

Aus dem Psychologischen Institut der Universität Bonn
(Direktor: Prof. Dr. H. Th o m a e)

Zur Wirkung der Prestige-Suggestion bei der Beurteilung politischer Texte

Hans Dieter Schmidt und Christiane Schmerl

I. Problemstellung

Eine große Anzahl von Experimenten zur Meinungs- und Einstellungsänderung hatte die Wirkung des „Etiketts“ von Botschaften zum Gegenstand. Gemäß der Annahme, daß für die Wirkung einer Kommunikation die Quelle ebenso wichtig sei wie der Inhalt, experimentierte man spätestens seit den 30er Jahren mit prestigehaltigen Etiketts, um Bedingungen der meinungs- und einstellungsändernden Wirkung durch derartige „Prestige-Suggestion“ zu erforschen. Stukat (1958) berichtet über entsprechende Experimente seit Binet (1900), bei denen „Prestige-Suggestion“ durch den Versuchsleiter bzw. dessen Prestige selbst gegeben war. Eine Sammlung weiterer Arbeiten findet sich bei Hovland u. a. (1959).

Gelegentlich ergaben sich Hinweise auf die Existenz eines eigenen Faktors „Prestige-Suggestibilität“, der als „tertiäre Suggestibilität“ die von Eysenck (1943) gefundenen orthogonalen Faktoren „Primäre Suggestibilität“ und „Sekundäre Suggestibilität“ (im Sinne eher motorischer und eher perzeptiver Reaktionen) ergänzen würde; das Für und Wider einer solchen Drei-Faktoren-Theorie der Suggestibilität behandelt Stukat (1958). Erheblich sind jedoch jeweils die bei Faktorisierungen von Variablen der Prestige-Suggestion gefundenen spezifischen bzw. situationsspezifischen Varianzanteile.

Bei Versuchen mit „Etiketts“ wurden als prestigehaltige Quellen vor allem vorgebliche Experten und Sachverständige, bekannte und beliebte Personen des öffentlichen Lebens (Politik, Wissenschaft, Literatur) oder einfach die Feststellung verwendet, es handle sich um die Meinung der Mehrheit.

Die Abhängigkeit des Grades der Meinungsänderung unter dem Einfluß von Prestige von der Beliebtheit und Glaubwürdigkeit des Kommunikators sowie vom Grad der erlebten Diskrepanz zwischen prestige-suggerierter und bereits bestehen-

der Meinung wurde vielfach untersucht und bestätigt; Erklärungsversuche für letztere bieten H o v l a n d u. a. (1957) unter dem Gesichtspunkt der „Ich-Beteiligung“, und F e s t i n g e r und A r o n s o n (1960) mit der Theorie der „cognitiven Dissonanz“ an. Die kombinierte Wirkung von Glaubwürdigkeit (des Verfassers) und Diskrepanz (zur eigenen Meinung) erforschten A r o n s o n u. a. (1960).

In den uns bekanntgewordenen Arbeiten wurden folgende „intervenierende Variable“ besonders erwähnt und teilweise untersucht: Alter der Vpn, Intelligenz, Affektivität, Suggestibilität, Konzentrationsfähigkeit, bisherige Einstellung, momentane Stimmung, Ich-Beteiligung, reales Wissen auf dem betreffenden Gebiet. Die Frage nach persönlichen Bedingungen unterschiedlicher Prestige-Suggestibilität soll im Rahmen dieser Arbeit nicht weiter berücksichtigt werden. Routinemäßig wird lediglich ein Intelligenz-Kurzmaß in die Hypothesenbildung einbezogen.

Während nun die überwiegende Anzahl der Autoren die im Verlaufe von Prestige-Suggestions-Versuchen gewöhnlich resultierende reduzierte Diskrepanz zwischen eigener und suggerierter Meinung im Sinne einer echten — wenn auch vielleicht vorübergehenden — Meinungsänderung interpretieren, nehmen A s c h (1940, 1948, 1957), L e w i s (1938, 1940, 1941) und B i r c h (1945) eher eine „Bedeutungsänderung“ an. Da der Sinn einer Feststellung nicht unabhängig von ihrer Quelle sei, vollziehe sich mit dem Wechsel des Etiketts auch ein Bedeutungswechsel des Textes. Als Beispiele hierfür lassen sich politische Begriffe wie „Freiheit“, „Sozialismus“ oder „Selbstbestimmungsrecht“ anführen, die je nach ihrem „Absender“ unterschiedliche Bedeutungen haben können. A s c h (1940) zieht daraus den Schluß, daß sich bei Versuchen zur Prestige-Suggestion weniger eine Änderung der Beurteilung des Gegenstandes als vielmehr eine Änderung des Gegenstandes der Beurteilung zeige.

Im übrigen haben prestigehaltige Etiketts keineswegs eine automatisch suggestive Wirkung. In den Experimenten von B i r c h (1945) hatten die unstimmgigen Etiketts „nur“ zur Folge, daß die Vpn die sprachlichen Formulierungen kritischer betrachteten.

Auch L e w i s (1941) fand in der Mehrzahl der Fälle keine Meinungsänderung; bei solchen Vpn, die ihre Meinung aber in der erwarteten Richtung änderten, machte sie die Beobachtung, daß diese Probanden nach einem neuen B e z u g s s y s t e m suchten, in das sie sowohl das Prestige-Etikett als auch den Text einordnen konnten. Nach N e w c o m b (1959) bestand die Wirkung der Prestige-Suggestion darin, nach einem anderen möglichen Bezugssystem zu suchen.

In diesem Zusammenhang dürfte von Bedeutung sein, wie sowohl das bisherige Bezugssystem als auch das dargebotene Material „strukturiert“ sind: scharf (eindeutig) oder unscharf (mehrdeutig). „Meinungsänderung“ (zu einem unveränderten Gegenstand) dürfte es nur bei eindeutigem Material geben; im Falle mehrdeutigen Materials wird man eher auf „Bedeutungsänderung“ im Sinne A s c h s schließen müssen. Dessen Kritik besteht sicherlich insofern zu Recht, als in der Regel relativ mehrdeutiges Material verwendet wurde. Dies ist allerdings nur allzu verständlich, denn wirklich „eindeutiges“ Material verrät in der Regel den echten Verfasser oder macht zumindest das Etikett unglaubwürdig. Im Übrigen wird sprachliches Material wohl stets mehrere Deutungen zulassen. Immerhin ist es trotz dieser Einwände nicht völlig ausgeschlossen, daß es echte „Meinungsänderung bei Bedeutungskonstanz“ gibt. Bei Arbeiten zur Prestige-Suggestion sollte die Art der verwendeten Texte besondere Berücksichtigung finden.

In der vorliegenden Arbeit soll die Änderung der Beurteilung nationalsozialistischer Texte (kurz: Nazi-Texte) unter dem Einfluß des Prestiges eines stimmigen („negativen“) und nicht stimmigen („positiven“) Etiketts geprüft werden. Es soll sich zeigen, ob die Methode der Prestige-Suggestion auch bei derartigem politischen Material, zum jetzigen Zeitpunkt und unter anderen Umweltbedingungen (die referierten Arbeiten sind zumeist älter und ausschließlich amerikanischen Ursprungs) zu interpretierbaren Ergebnissen führt. Dabei werden lediglich die Glaubwürdigkeit bzw. Beliebtheit des Etiketts und die Eindeutigkeit der Texte variiert.

Insofern „stark nazistische“ Texte einen größeren Grad erlebter Eindeutigkeit aufweisen als „schwach nazistische“, soll mit dem genannten Material gleichzeitig ein Beitrag zur Frage der Abhängigkeit der Meinungsänderung vom Grad der Eindeutigkeit/Mehrdeutigkeit des Reizmaterials geleistet werden. Es ist anzunehmen, daß eine Meinungsänderung unter dem Einfluß von Prestige dann gefördert wird, wenn weniger eindeutiges Textmaterial zu beurteilen ist.

II. Methode

a) Vorversuche

Zur Frage, welche politischen Personen als „positives“ und welche als „negatives Etikett“ geeignet seien (Zeitpunkt: Frühjahr 1966) sowie zur Erprobung der Beurteilung politischer Texte an Vpn überhaupt wurden 23 Psychologiestudenten beiderlei Geschlechts (Erstsemester) einem ersten Vorversuch unterzogen. Um sicherzustellen, daß allgemein beliebte und allgemein unbeliebte Politiker tatsächlich mit einiger

Wahrscheinlichkeit positives oder negatives Prestige haben, wurde zunächst die relative Beliebtheit von 12 Politikern festgestellt. Die Vpn erhielten folgende Instruktion:

„Bitte ordnen Sie die folgenden politischen Personen danach, wie sympathisch sie Ihnen persönlich sind (von 1 = dem Beliebtesten bis 12 = dem Unbeliebtesten): Konrad Adenauer, Willi Brandt, Ludwig Erhard, Josef Goebbels, Theodor Heuss, Adolf Hitler, Heinrich Lübke, Erich Mende, Carlo Schmid, Franz-Josef Strauß, Walter Ulbricht, Herbert Wehner.“

Es ergaben sich folgende Rangplatzmittelwerte: Heuss 1,8; Adenauer 2,9; Erhard 4,5; Schmid 4,7; Brandt 5,8; Wehner 5,8; Lübke 5,83; Strauß 7,7; Mende 8,3; Goebbels 10,1; Ulbricht 10,4; Hitler 11,2. Die Übereinstimmung der 23 Beurteiler war erheblich ($W = 0.84$).

In der gleichen Sitzung mußten die gleichen Vpn noch einen Persönlichkeitsfragebogen ausfüllen, Polaritätenprofile erstellen und Wortreihen lernen und reproduzieren; diese Aufgaben stammten aus anderen Untersuchungen.

Sodann wurden jeder Vp zwei willkürlich ausgewählte Nazi-Texte vorgelegt. 6 Vpn erhielten Text 1 mit dem Etikett Heuss und Text 2 mit dem Etikett Erhard; 6 Vpn erhielten Text 1 mit Erhard und Text 2 mit Heuss; 11 Vpn erhielten die Texte ohne Etikett. Bei den 12 Vpn mit Etikett handelte es sich um solche, die Heuss oder Erhard den 1. oder 2. Platz gegeben hatten. Gefordert war, die Texte nach dem Grad der persönlichen Zustimmung/Ablehnung auf einer graphischen 7-Punkte-Skala zu bewerten.

Zwischen den Beurteilungen der Vpn „mit“ und „ohne“ Etikett zeigten sich keine statistisch erheblichen Unterschiede für die beiden Texte; die Differenzen wiesen lediglich in die zu erwartende Richtung.

Aufgrund dieses Vorversuchs vermuteten wir, daß bei einer größeren Stichprobe und bei Kontrolle der Variablen „Nazi-Texte“ interpretierbare Resultate zu erwarten seien. Die Versuchsanordnung selbst (Art der Aufgabe, Art der Texte, Rating-Form) erschien als unbedenklich.

Um nun „eher eindeutige“ und „eher mehrdeutige“ Texte unterscheiden zu können, empfahl es sich, eine Reihe von Nazi-Texten nach ihrer „Stärke“ eindimensional zu skalieren. Die Schwierigkeit der Textauswahl bestand darin, daß Texte mit Nazi-Inhalt im allgemeinen durch den Gebrauch drastischer Ausdrucksweisen und bestimmter Begriffe leicht als solche zu erkennen sind. Bei intensiver Lektüre von Hitlers „Mein Kampf“ fanden sich nur ein knappes Dutzend Passagen von jeweils etwa 80 Wörtern Länge, die a) auch aktuelles Interesse beanspruchen können, b) typische Nazigedanken enthalten und c)

auch äußerlich unverfänglich erscheinen können. Sie wurden durch einige Zitate von anderen Nazi-Autoren ergänzt.

Zehn Nazi-Texte von ähnlicher Länge wurden in einem weiteren Versuch 66 Psychologiestudenten im 3. und 4. Semester in Paarvergleichsordnung gegeben. Es sollte jeweils bestimmt werden, welcher von beiden Texten „mehr nationalsozialistisch“ sei. Aufgrund dieses Versuchs wurden sämtliche Texte nach der Methode des Paarvergleichs von Thurstone (case 2; Guilford, 1954) skaliert.

Für den Hauptversuch ausgewählt wurden die zwei „stärksten“ (A, B) und „schwächsten“ (C, D) Nazi-Texte:

Text A: (Hitler, Mein Kampf)

Man begreift nie, daß die Stärke einer politischen Partei keineswegs in einer möglichst großen und selbständigen Geistigkeit der einzelnen Mitglieder liegt, als vielmehr im disziplinierten Gehorsam, mit dem ihre Mitglieder der geistigen Führung Gefolgschaft leisten. Das Entscheidende ist die Führung selbst. Wenn zwei Truppenkörper miteinander kämpfen, wird nicht derjenige siegen, bei dem jeder einzelne die höchste strategische Ausbildung erhielt, sondern derjenige, der die überlegenste Führung und zugleich die disziplinierteste, blindgehorsamste, bestgedrillte Truppe hat.

Text B: (Kriek, Nationalsozialistische Erziehung)

Politik und Staat können naturgemäß in Kunst und Wissenschaft nicht selbst schöpferisch sein, aber sie können einer solchen Schöpfung die Bahn und den Boden bereiten. Ferner können und müssen sie, um der Sinn- und Gleichrichtung des ganzen Lebens willen, worüber sie verantwortlich zu wachen haben, über Kunst und Wissenschaft ein Richteramt üben: das Hemmende und Zerstörende ist zu unterdrücken, das dem Ziel förderliche mit Nachdruck zu pflegen und zu unterstützen.

Text C: (Hitler, Mein Kampf)

Ich bin der Überzeugung, daß der Mann sich im allgemeinen — Fälle ganz besonderer Begabung ausgenommen — nicht vor seinem dreißigsten Jahre in der Politik öffentlich betätigen soll. Er soll dies nicht, da ja bis in diese Zeit hinein zumeist erst die Bildung einer allgemeinen Plattform stattfindet. Auch der Dreißigjährige wird im Lauf seines Lebens noch vieles zu lernen haben, allein es wird dies nur eine Ergänzung und Ausfüllung des Rahmens sein, den die grundsätzlich angenommene Weltanschauung ihm vorlegt.

Text D: (Hitler, Mein Kampf)

Schon die Möglichkeit der Erhaltung eines gesunden Bauernstandes als Fundament der gesamten Nation kann niemals hoch genug eingeschätzt werden. Viele unserer heutigen Leiden sind nur die Folge des ungesunden Verhältnisses zwischen Land- und Stadtvolk. Ein fester Stock kleiner und mittlerer Bauern war noch zu allen Zeiten der beste Schutz gegen soziale Erkrankungen, wie wir sie heute besitzen. Dies ist aber auch die einzige Lösung, die eine Nation das tägliche Brot im inneren Kreislauf einer Wirtschaft finden läßt.

Es soll nicht behauptet werden, A und B stellten die „typischsten“, C und D die „untypischsten“ Nazi-Texte dar. Ihre Auswahl ergibt sich vielmehr aus der Skalierung von zehn ohnehin „gemäßigten“ Nazi-Texten, so daß alle vier Texte sicherlich ganz allgemeine faschistische bzw. autoritäre Denkweisen ausdrücken.

b) Hauptversuche

Versuchspersonen der Hauptversuche waren 85 angehende Sozialarbeiter; es handelte sich um drei Klassen einer Berufsschule für Sozialarbeiter in Westfalen, alle männlich, mit einem mittleren Alter von 24¹/₂ Jahren.

Zunächst wurden den Vpn zwei zusammengeheftete Bogen vorgelegt, deren erster einige Angaben zur Person verlangte. Auf dem zweiten Blatt sollte analog dem Vorversuch eine Beliebtheitsrangreihe der Politiker Heuss, Adenauer, Erhard, Brandt, Lübke, und Goebbels aufgestellt werden. Da „Hitler“ sich bereits als durchgängig „negativer“ Politiker erwiesen hatte, diente dieser Versuch nur zur Auswahl des „positiven“ Etiketts.

Sodann wurden den Vpn eine von *Lienert u. Leuchtmann* (1958) empfohlene Kurzform des Amthauer-Intelligenztests, bestehend aus den Untertests SE, AN, GE, vorgelegt.

Der eigentliche Hauptversuch, die Beurteilung der politischen Texte, vollzog sich nach folgendem Plan:

Die Vpn wurden unwissentlich in drei gleich große Gruppen aufgeteilt; alle bekamen gleichzeitig zwei verschiedene Texte vorgelegt. Jede Vp bekam einen „starken“ und einen „schwachen“ Nazi-Text, und zwar entweder A und C oder B und D.

Die Texte der Gruppe I waren mit einem *positiven* Etikett versehen; als solches fungierte der von jeder einzelnen Vp mit Rangplatz 1 bezeichnete Politiker. Diese Zuordnung war während der Ausführung des Intelligenztests unbemerkt vorgenommen worden.

Die Texte der Gruppe II waren mit einem *negativen* Etikett versehen; als „Verfasser“ fungierte hier ausnahmslos „Hitler“.

Gruppe III erhielt die politischen Texte *ohne* Etikett.

Alle Vpn sollten jeden der beiden Texte analog dem Vorversuch auf einer 7-Punkte-Skala nach dem Grad der persönlichen Zustimmung beurteilen. Außerdem wurde Platz für schriftliche Bemerkungen gelassen. Die Vpn der Gruppe III wurden anschließend um eine Schätzung des möglichen Verfassers gebeten.

Allen Vpn war mehrfach versichert worden, daß alle Auswertungen anonym bleiben würden und daß kein einziges Ergebnis irgendwelche Konsequenzen haben würde.

Die Auswahl der Versuchsgruppen erfolgte so, daß alle drei Gruppen auch in bezug auf ihre Beliebtheitsschätzungen homogen waren. Intelligenzhomogenität konnte jedoch nur angenommen werden. Die Darbietungsfolge „stark NS — schwach NS“ kam in jeder Gruppe gleich häufig vor wie die umgekehrte Reihenfolge. Den Versuchsplan gibt Tab. 1 wieder.

Tabelle 1

Versuchsplan

	Gruppe I (pos. Etikett)				Gruppe II (neg. Etikett)				Gruppe III (ohne Etikett)			
1. Text	A	C	B	D	A	C	B	D	A	C	B	D
2. Text	C	A	D	B	C	A	D	B	C	A	D	B

Aus den anfangs erwähnten Annahmen lassen sich folgende Hypothesen ableiten:

- H₁: Die Textbeurteilungen der Gruppen I, II und III unterscheiden sich insgesamt voneinander.
- H₂: Die unter H₁ erwarteten Unterschiede zeigen sich innerhalb jedes Textes (A, B, C, D) zwischen Vpn mit positivem und negativem Etikett (I und II).
- H₃: Bei den schwachen Nazi-Texten (C, D) ist die Meinungsänderung unter dem Einfluß des positiven Etiketts (geschätzt aus der unter H₂ erwarteten Differenz) größer als bei den starken d. h. eher eindeutigen Nazi-Texten.

Nimmt man bei weniger intelligenten Vpn eine größere allgemeine Beeinflußbarkeit an (wie dies Zillig schon 1928 bei Schulkindern fand) so läßt sich aufgrund der Intelligenz-Kurzprüfung erwarten:

- H₄: Der mittlere IQ von Vpn, die eher positive Urteile abgeben (< 4) ist geringer als bei Vpn mit eher negativen Urteilen (> 4).
- H_{4a}: Diese Differenz ist größer bei positivem als bei negativem Etikett.

Betrachtet man die zusätzlichen schriftlichen Bemerkungen der Vpn nach der Urteilsabgabe als Ausdruck eines Konfliktes, der sich aus der Unstimmigkeit von Text und Etikett ergibt, so läßt sich vermuten:

- H₅: Bei Texten mit positivem Etikett werden häufiger Bemerkungen auftreten als bei Texten mit negativem Etikett.
- H_{5a}: Dieses Verhältnis wird bei starken Nazi-Texten (A, B) größer sein als bei schwachen.

Unterscheidet man zwischen zustimmenden und ablehnenden Bemerkungen, so läßt sich wegen der Art der ausgewählten Texte erwarten:

- H₆: Es werden mehr ablehnende als zustimmende Bemerkungen auftreten.
- H_{6a}: Dies gilt für alle Texte und für jedes Etikett.

Bezeichnet man ferner eine Bemerkung als „konkordant“ bzw. „diskordant“, je nachdem ob das zugehörige Rating in die gleiche (zustimmende oder ablehnende) Richtung geht wie die verbale Bemerkung, so müßte man erwarten:

H₇: Das Verhältnis von diskordanten zu konkordanten Bemerkungen wird bei Texten mit positivem Etikett größer sein als bei solchen mit negativem Etikett.

Die statistische Prüfung der Hypothesen soll mit entsprechenden parametrischen Verfahren, u. U. unter Inanspruchnahme des Vorteils der Robustheit der einzelnen Methoden, erfolgen. Als Signifikanzniveau wird $\alpha = 5\%$ festgelegt, und es wird die Möglichkeit in Anspruch genommen, nicht signifikante Differenzen im Sinne von „Tendenzen“ zu interpretieren.

III. Ergebnisse

Die Ratings aller drei Stichproben und der Gesamtstichprobe sind in völlig homogener Weise zweigipflig verteilt. Nachdem die statistische Unabhängigkeit der nach dem Versuchsplan abhängigen Textbeurteilungen festgestellt werden konnte ($r_{AC} = 0,00$; $r_{BD} = 0,02$) wurden die von den 85 Vpn abgegebenen 170 Urteile wie nicht korrelierende Daten behandelt.

Tab. 2 gibt die wichtigsten deskriptiven Statistiken für die Textbeurteilungen unter den verschiedenen Bedingungen an.

Zwischen den Urteilen der Gruppen I, II und III zeigte sich ein gesicherter Unterschied:

Quelle	QuS	df	Varianz	F
Zw. Etiketts	45,18	2	22,59	4,28 ⁺
Innerhalb	882,07	167	5,28	
Total	927,25	169		

Damit wurde H₁ bestätigt. Als einzige Mittelwertsdifferenz war diejenige zwischen I und II (positives und negatives Etikett) signifikant ($M_I = 3,91$; $M_{II} = 5,10$; $t = 3,13$; $p < 0,01$). Unter dem Einfluß des „positiven“ Etiketts werden die gleichen politischen Texte erheblich günstiger beurteilt.

Entsprechende Unterschiede innerhalb jedes einzelnen der vier Texte zeigten sich nicht (H₂ verworfen); lediglich für Text B ergab sich ein gesicherter Unterschied ($M_{BI} = 3,50$; $M_{BII} = 5,35$; $t = 2,28$; $p < 0,05$). Die übrigen Differenzen liegen jedoch, wie Tab. 2 zeigt, sämtlich in der erwarteten Richtung.

Tabelle 2

Mittelwerte der Beurteilungen von „starken“ (A, B) und „schwachen“ (C, D) Nazi-Texten mit positivem (+), negativem (–) und ohne Etikett (o). Niedrige Werte bedeuten Zustimmung, hohe Werte bedeuten Ablehnung.

Text	M	s	N
A +	4,06	2,44	15
A –	5,53	1,95	15
A o	5,35	2,12	14
B +	3,50	2,03	14
B –	5,35	2,13	14
B o	5,15	2,13	13
C +	4,33	2,52	15
C –	4,60	2,25	15
C o	4,21	2,04	14
D +	3,72	1,98	14
D –	4,93	1,71	14
D o	4,70	2,02	13

Faßt man die „schwachen“ (C, D) und die „starken“ Nazi-Texte (A, B) zusammen, so ergibt sich zwischen den Bedingungen I und II ein erheblicher Unterschied in der erwarteten Richtung für die „starken“ Nazi-Texte ($M_{A, BI} = 3,80$; $M_{A, BII} = 5,45$; $t = 2,91$; $p < 0.01$), nicht aber für die „schwachen“ Nazi-Texte. Damit ist H_3 falsifiziert. Das Gegenteil von H_3 trifft zu: Eine eindeutige Meinungsänderung in der „suggerierten“ Richtung tritt nur bei dem eher eindeutigen Textmaterial auf.

Als Intelligenz-Kurzmaß diente der Median der Amthauer-SW der drei verwendeten Untertests; dieses Maß war bei der Gesamtstichprobe und bei allen drei Versuchsgruppen in völlig homogener Weise eingipflig und leicht rechtsschief verteilt. Vpn, die beiden Texten eher zustimmten, unterschieden sich jedoch hinsichtlich des Intelligenzmaßes nicht von solchen, die eher ablehnende Urteile abgaben. Damit ist H_4 falsifiziert. Die Prüfung von H_{4a} kann entfallen.

Schriftliche Bemerkungen wurden von den Vpn mit 72 von 170 in 42 % aller Fälle gegeben. Bei den Texten mit positivem Etikett traten in 39,4 %, bei denjenigen mit negativem Etikett in 53,4 % aller Fälle schriftliche Bemerkungen auf. Diese Differenz ist — geprüft mit dem

t-Test für Proportionen — nicht signifikant und liegt in der H_5 entgegengesetzten Richtung (H_5 falsifiziert). Bei Bedingung III (ohne Etikett) wurden mit 33,3 % signifikant weniger Bemerkungen abgegeben als bei Bedingung II (negatives Etikett). Die Prüfung von H_{5a} kann entfallen.

Es zeigte sich weiterhin, daß bei den „starken“ Nazi-Texten wesentlich mehr Bemerkungen abgegeben wurden (50,6 %) als bei den „schwachen“ Texten (34,0 %). Die Differenz der Anteile von schriftlichen Bemerkungen an den Beurteilungen ist auf dem 1 %-Niveau gesichert.

Die qualitative Einteilung der von den Vpn abgegebenen schriftlichen Bemerkungen in „ablehnende“ und „zustimmende“ wurde von zwei unabhängigen Beurteilern vorgenommen, die eine Übereinstimmung von $\Phi = 0.7$ erzielten. Zur Prüfung der in H_6 hypostasierten Unterschiede wurde die Einteilung desjenigen Beurteilers verwendet, der weniger nicht entscheidbare Fälle lieferte. Danach gab es mit 62 negativen und 9 positiven Bemerkungen ein mit dem Binomialtest sicheres Überwiegen negativer Anmerkungen (H_6 verifiziert). Dieses Ergebnis zeigte sich in ähnlich eindeutiger Weise unter allen drei Versuchsbedingungen und für die Texte A, B und D (H_{6a} verifiziert).

Das Verhältnis von „diskordanten“ zu „konkordanten“ Bemerkungen war bei Texten mit positivem Etikett (7 : 16) nicht signifikant größer als dasjenige bei Texten mit negativem Etikett (4 : 27); damit ist H_7 verworfen. Es zeigte sich jedoch eine Tendenz in der erwarteten Richtung; außerdem war das erstgenannte Verhältnis erheblich größer als dasjenige bei Texten ohne Etikett (1 : 17; $t = 2,40$; $p < 0.02$).

IV. Diskussion

In Übereinstimmung mit den zitierten älteren amerikanischen Experimenten hat sich auch in unserem Versuch eine Meinungsänderung durch Prestige-Suggestion ergeben. Wir verwendeten dabei keine Version eines Vorher-Nachher-Versuchsplans, bei dem zunächst eine Einstellungsmessung, dann die Darbietung des etikettierten Textes, und schließlich eine zweite Einstellungsmessung erfolgt. Wir bevorzugten einen einfachen Planversuch, der im einfachsten Falle auch durch einen Zweigruppenversuchsplan ersetzt werden könnte. Auf diese Weise war es unnötig, die Einstellung der Vpn gegenüber dem Inhalt des politischen Textes zweimal zu messen; dies hätte die für Messungswiederholungen im psychologischen Bereich charakteristischen Schwierigkeiten mit sich gebracht.

Das Ergebnis, daß sich eine meinungsändernde Wirkung des positiven Etiketts eher bei „starken“ als bei „schwachen“ Nazi-Texten zeigt, kann nur mit sehr großer Vorsicht verallgemeinernd interpretiert werden. Denn einmal ist es möglich, daß eine Skalierung der Nazi-Texte nach ihrer „Stärke“, wie sie im Vorversuch vorgenommen wurde, bei den Vpn des Hauptversuchs abweichend ausgefallen wäre. Zum anderen hängt die Interpretation dieses Ergebnisses davon ab, ob man „starke“ Texte als „eindeutig“, und „schwache“ als „mehrdeutig“ interpretieren will.

Bezeichnet man die Texte A und B als „eindeutiges Material“, so hat sich für diese Texte erwiesen, daß Meinungsänderungen durch Prestige-Suggestion auch bei „bedeutungskonstantem“ Material möglich sind. Die Verwerfung von H_3 spräche dann gegen die Vermutung, daß die Prestige-Suggestibilität mit der Mehrdeutigkeit des Materials anwachsen muß.

Einiges spricht dafür, daß unsere Skala „stark—schwach nazistisch“ sich nicht mit einer „Eindeutigkeit—Mehrdeutigkeit“-Skala deckt. Dies ergibt sich zunächst aus der von vornherein notwendigen Beschränkung auf „gemäßigte“ Texte: Dadurch, daß extreme Nazi-Formulierungen bei der Auswahl des Materials bereits ausgeschlossen wurden, erhalten auch die Texte A und B etwas „Mehrdeutiges“ in dem Sinne, daß entsprechende faschistische bzw. autoritäre Äußerungen auch von anderen als nazistischen Politikern stammen könnten.

Zum anderen könnte die absolute Homogenität der Rating-Urteile darauf hinweisen, daß die vier Texte als in ähnlicher Weise „eindeutig“ erlebt wurden. Es gibt aber — insbesondere aus schriftlichen Zusatzbemerkungen — Anzeichen dafür, daß Text B, der die Reglementierung von Kunst und Wissenschaft empfiehlt, von den Vpn des Hauptversuchs als „stärkster“ Nazi-Text angesehen wurde. Daß dieser Text — wie aus Kommentaren von Vpn hervorgeht — als besonders „aktuell“ bezeichnet werden kann, spricht ebenfalls für seine „Eindeutigkeit“. Die Annahme einer Abhängigkeit der Meinungsänderung von der Mehrdeutigkeit des Textmaterials ist jedenfalls nicht eindeutig bestätigt worden. „Meinungsänderung bei Bedeutungskonstanz“ ist demnach durchaus denkbar.

Das Überwiegen ablehnender Beurteilungen — am besten ersichtlich aus den Daten der Bedingung III (ohne Etikett); vgl. Tab. 2 — kennzeichnet die verwendeten Nazi-Texte als eher „negativen“ Meinungsgegenstand mit einer Position von $-1,3$ (A, B) bzw. $-0,5$ (C, D) auf einem Meinungskontinuum in der Art von Osgood und Tannenbaum (1955). Insofern bedeutet die verstärkte Zustim-

mung zu diesen Texten unter Bedingung I, daß die eindeutig positiven Etiketts „Heuss“, „Erhard“, „Adenauer“, „Brandt“) in der Lage sind, die „negative“ Wirkung des unetikettierten oder negativ (mit „Hitler“) etikettierten Textes aufzuheben und als Meinungsgegenstand zu neutralisieren. Dieser Sachverhalt wird noch deutlicher, wenn man Urteile < 4 als „Zustimmung“, und Urteile > 4 als „Ablehnung“ bezeichnet. Der Binomialtest zeigt dann lediglich unter Bedingung I kein signifikantes Abweichen von der Gleichverteilung „zustimmender“ und „ablehnender“ Urteile an. Verallgemeinert bedeuten diese Ergebnisse, daß Prestige-Suggestion nicht unbedingt ablehnende Meinungen in zustimmende verwandeln kann, daß sie aber ablehnende Haltungen ausgleichen und zu neutralen Urteilen führen kann.

Das Ausmaß der zu erzielenden Meinungsänderung bezüglich eines Nazi-Textes müßte sich — wenn die „Kongruenztheorie“ von O s g o o d zutrifft — nach der von O s g o o d u. T a n n e n b a u m (1955) angegebenen Beziehung

$$AC_{OJ_2} = \frac{d_{OJ_1}}{d_{OJ_1} + d_{OJ_2}} P_{OJ_2}$$

vorhersagen lassen, vorausgesetzt, die Positionen des Etiketts und des Nazi-Textes auf dem Meinungskontinuum (von -3 bis $+3$) sind bekannt. Setzen wir z. B. für Text B den Mittelwert der Urteile unter Bedingung III („ohne Etikett“) ein ($5,15$, nach O s g o o d also $-1,2$) und für das „positive Etikett“ den Wert $+2,0$, so ergibt sich unter Berücksichtigung des Maßes für den „Druck in Richtung Kongruenz“ ($P_{OJ_2} = 3,2$) das folgende Ausmaß von zu erwartender Meinungsänderung:

$$AC_{OJ_2} = \frac{(+2)}{(+2) + (-1,2)} (3,2) = +2,0$$

Der Gegenstand der Meinungsänderung (Text B) würde danach auf dem Meinungskontinuum um $2,0$ Skaleneinheiten nach oben rücken d. h. auf $+0,8$. Dieser „neue“ Wert entspricht — obgleich er noch in Bezug auf „Glaubwürdigkeit“ korrigiert werden müßte (O s g o o d u. T a n n e n b a u m, 1955, S. 47 ff.) — in etwa dem empirisch d. h. unter Bedingung I („positives Etikett“) gefundenen Mittelwert von $+0,5$. Da in unserem Fall die Einstellung gegenüber OJ_1 , dem beliebtesten Politiker, nicht auf einem O s g o o d schen Meinungskontinuum gemessen wurde, ließe sich spekulieren, daß eine Position von $+1,7$ für „Heuss“ usw. zutreffender gewesen wäre als der Schätzwert von $+2,0$. In jedem Falle erscheint es denkbar, bei weiteren Experimenten zur Prestige-Suggestion im Rahmen der O s g o o d schen Kongruenztheorie Erwartungen zu formulieren und diese Theo-

rie für den Fall der Prestige-Suggestion bei politischen Texten weiterhin zu überprüfen.

Betrachtet man nur die 42 % abgegebenen schriftlichen Bemerkungen der Vpn, so ergibt sich das gegenüber der Rating-Beurteilung geradezu umgekehrte Bild: Die „positive“ Etikettierung bleibt wirkungslos, und Prestige-Suggestion zeigt sich nur bei den Texten C und D.

Da bei diesem Experiment den Vpn deutlich gemacht worden war, daß allein ihr Rating entscheidend sei, und da hier keiner der üblichen Einwände gegen Rating-Methoden ins Gewicht fallen dürfte, bietet sich als Erklärung an, daß die schriftlichen Bemerkungen eine kompensatorische und/oder konfliktmindernde Funktion übernehmen. Geht man davon aus, daß eine schriftliche Zusatzbemerkung als Ausdruck eines irgendwie gearteten erlebten Konflikts anzusehen sei, so läßt sich die Art des schriftlichen Kommentars im Sinne einer Spannungsminderung interpretieren. So berichtet z. B. M c K e a c h i e (1963) über eine Reihe von Arbeiten, bei denen sich die Möglichkeit, bei einer Beantwortung in starren Kategorien (z. B. multiple choice) auch noch schriftliche Kommentare abgeben, als angstreduzierendes und — bei Examensklausuren — leistungssteigerndes Mittel erwies. Zwar konnte im Falle unserer H₇ die Nullhypothese nicht zurückgewiesen werden, doch lassen die aufweisbaren deutlichen Tendenzen erkennen, daß der „Konflikt“, der bei der Beurteilung eines „an sich“ negativ zu beurteilenden Nazi-Textes mit einem „positiven“ Etikett auftreten müßte, in gewissem Maße durch eine „diskordante“ Äußerung verarbeitet wurde.

Als Beispiele für solche „diskordanten“ Zusatzbemerkungen von Vpn, die unter Bedingung I zu urteilen hatten, seien genannt:

Vp X zu Text B (Etikett: „Heuss“; Rating: 3):

„Ein Richteramt über Kunst und Wissenschaft steht dem Staat nur bedingt zu. Wie konnte Heuss so etwas sagen? Mangelnde soziologische Kenntnisse!“

Vp Y zu Text B (Etikett: „Heuss“; Rating: 3):

„Ein Richteramt über Kunst zu üben, halte ich für gefährlich und entwicklungs-hemmend (NS-Zeit). Die ersten drei Zeilen volle Zustimmung.“

Vp Z zu Text A (Etikett: „Adenauer“; Rating: 3):

„Diese Zustimmung bedeutet keineswegs ein Gutheißen dessen, was hinter dem Text steht, sondern höchstens ein Zugeben von Tatsachen in dem Text, die abzustreiten ich für albern halte.“

Es erscheint aussichtsreich, Versuche zu planen, in denen die Verwendung von Vpn-Kommentaren als möglicher Indikator erlebter Unstimmigkeit systematisiert und die entsprechenden Zusammenhänge näher untersucht werden. Für die Wirkweise der (positiven) Prestige-

Suggestion bedeutet das diskutierte Nebenergebnis jedenfalls, daß die in welcher Richtung auch immer geänderten Meinungen nicht ohne weiteres als absolut oder gar persistent verändert gelten können.

Für die Prüfung des Einflusses der Intelligenz auf das Ausmaß der Prestige-Suggestibilität war die untersuchte Stichprobe nicht gut geeignet. H_4 und H_{4a} ließen sich sinnvoll wohl nur an verschiedenen Intelligenz- und Schulbildungspopulationen prüfen. Weitere Persönlichkeitsvariable, insbesondere „allgemeine Suggestibilität“ im Sinne von Eysencks „sekundärer Suggestibilität“, wurden hier noch nicht berücksichtigt; dies soll einer späteren Arbeit vorbehalten bleiben.

Z u s a m m e n f a s s u n g

Die Möglichkeit, politische Urteile durch Manipulation des „Etiketts“ des Beurteilungsgegenstandes zu verändern, wurde überprüft und bestätigt. 85 angehende Sozialarbeiter hatten „stärkere“ und „schwächere“ Nazi-Texte zu beurteilen, die teilweise angeblich von einem sehr beliebten Politiker stammten. Normalerweise abgelehnte Texte wurden dadurch auf dem Meinungskontinuum mindestens bis zur Indifferenzlage angehoben. Entgegen der Erwartung war die Urteilsänderung bei „stärkeren“ Nazi-Texten besonders groß. Bei einem Teil der Vpn zeigte sich die Tendenz, das positivere Urteil durch zusätzliche Kommentare einzuschränken.

S u m m a r y

By labelling political statements in different ways judgments could be altered. 85 social workers were asked to give their opinion to different Nazi texts labelled under the name of well known and liked German statesmen. By this most of the judgments could be changed from the rejection area to at least indifference, particularly judgments about "strong" Nazi texts; this was against the hypothesis. In their verbal comments the subjects showed the tendency to alleviate their judgments.

R e s u m é

85 sujets étaient priés de juger des textes politiques dont les «étiquettes» étaient changées par l'expérimentateur. On se servit de textes «nazis» d'une intensité différente et, comme étiquette, de noms d'hommes politiques actuels très populaires. Le changement obtenu des jugements de l'expression négative jusqu'à l'indifférence était particulièrement accentué quand il s'agissait de textes nazis très intenses. Un certain nombre de juges eut la tendance de restreindre le jugement positif initial par des commentaires supplémentaires.

Literatur

- Asch, S. E.: Studies in the principles of judgments and attitudes: II. Determination of judgments by groups and by ego standards. *J. soc. Psychol.* 12, 1940, 433—465.
- Asch, S. E.: The doctrine of suggestion, prestige, and imitation in Social Psychology. *Psychol. Rev.* 55, 1948, 260—276.
- Asch, S. E.: *Social Psychology*. Englewood Cliffs, N. J. (1957)⁴.
- Birch, H. G.: The effect of socially disapproved labeling upon a well structured attitude. *J. abn. soc. Psychol.* 40, 1945, 301—310.
- Eysenck, H. J.: Suggestibility and hypnosis — an experimental analysis, *Proc. Roy. Soc. Med.* 36, 1943, 349—354.
- Eysenck, H. J. and Furneaux, W. D.: Primary and secondary suggestibility: an experimental and statistical study. *J. exp. Psychol.* 35, 1945, 485—503.
- Festinger, L. and Aronson, E.: The arousal and reduction of dissonance in social contexts. In: Cartwright, D. and Zander, A. (eds): *Group dynamics: Research and theory*. Evanston, Ill. 1960; 214—231.
- Fröhlich, W. D.: *Forschungsstatistik*. Bonn (1965)⁴.
- Guilford, J. P.: *Psychometric methods*. N. Y. — Toronto — London (1954)².
- Hovland, C. I., Harvey, O. J. and Sherif, M.: Assimilation and contrast effects in reactions to communication and attitude change. *J. abn. soc. Psychol.* 55, 1957, 244—252.
- Hovland, C. I., Janis, I. L. and Kelley, H. H.: *Communication and persuasion*. New Haven (1959)³.
- Lewis, H. B.: *The study of attitudes*. Psychologists League J., N. Y. 1938.
- Lewis, H. B.: Studies in the principles of judgments and attitudes: II. The influence of political attitudes on the organization and stability of judgments. *J. soc. Psychol.*, 14, 1940, 121—146.
- Lewis, H. B.: Studies in the principles of judgments and attitudes: IV: The operation of „Prestige Suggestion“. *J. soc. Psychol.* 14, 1941, 522—532.
- Lienert, G. A.: *Verteilungsfreie Methoden in der Biostatistik*. Meisenheim 1962.
- Lienert, G. A. u. Leuchtmann, T.: Die Möglichkeiten einer Kurzform des IST-Amthauer. *Psychol. Prax.* 2, 1958, 177—182.
- McKeachie, W. J.: Research on teaching at the college and university level. In: Gage, N. L. (ed): *Handbook of research on teaching*. Chicago 1963; 1118—1172.
- Newcomb, T. M.: *Sozialpsychologie*. Meisenheim 1959.
- Osgood, C. E. and Tannenbaum, P. H.: The principle of congruity in the prediction of attitude change. *Psychol. Rev.* 62, 1955, 42—55.
- Stukat, K. G.: *Suggestibility. A factorial and experimental analysis*. Stockholm 1958.
- Zillig, M.: Einstellung und Aussage. *Z. Psychol.* 106, 1928, 58—106.

Anschrift der Verfasser:

Dr. H. D. Schmidt und Christiane Schmerl

Psychologisches Institut

53 Bonn