unterstützt die Vermutung, daß das Niveau der Klasse relativ einheitlich ist (vgl. Standardabweichung). Aufgabe 15 zeigt mit dem Index 90 die absolut größte Trennschärfe.

Erstaunlich zeigt sich der relativ geringe Wert von Aufgabe 20, zumal diese den höchsten Schwierigkeitsindex aufweist. Man muß hier auf eine ziemlich hohe Ratewahrscheinlichkeit schließen, obwohl dies aus der Distraktorenanalyse nicht so deutlich hervorgeht.

Literatur

- Hans Jürgen Becker/Günter Jüngel:* Erarbeitung und Erprobung eines Chemietests (Sekundarstufe 1).
- Gaude, P./Teschner, W.-P.:* Objektiviert Leistungsmessung in der Schule, Frankfurt a.M. 1973 3. Auflage.
- Lienert, G.:* Testaufbau und Testanalyse, Weinheim 1969.
- Kuschmann, W./Zwartz, H.:* Objektive Leistungsmessung in der Schulpraxis, Neue Unterrichtspraxis, 1975, Heft 3 (Teil 1).

Das Programm zur Sortierung der Datenspeicher (vgl. Häufigkeitsverteilung) verdanken wir *Chahriar Bidardel*, Koblenz.