

Ein Fragebogen zur Messung toleranter Einstellungen bei 12 bis 14 jährigen Volksschülern*)

Hans Dieter Schmidt und Harald Euler

Ziel der Arbeit ist es, ein Meßmittel toleranter Einstellungen für Kinder der Volksschul-Oberstufe vorzulegen. Wir sind uns der grundsätzlichen Schwierigkeiten der Konstruktion von Einstellungsfragebogen für Kinder bewußt, sind aber durch die bisherigen Versuche ermuntert worden, die vorläufige Endfassung des Fragebogens zur Diskussion zu stellen. Die Erörterung der theoretischen Probleme einschließlich weiterer experimenteller Ergebnisse geschieht in einer umfassenderen Arbeit, die in Vorbereitung ist. Hier soll über die einzelnen Schritte der *Konstruktion* des Fragebogens berichtet werden.

1.

Aus bereits bestehenden Erwachsenen-Fragebogen (Saarbrücker Liste (Spreen), „Intoleranz gegen Ambiguität“ (Frenkel-Brunswick), „Dogmatismus“ (Rokeach), unveröffentlichte „Demokratie“- und „Antisemitismus“-Fragebogen) wurden Items, die eine tolerante/intolerante Einstellung gegenüber Mitmenschen zu implizieren schienen, ausgewählt. Aus Gesprächen mit Pädagogen und aus schöngestiger Literatur (Lessing, Nathan der Weise) wurden Hinweise zur Formulierung entsprechender Feststellungen entnommen. Eine Vielzahl weiterer Items, die tolerante/intolerante Einstellungen verschiedenster Art aus verschiedenen Lebensbereichen zu enthalten schienen (Familie, Erziehung, religiöser Bereich, Spiel und Sport, Straßenverkehr, fremde Volksgruppen und Minderheiten) wurden eigens entworfen. Auf diese Weise konnten über 100 Feststellungen gesammelt werden.

2.

Diese Items wurden unter folgenden Gesichtspunkten revidiert:

a) *kindgemäße Form*: Jede einzelne Feststellung wurde daraufhin untersucht, ob uns Sinn und Formulierung für 11/12jährige schon verständlich, und für 14jährige nicht mehr zu kindlich erschienen. Die einzelnen sprachlichen Formulierungen wurden dementsprechend abgeändert.

b) *Kürze und Prägnanz*: Die einzelnen Sätze erhielten eine möglichst kurze und prägnante Form; Nebensätze und Partizipialkonstruktionen wurden möglichst vermieden.

*) Folgenden Pädagogen und/oder Psychologen danken wir für freundliche Unterstützung: Frä. Angela Fahrenberg, Dortmund; Frä. Brigitte Herrmann, Duisburg; Frau Almut Simons, Mondorf; Frä. Bärbel Staub, Wiedenest; Herrn V. Becker, Hamm; Herrn H. T. Euler, Meschenich; Herrn H. Hoffmeister, Pemze; Herrn W. Seifert, Solingen; Herrn Dipl.-Psych. H. Simons, Bonn.

c) *Positive und negative Formulierungen*: „Positive“ Formulierungen („Ich finde es gut, daß bei uns so viele Ausländer sind“) wechselten mit negativen („Wer mich einmal verpetzt hat, braucht sich bei mir nie wieder sehen zu lassen“) ab. Dabei überwogen die positiven Formulierungen.

d) *Positive und negative Antworten*: Einmal scheinen Kinder (wie viele Erwachsene) eher bereit zu sein, mit „Ja“ bzw. „Stimmt“ zu reagieren. Zum anderen sollte gewährleistet sein, daß zur „toleranten“ Einstellung auch und insbesondere die Ablehnung „intoleranter“ Aussagen gehört. Deshalb wurde eine Vielzahl von „intoleranten“ Sätzen gebildet, deren negative Beantwortung als Hinweis auf „Toleranz“ gelten sollte.

e) *Antwortkategorien*: Dem Charakter der Items entsprechend (ausschließlich Feststellungen statt Fragen) wurde das Kategorienpaar „Stimmt nicht — Stimmt“ als Antwortmöglichkeit angeboten.

f) *Reihenfolge der Items*: Die Items wurden so gemischt, daß sinngemäß ähnliche möglichst weit voneinander getrennt waren, längere Sätze mit kürzeren abwechselten und ähnliche Satzanfänge nicht aufeinander folgten.

3.

Drei Jungen und zwei Mädchen im Alter von 11 bis 13 Jahren wurden sämtliche Sätze im Einzelversuch vorgelesen. Der VI achtete bei jedem Satz darauf, ob die Feststellung verstanden wurde bzw. welche Formulierung zum *Verständnis* besser geeignet war. Unverständliche und zu „schwierige“ und zu „leichte“ Items wurden schon jetzt eliminiert. Damit verminderte sich der Bestand an Items auf 72.

4.

Die 72 Items und zwei vorgeschaltete „Eisbrecher“ wurden durchlaufend numeriert auf drei DIN-A-4-Bögen zusammengestellt; am rechten Blattrand waren zwei Spalten mit den Überschriften *Stimmt nicht — Stimmt* zum Ankreuzen freigelassen. Die vom VI vor der ganzen Klasse mündlich zu gebende und mehrmals zu wiederholende *Instruktion* enthielt folgende Punkte:

a) Auf dem Blatt stehen Meinungen; jeder Schüler soll hinter jedem Satz ankreuzen, ob das nach seiner Meinung stimmt oder nicht stimmt. Man kann dabei nichts richtig oder falsch machen.

b) Jeder kann ganz ehrlich sein und ganz für sich allein ankreuzen; der Lehrer wird sich die Fragebogen nicht durchlesen; kein Schüler soll bei sich abgucken lassen.

c) Kein Punkt soll ausgelassen werden; wenn jemand nicht sicher ist, welches Kästchen er ankreuzen soll, dann soll er sich für das entscheiden, was am ehesten zutrifft.

5.

Der gesamte Fragebogen wurde insgesamt 203 Schülern des sechsten (N = 50), siebenten (N = 86) und achten Schuljahrs (N = 67) aus drei verschiedenen Schulen gegeben. Die Daten wurden nach *Schwierigkeitsgrad* und

Trennschärfe analysiert (vgl. Lienert, 1961, 84 ff.). Eliminiert wurden solche Items, die

- a) einen Schwierigkeitsindex (p) von 0—20 % und von 80—100 % hatten,
- b) einen Trennschärfeindex (p_{bisfit}) aufwiesen, der nicht mindestens auf dem 1 %-Niveau signifikant war.

Dadurch verringerte sich die Anzahl der Fragebogen-Items auf 28.

6.

Die 28 Feststellungen wurden 10 *Experten* (5 Pädagogen und 5 pädagogisch-psychologisch arbeitenden wissenschaftlichen Psychologen) auf Kärtchen einzeln dargeboten. Anhand von 28 15stufigen (bipolaren) graphischen Rating-Skalen sollte der Grad beurteilt werden, in dem jedes Item tolerantes oder intolerantes Verhalten 12 bis 14jähriger Volksschüler gegenüber anderen ausdrückt. 3 Items, über die das Urteil der Experten auseinanderging, d. h. die im Rating einen Mittelwert in der Mitte zwischen dem „toleranten“ und „intoleranten“ Pol erhielten, wurden eliminiert. — Dieses Beurteilungsexperiment wurde im Gruppenversuch mit 24 Studierenden der Psychologie wiederholt. Die Übereinstimmung ihrer Beurteilungen mit denen der Experten war ungewöhnlich hoch: Die Durchschnittswerte der Ratings beider Gruppen korrelierten mit $Rho = 0.908$.

7.

In diesem Stadium der Konstruktion wurde eine *Inhaltsanalyse* der 25 verbleibenden Items vorgenommen. Der im Fragebogen operational definierte Begriff der Toleranz enthält demnach thematisch die Ablehnung folgender Spielarten intoleranten Verhaltens (geordnet nach der Häufigkeit der Themen):

- Forderung nach autoritärer Erziehung und Führung
- Vorurteile allgemeiner Art
- Intoleranz gegenüber Menschen mit anderer Religion
- Intoleranz gegenüber Menschen anderer Nationalität
- Intoleranz gegenüber Minderheiten
- Forderung nach persönlichen Vorrechten
- Strenge Auswahl von Freunden und Spielgefährten
- Intoleranz gegenüber Mehrdeutigkeit (Ambiguität)
- Forderung nach Nichteinmischung
- Allgemeine Antipathie gegen andere
- Intoleranz gegenüber Schulkameraden.

8.

Zur Kontrolle möglicher *Verfälschungstendenzen* bzw. zur Erfassung der allgemeinen Einstellung gegenüber dem Fragebogen wurden die 11 „Lügen“-Items aus der Kinder-Form der MAS-Skala von Castaneda u. a. in den Fragebogen aufgenommen. Es handelt sich um eine noch nicht standardi-

sierte Übersetzung der amerikanischen Form; an der Itemanalyse der Kinder-Lügenliste wird gearbeitet. Trotzdem wurden bei den weiteren Untersuchungen die Fragebogen derjenigen Kinder nicht weiter berücksichtigt, die mehr als die Hälfte der L-Items „verfälschend“ beantworteten.

9.

58 Schüler des 6. bis 8. Schuljahrs füllten den gesamten Fragebogen aus; nach Eliminierung von sieben Protokollen mit zu hohen L-Scores erfolgte eine erneute Bestimmung der *Trennschärfe*-Indices. Dabei erreichten 3 Koeffizienten nicht das geforderte Signifikanzniveau von 1 %. Bei einer dritten Trennschärfebestimmung an 70 vergleichbaren Schülern ergaben sich keine wesentlichen Änderungen mehr. Dies ist die vorläufig endgültige Form des Fragebogens:

	Schwier. (p)	Trennsch. ($r_{pbis\text{rit}}$)
	(N = 203)	
1. E Wenn die Schule aus ist, gehe ich gerne spielen.		
2. E Wenn jemand viel Geld hat, kann ich ihn gut leiden.		
3. L+ Ich kann jeden gut leiden, den ich kenne.		
4. T- Wenn ich später Kinder habe, würde ich genau aufpassen, mit wem sie spielen.	50.78	.2506**
5. L- Ich möchte in einem Spiel lieber gewinnen als verlieren.		
6. T+ Wenn in der Klasse ein Streber ist, darf man ihn trotzdem nicht ärgern	79.86	.2114**
7. T- Manchen Leuten kann man schon von außen ansehen, daß sie nichts taugen.	55.44	.3068**
8. T- Das Land, wo man geboren ist und wo man zur Schule geht, ist auch das beste Land.	68.31	.2429**
9. L+ Ich bin stets freundlich.		
10. T- Mohammedaner, Christen und Juden glauben alle an etwas Verschiedenes, aber nur einer von ihnen hat Recht.	57.42	.3473**
11. T- In einer richtigen Familie sollen sich alle nach dem Vater richten; andere haben nichts zu sagen.	76.90	.3407**
12. L+ Ich benehme mich immer gut.		
13. T+ Mit Zigeunerkindern würde ich genau so gerne spielen wie mit anderen Kindern.	67.81	.4020**
14. L+ Ich bin zu allen Leuten stets nett.		

15.	T-	Wenn man Recht hat, braucht man nicht darauf zu hören, was die anderen sagen	68.31	.3722**
16.	L+	Ich sage immer die Wahrheit.		
17.	T-	Wenn wir mit meinen Sachen spielen, hab ich am meisten zu sagen.	48.80	.3948**
18.	T-	Ich kann mir nicht vorstellen, daß einer gerne in Rußland lebt.	70.01	.3392**
19.	L+	Ich werde nie ängstlich.		
20.	T-	Wenn man immer darauf hört, was andere sagen, hat man es nicht so gut.	37.47	.2523**
21.	T-	Schlechte Menschen haben keine guten Eigenschaften.	63.10	.3459**
22.	L+	Ich sage nie etwas, was ich nicht sagen soll.		
23.	T+	Ein Indianerhäuptling muß auch darauf hören, was der Dümme sagt.	67.05	.2697**
24.	T-	Menschen, die keine Arbeit finden, weil sie immer meckern, denen soll man auch nichts geben.	66.06	.3370**
25.	T-	Es gibt nur zwei Sorten von Leuten: gute und schlechte. Dazwischen gibt es keine.	71.49	.2678**
26.	T-	Wenn ich zu sagen hätte, dann würde ich dafür sorgen, daß alle Menschen den gleichen Glauben bekommen.	54.72	.2335**
27.	L+	Ich bin immer gut.		
28.	T-	Bei Fußball-Länderspielen muß man immer mit der deutschen Mannschaft halten, auch wenn sie mal schlecht und gemein spielt.	58.67	.4824**
29.	T-	Man kann ruhig jemanden zu etwas zwingen, was er nicht mag; Hauptsache, es ist gut für ihn.	67.05	.2219**
30.	L-	Es ist gut, in der Schule gute Noten zu bekommen.		
31.	T-	Die in der Minderheit sind, müssen den Mund halten; nur die Mehrheit entscheidet	73.95	.4640**
32.	L+	Ich lüge nie.		
33.	T-	Es gibt viele verschiedene Religionen auf der Erde, aber nur die Christen haben Recht.	65.57	.2843**

Fortsetzung siehe nächste Seite

	Schwier. (p)	Trennsch. (r_{pbis}) (N = 203)
34. T- Es gibt Leute, die man wirklich zu nichts gebrauchen kann.	35.64	.4362**
35. T- Wovon ich fest überzeugt bin, da kann mir niemand was reinreden.	42.08	.3133**

(E = „Eisbrecher“-Items; L = „Lügen“-Items; T = „Toleranz“-Items; „+“ = „Stimmt“-Reaktion erhält 1 Punkt; „-“ = „Stimmt-nicht“-Reaktion wird gezählt.)

10.

Die Überprüfung der Verteilung der T-Scores bei 70 Schülern (vgl. 9.) auf *Normalität* nach Kolmogoroff-Smirnoff ergab, daß die empirischen Werte nur mit einer Wahrscheinlichkeit von $p > 20\%$ von den theoretischen Werten abwichen ($D = 0.0811$ bei $M = 9.614$ und $\sigma = 3.530$).

11.

Die Ermittlung der *Retest-Reliabilität* erfolgte zunächst an 29 Schülern des 7. und 8. Schuljahrs mit *einer* Woche Abstand zwischen beiden Untersuchungen, und an 13 Schülern des 6. bis 8. Schuljahrs einer anderen Schule mit einem Abstand von *acht* Wochen (dazwischen lagen die großen Ferien). Die Retest-Koeffizienten betragen im ersten Falle $Rho = 0.859$ und im zweiten Falle $Rho = 0.804$.

Zur Berechnung der *split-half-Reliabilität* wurde bei 70 Schülern verschiedener Schulen und Klassen (vgl. 9.) ein Zusammenhang zwischen den T-Werten des nach geradzahligen und ungeradzahligen Items halbierten Fragebogens von $r_{12} = 0.612$ gefunden. Bei nahezu gleichen Streuungen entspricht dies nach der Spearman-Brown-Formel (Lienert, 1961, 217) einer Stabilität des gesamten T-Fragebogens von $r_{tt} = 0.759$.

12.

Um den Zusammenhang des Fragebogens mit geeignet erscheinenden *Außenkriterien* zu ermitteln, wurden vorläufig zwei Aspekte berücksichtigt.

a) Bei der Schwierigkeit, mögliche Außenkriterien „toleranter“ Einstellungen 12 bis 14jähriger Volksschüler zu finden, konnte auf das *Lehrerurteil* nicht verzichtet werden. Die Lehrer wurden dabei mündlich instruiert, daß unter „Toleranz“ der Grad zu verstehen sei, in dem sich 12 bis 14jährige gegenüber ihren Mitmenschen (vor allem ihren Mitschülern) konzilient verhielten; gemeint sei aber nicht einfach ein passives Gewährenlassen und „Erdulden“ anderer, sondern auch die aktive Billigung von Menschen, die anders sind, anders denken usw.

Nachdem es sich bei vier Schulklassen zunächst herausgestellt hatte, daß sich die Lehrer überfordert fühlten, wenn sie alle Schüler hinsichtlich deren „toleranten“ und „intoleranten“ Verhaltens aufteilen sollten, erhielten die Lehrer von 4 anderen Klassen (6. bis 8. Schuljahr) die Instruktion, entsprechend extreme Schüler in eine Rangreihe einzustufen und dabei so weit wie möglich in den Mittelbereich vorzustoßen. Die Lehrer, die übrigens keinerlei Kenntnis von den Fragebogenresultaten hatten, stellten dabei wie von selbst eine vollständige Rangreihe her. Die Spearman'schen Rangkorrelationskoeffizienten dieser Lehrerbeurteilungen mit den T-Werten des Fragebogens zeigten eine befriedigende Größe:

Lehrer	Schuljahr	Anzahl der Schüler	Rangkorrelation Lehrerurteil/ T-Fragebogen	p
A	6.—8.	12	0.518	> .05
B	6.—8.	17	0.492	< .05
C	6.—8.	24	0.432	< .05
D	6.—8.	31	0.456	< .01

Aus den in Rangreihenform vorliegenden Lehrerurteilen wurden die „tolerantesten“ und „intolerantesten“ 25 Prozent ausgewählt und hinsichtlich ihrer Fragebogenscores miteinander verglichen. Ein Vergleich mit dem U-Test (Mann u. Whitney) ergab einen auf dem 0.005-Niveau gesicherten Unterschied ($z = 3.091$). Der Mittelwert der 16 von Lehrern als „tolerant“ eingestuften Schüler betrug 11.88, derjenige der 16 „intoleranten“ 7.50 Fragebogenscores. Auch der Mittelwertsvergleich mit dem t-Test ergab einen sehr signifikanten Unterschied ($t = 4.293$) zwischen den Extremgruppen.

b) Um den Zusammenhang des „Toleranz“-Fragebogens mit einem mehr direktem Kriterium toleranten Verhaltens zu ermitteln, unterzogen wir zunächst 29 und dann 70 Schüler der drei Volksschul-Oberklassen einem sozialpsychologischen Experiment. Jeder Schüler erhielt eine Liste mit 42 (im ersten Versuch) bzw. 60 (im zweiten Versuch) Namen von Nationen, Völkern und Stämmen (z. B. Amerikaner, Ägypter, Russen, Bayern, Mongolen, Schwaben usw.). Unter ihnen befanden sich auch — in Anlehnung an ein Experiment von Murphy und Likert (vgl. Hartley u. Hartley) — fiktive Volksbezeichnungen wie Antagonier, Dardanesen, Hortonen. Die Schüler sollten diejenigen Völker unterstreichen, die ihnen sympatisch waren, und diejenigen durchstreichen, die sie unsympathisch fanden und ablehnten. Als ein Index toleranten Verhaltens gegenüber anderen Gruppen wurde der Quotient aus der Anzahl der unterstrichenen Völkernamen zur Anzahl der überhaupt angestrichenen (also unterstrichenen plus durchgestrichenen) Völker berechnet ($p_{Tol/tot}$).

Im ersten Experiment ($N = 29$) ergab sich ein Zusammenhang zwischen T-Scores und $p_{Tol/tot}$ -Scores von $Rho = 0.447$ ($p < .05$). Im zweiten Ver-

er wesentlich größeren Zahl von Schülern verschiedener Schulen und Klassen und einer vermehrten Anzahl von zu beurteilenden Völkern ergab sich ein Korrelationskoeffizient von $r_{xy} = 0.361$ ($p < .01$). Zusätzlich wurde ein Extremgruppenvergleich angeschlossen. Die $p_{Tol/tot}$ -Werte der extremen (25 %) Fragebogen-Scorer ($N_{T+} = 13$; $N_{T-} = 12$; $N = 70$) wurden mit dem U-Test auf ihren Unterschied geprüft; es ergab sich eine statistisch gesicherte Differenz ($z = 2.013$; $p < .05$). Bei fast gleichen Streuungen der Fragebogenscore der beiden Extremgruppen ergab sich auch mit dem t-Test ein erheblicher Unterschied dieser Gruppen in den Toleranz-Werten des Völker-Experiments ($M_{T+} = 0.755$; $M_{T-} = 0.568$; $t = 2.435$; $p < .05$). Es hat sich also gezeigt, daß Schüler, die im Toleranz-Fragebogen eine höhere Punktzahl erreichen als andere, auch stärker als diese geneigt sind, andere Gruppen eher zu akzeptieren als abzulehnen.

Nachdem sich ergeben hat, daß der T-Fragebogen den Ansprüchen der Testkonstruktion genügt, wie sie für Fragebogen gefordert werden (Eysenck, 1953; Lienert, 1961) muß nun die bisher erfolgreich vorgenommene Kriteriumsvalidierung erweitert und in Richtung auf eine möglichst umfassende Construct-Validierung (Cronbach u. Meehl, 1955) überschritten werden.

Zusammenfassung

Ein Fragebogen zur Messung „toleranter“ Einstellungen bei 12 bis 14-jährigen Volksschülern (6. bis 8. Schuljahr) wird vorgelegt, indem über die einzelnen Schritte der Konstruktion berichtet wird. Das Meßmittel genügt offenbar den Anforderungen der Testkonstruktion, wie sie für Fragebogen zu stellen sind. Weitergehende Untersuchungen werden angekündigt.

Literatur

- Castaneda, A., McCandless, B. R., and Palermo, D. S. (1956) The children's form of Manifest Anxiety Scale. *Child Develop.* 27, 317—326.
- Cronbach, L. J., and Meehl, P. E. (1955) Construct validity in psychological tests. *Psychol. Bull.* 52, 281—302.
- Eysenck, H. J. (1953) Fragebogen als Meßmittel der Persönlichkeit. *Z. exp. angew. Psychol.* 1, 291—335.
- Frenkel-Brunswik, E. (1949) Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.* 18, 108—143.
- Guilford, J. P. (1956) *Fundamental statistics in psychology and education.* New York—Toronto—London (3rd ed.).
- Hartley, E. L., and Hartley, R. E. (1946) *Problems of prejudice.* Zit. n. Hartley and Hartley, (1955) *Die Grundlagen der Sozialpsychologie.* Berlin.
- Lienert, G. A. (1961) *Testaufbau und Testanalyse.* Weinheim.
- Lienert, G. A. (1962) *Verteilungsfreie Methode in der Biostatistik.* Meisenheim a. Glan.
- Rokeach, M. (1954) The nature and meaning of dogmatism. *Psychol. Rev.* 61, 194—204.

- S p r e e n, O. (1961) Konstruktion einer Skala zur Messung der manifesten Angst
in experimentellen Untersuchungen. Psychol. Forsch. 26, 205—223.
- S ü l l w o l d, F. (1962) Entwicklung und Verwendung von Einstellungsskalen.
Stud. Gen. 15, 126—135.

Anschrift der Verfasser:

Dr. H. D. Schmidt und H. Euler,
Psychol. Inst. d. Univ. Bonn,
53 Bonn, Am Hof 4