

B I E L E F E L D E R A R B E I T E N Z U R S O Z I A L -
P S Y C H O L O G I E

Nr.18

(April 1977)

Hans Dieter Mummendey,
Bernd Schiebel und Uwe Troske:

Experimentelle Untersuchung der Stabi-
lität instrumentell-aggressiven Verhal-
tens bei Meßwiederholungen in verschie-
denen Zeitabständen

Zusammenfassung:

Zur Untersuchung der Zeitstabilität instrumentell-aggressiven Verhaltens und des Effektes eines aggressiven Modells wurde ein Experiment mit 59 nichtstudentischen erwachsenen Personen (Lehrlingen) geplant und ausgeführt. Es zeigte sich eine sehr große Konstanz der Aggressionsindikatoren bei Meßwiederholungen nach zwei Wochen und zwei Monaten bei den unbehandelten Kontrollgruppen (Stabilitätskoeffizienten um .86) sowie eine leichte Abnahme bei den Versuchsgruppen, die bei der ersten Messung ein aggressives Modell beobachtet hatten. Zu einem deutlichen Modell-Effekt kam es nicht; dies wurde auf das Fehlen eines aggressionsreduzierenden nicht-aggressiven Modells zurückgeführt.

Problemstellung

Die Möglichkeit, instrumentell-aggressives Verhalten im psychologischen Experiment zu erfassen, ist vielfach belegt worden. Um wichtige Fragestellungen - z.B. solche nach den Bedingungen des Modellerns aggressiven Verhaltens - zu beantworten, scheinen experimentelle Anordnungen gegenüber feldexperimentellen und nicht-experimentellen Verfahren den Vorzug größerer Reliabilität der Messung und größerer interner Validität des Versuchsplans zu genießen. Dennoch sind wesentliche Gütekriterien der Verhaltensmessung bzw. wichtige Voraussetzungen der internen Validität entsprechender Versuchspläne selten empirisch untersucht worden, sondern sie wurden und werden von den Experimentatoren meist stillschweigend als gegeben unterstellt. Damit liefern letztere wiederum berechnete oder aber unnötige Angriffsflächen für Kritik am psychologischen Experiment.

In der vorliegenden Arbeit soll eine solche, als relevant für die innere und äußere Gültigkeit von Experimenten zu Bedingungen aggressiven Verhaltens anzusehende Variable experimentell untersucht werden: die Stabilität von im Experiment gezeigtem instrumentell-aggressivem Verhalten bei Meßwiederholungen in verschiedenen Zeitabständen.

Die Frage, ob die Stabilität (oder Veränderung) individuellen aggressiven Verhaltens bei einer Meßwiederholung nach einer bestimmten Zeit auf mangelnde (oder hinreichende) Reliabilität des Meßmittels hinweist oder ob sie als Indikator für eine erfolgte (oder ausgebliebene) Veränderung des individuellen Ausprägungsgrades aggressiven Verhaltens in der Zeit - oder auch nur einer veränderten (oder konstanten) Reaktion oder Reaktionseinstellung gegenüber dem Experiment selbst - anzusehen ist, kann nicht leicht beantwortet werden. Meßgenauigkeit im Sinne der Retest-Reliabilität und Merkmalskonstanz im Sinne der Stabilität von Persönlichkeitsmerkmalen sind operational nur schwer zu unterscheiden. (Von HEISE, 1962, stammt ein Versuch, Reliabilität und Stabilität im genannten Sinne voneinander zu trennen, doch bedarf es zu dessen pfadanalytischem Verfahren der Erfassung des kritischen Verhaltens zu mindestens drei Meßzeitpunkten. In Experimenten ist es jedoch gewöhnlich

bereits sehr schwierig, einmalige Meßwiederholungen, die aussagekräftig sein sollen, zu organisieren.) So soll die vorliegende Arbeit auch zu der meßtheoretischen Frage "Reliabilität vs. Stabilität" nicht viel beitragen, sondern es erscheint bereits als ein gegenüber der allgemeinen Praxis experimenteller Forschung beträchtlicher Gewinn, überhaupt Meßwiederholungen zu verschiedenen Zeitpunkten mit den dafür möglichen experimentellen Kontrollen vorzunehmen.

Aus mehrjähriger Arbeit mit einem erstmals von SCHMIDT-MUMMENDEY (1972) beschriebenen Aggressionssimulator wurde in letzter Zeit eine Versuchsanordnung entwickelt, mit der es gelang, wichtige Ergebnisse der experimentellen Erforschung aggressiven Verhaltens auch an westdeutschen, nicht-studentischen erwachsenen Versuchspersonen zu replizieren:

- zum "Waffen-Effekt", d.h. der Förderung oder Hemmung aggressiven Verhaltens beim Vorhandensein aggressiver Stimuli (SCHMIDT & SCHMIDT-MUMMENDEY 1974) und
- zur Wirkung von Modell-Eigenschaften auf das aggressive Verhalten erwachsener Personen (SCHMIDT, SCHMIDT-MUMMENDEY, SCHRÖER & STALLBERG 1976a; SCHMIDT, SCHMIDT-MUMMENDEY, GOERTZ, MAUS, SCHRÖER & TROSKE 1976b).

Außerdem erwies sich die Erfassung des aggressiven Verhaltens mit dieser Versuchsanordnung als weitgehend stabil

- gegenüber einer Substitution der verwendeten Strafreize (Geldstrafen statt Elektroschocks)
- gegenüber Substitutionen des Versuchsleiters (der V1 spielt eine wichtige Rolle, da er die Vp auffordert, ihren Gegenspieler für nicht-nachgiebiges Verhalten zu bestrafen)
- gegenüber Merkmalen der Versuchs-Situation (mehrfacher Wechsel des Labors durch Umzug)
- gegenüber Vpn-Populationsunterschieden (SCHMIDT, SCHIEBEL & TROSKE 1976c).

Allerdings wurde bisher noch nicht untersucht, ob und ggf. wie sich die Meßwerte der abhängigen Variablen in diesen Experimenten (Straf-Stärke und Straf-Dauer) bei Meßwiederholungen verändern. Zwar läßt der Aufweis

großer Stabilität gegenüber Populations- und Situationsunterschieden darauf schließen, daß experimentelle Situation und Instruktion in ausreichendem Maße und in unverfänglicher Weise zur Gewinnung konsistenter Daten bei unterschiedlichen Personen führen werden, doch steht der Nachweis intraindivideller Stabilität/Variabilität noch aus. Es könnte ja sein, daß es sich bei dem gemessenen Verhalten um ein Muster einzigartiger Reaktionen auf eine vollständig neue und überraschende experimentelle Situation handelt, so daß die Ergebnisse von Meßwiederholungen nach längerer Zeit keinerlei sinnvollen Bezug zur jeweils ersten Messung dieses Verhaltens erkennen lassen.

Versuchsplanung

Zur Untersuchung der zeitlichen Stabilität ist mindestens eine Meßwiederholung pro Person erforderlich. Für einmalige Wiederholungsmessungen erscheint der notwendige Kontrollgruppenaufwand (vgl. CAMPBELL & STANLEY 1963) auch gerade noch vertretbar. Mehrfache Meßwiederholungen verbieten sich - stärker noch als einmalige - außerdem wegen einer Reihe von störenden Lerneffekten (z.B. Reaktionen auf behaltene frühere Reaktionen, zunehmende Transparenz der Versuchsanordnung etc.) und wegen der bereits bei einmaligen Wiederholungen großen experimentellen "Sterblichkeit" bzw. Drop-out-Quote. Daher erschien als einfachster, gerade noch mit Aussicht auf Erfolg praktisch realisierbarer Versuchsplan eine Anordnung sinnvoll, bei der eine Stichprobe von Personen nach Zufall in zwei Gruppen aufgeteilt wird:

Die erste Gruppe wiederholt den Versuch nach einer bestimmten Zeit a, und die zweite Gruppe wiederholt den Versuch nach einer hiervon abweichenden, längeren Zeitspanne b. Als Zeitspannen sollten a) wenige Wochen, und b) wenige Monate gewählt werden. In der Praxis des Experiments waren a = 2 Wochen, und b = 2 Monate:

Gruppe 1:	1.Messung	--	(2 Wochen)	--	2.Messung
Gruppe 2:	1.Messung	--	(2 Monate)	--	2.Messung

Zwar stand die Frage nach der Reliabilität bzw. Stabilität instrumentell-aggressiven Verhaltens im Experiment im Mittelpunkt unseres Interesses,

doch ergab sich zugleich die Möglichkeit, zumindest einen der in den vergangenen, genannten Untersuchungen nachgewiesenen Modell-Effekte aggressiven Verhaltens in Bezug auf seine Persistenz zu untersuchen. Wir entschieden uns für die Untersuchung des am eindeutigsten gesicherten Modell-Effektes, nämlich Modell-Aggression.

In den bereits erwähnten Arbeiten hatte sich durchgängig ergeben, daß Personen, denen ein kurzer Modell-Film (Video-Film in Fernseh-Monitor) vorgeführt worden war, auf welchem eine Person ihren Gegenspieler jeweils intensiv und dauerhaft strafte (mit Elektroschocks oder Geldentzug), selbst im anschließenden Spiel am Aggressionssimulator erheblich intensiver und dauerhafter strafte als Personen, die ein ausgesprochen nicht-aggressives Modell beobachtet hatten. Da wir davon ausgingen, daß dieser Modell-Effekt auch im Vergleich einer Personengruppe, die ein aggressives Modell sieht, mit einer unbehandelten Kontrollgruppe zutage tritt, sollten die oben aufgeführten Gruppen 1 und 2 nochmals in jeweils zwei, in Bezug auf die genannte Modell-Bedingung unterschiedene Versuchs- und Kontrollgruppen wie folgt aufgespalten werden:

Gruppe 1:	Versuchsgruppe 1: Film -- 1.Messung -- (2 Wochen) -- 2.Messung
	Kontrollgruppe 1: ----- 1.Messung -- (2 Wochen) -- 2.Messung
Gruppe 2:	Versuchsgruppe 2: Film -- 1.Messung -- (2 Monate) -- 2.Messung
	Kontrollgruppe 2: ----- 1.Messung -- (2 Monate) -- 2.Messung

Das Ausmaß der Stabilität/Variabilität des Modell-Effekts wird sich dann im Vergleich des Unterschiedes der Aggressionswerte von Versuchs- und Kontrollgruppe zwischen erster Messung, zweiter Messung nach zwei Wochen, und zweiter Messung nach zwei Monaten ergeben.

Das Ausmaß der Stabilität/Variabilität aggressiven Verhaltens (ohne Berücksichtigung des Modell-Einflusses) wird sich im Vergleich der Aggressionswerte der Kontrollgruppen bei erster Messung, zweiter Messung nach zwei Wochen, und zweiter Messung nach zwei Monaten darstellen lassen.

Ein bedeutsamer Aspekt der Versuchsplanung ist - ebenso wie bei unseren bisherigen Experimenten - die sorgfältige Auswahl der Versuchspersonen.

Diese müssen mindestens den folgenden Ansprüchen genügen:

- keine Studenten oder ähnliche, nicht-repräsentative Personengruppen
- keine Vorkenntnisse über Versuche dieser Art, sei es aus eigener Anschauung oder aus den Massenmedien
- möglichst erwachsene Versuchspersonen.

Ausführung der Versuche

Die Experimente fanden von November 1976 bis Januar 1977 in den psychologischen Experimentierräumen der Universität Bielefeld (Ebene T1) statt. An ihnen nahmen insgesamt 59 Personen als Vpn teil.

Als Vpn kamen männliche Berufsschüler aus der ca. 50 km entfernten Stadt Paderborn (55 Berufsschüler, vier Berufsfachschüler) mit einem Durchschnittsalter von 16 Jahren (jüngste Vp 15, älteste 18 Jahre) zu den Versuchen. Sie wurden von ihrem Religionslehrer zu einer Universitätsbesichtigung mit Teilnahme an einem Experiment angeworben. Im Vordergrund des Interesses stand eher die Universität, weniger das Experiment. Hin- und Rücktransport wurden von instruierten Studenten mit PKW vorgenommen. Dabei ergab sich die Möglichkeit, informell Äußerungen über das Experiment zu erhalten. Nach den Aussagen der Fahrer und des Lehrers, der noch eine Nachbefragung vornahm, lehnte keine Person das Experiment vor oder nach seiner Ausführung ab, auch ergaben sich keine Hinweise auf Transparenz des Versuchs oder Hypotheseneinsicht. Alle Vpn nahmen an, die ihnen jeweils dargebotenen Versuchsbedingungen seien allen anderen Vpn in gleicher Weise appliziert worden. Niemandem fiel also auf, daß einigen Personen ein Video-Film vorgespielt worden war, anderen dagegen nicht. Die genannten Beobachtungen waren für uns eine wichtige Voraussetzung für die Möglichkeit, den Versuch später zu wiederholen.

Alle Versuchsteilnehmer sagten zu, nach einer bestimmten Zeit (die ihnen erst später vom Lehrer mitgeteilt wurde), nochmals zu einem anderen Versuch in die Universität zu kommen. Bei näherem Nachfragen wurde etwas von "Ausfüllen von Fragebogen" gesagt. Jeder Vp wurden als Vergütung für die Teilnahme an beiden Versuchen fünfzehn Mark angeboten, von denen fünf als Fahrtkostenbeitrag für die chauffierenden Studenten einbehalten werden

sollten; der Betrag wurde erst nach der jeweils zweiten Untersuchung ausgezahlt. Übereinstimmend berichteten Fahrer und Lehrer, daß der finanzielle Anreiz bei der Anwerbung der Versuchsteilnehmer keine bedeutende Rolle zu spielen schien.

Große Sorgfalt wurde auch auf die Aufteilung der Personen auf verschiedene Termine und Versuchsbedingungen verwandt. Der Lehrer warb insgesamt vier ansonsten vergleichbare Klassen an, die an vier unterschiedlichen Wochentagen Berufsschulunterricht hatten und einander sonst nicht kannten: eine Donnerstags-, Freitags-, Montags- und Mittwochs-Gruppe. Die ersten beiden Gruppen kamen am 25./26. November und kamen zum zweiten Mal am 9./10. Dezember (zwei Wochen Abstand). Die beiden restlichen Gruppen kamen am 24./29. November zum ersten Mal, und am 26./31. Januar zum zweiten Mal (zwei Monate Abstand).

Bei der jeweils ersten Untersuchung wurden allen Personen abwechselnd die Bedingungen "Film" oder "kein Film" appliziert, um sequentielle Effekte zu vermeiden. Dennoch schien es an einem der vier Tage so, als herrschte unter jenen Personen, die am späteren Nachmittag untersucht wurden, im Vergleich zu denen, die am frühen Nachmittag an der Reihe waren, eine gereizte Stimmung, denn an diesem Nachmittag versagte der von weiteren Studenten vorgenommene Besichtigungs-Service. Zur Kontrolle wurden später Korrelationen zwischen Personen-Reihenfolge und Aggressionsindikator berechnet.

Bei der jeweils ersten Untersuchung kam die Person in den Raum mit der "Aggressionsmaschine" (beschrieben bei SCHMIDT-MUMMENDEY 1972, 77 ff.) und erhielt vom VI die Instruktion (beschrieben bei SCHMIDT et al. 1976b, 9 ff.), in welcher als Strafreize der Entzug von Vpn-Geld verwendet wird. Dabei wurde entweder der Film "Hoch-aggressives Modell" (beschrieben bei SCHMIDT et al. 1976a, 11 ff.) gezeigt, oder nicht. Die Vp wetteiferte dann in der üblichen Weise mit ihrem im Nebenraum sitzenden Gegenspieler darum, erster am Ziel zu sein und bemühte sich, den Gegner durch unterschiedlich intensive und dauerhafte Strafen (Tastendruck) zu vertreiben, wenn dieser in der "Einbahnstraße" den Weg nicht freimachte. Wenn die Person mit dem Versuch fertig war, ging sie zu ihrer Besichtigungs-Gruppe oder in die

nahegelegene Cafeteria. Die jeweils noch nicht untersuchten Personen wurden zur gleichen Zeit von einem anderen Studenten betreut.

Bei der jeweils zweiten Untersuchung absolvierte jede Vp tatsächlich eine Paper-Pencil-Untersuchung, indem sie Bogen zur Selbsteinschätzung und zum Locus of Control ausfüllte. Ganz nebenbei wurde sie jedoch gebeten, vorher kurz noch einmal "einen Versuch wie neulich" zu machen. Hierbei erfolgte - wie bereits geschildert - dann nur das Spiel am Aggressionssimulator, und nicht mehr die Instruktion mit dem Modellfilm. Diese jeweils zweiten Experimente waren übrigens für den Versuchsleiter "Blindversuche" in Bezug auf die Versuchs-Kontrollgruppen-Bedingung, da dem V1 die Liste der Mitglieder beider Gruppen nicht mehr vorlag.

Ergebnisse

Versuchspersonen. Von den planmäßig 60 Versuchsteilnehmern kamen 59 zur ersten Untersuchung; 30 wurden der Versuchsgruppe, 29 der Kontrollgruppe zugeteilt. Je 15 von ihnen wurden zur zweiten Untersuchung nach zwei Wochen gebeten, die restlichen zur zweiten Untersuchung nach zwei Monaten. Von ihnen erschienen nach zwei Wochen insgesamt 21 (11 der Versuchs-, 10 der Kontrollgruppe), und nach zwei Monaten 15 (6 der Versuchs-, 9 der Kontrollgruppe). Der stärkere Abfall der Zahl der Wiederholer nach zwei Monaten schien weniger etwas mit mangelndem Interesse als vielmehr mit zur gleichen Zeit stattfindenden Prüfungen zu tun zu haben. Da jedoch nicht auszuschließen war, daß systematische Selektionseffekte wirksam gewesen sein könnten, wurden die folgenden Vergleiche der Indikatoren aggressiven Verhaltens, sofern sie sich auf Gruppenmittelwerte beziehen, jeweils für die (geringere) Zahl derjenigen Personen angestellt, die beim ersten und beim zweiten Mal zum Versuch erschienen waren.

Zeitstabilität der Meßwerte. Die Meßwerte instrumentell-aggressiven Verhaltens im vorliegenden Experiment sind wie üblich

- die durchschnittliche Höhe der in allen 27 Durchgängen des Versuchs von der Vp gedrückten Straftasten, d.h. die Stärke der Bestrafung
- die durchschnittliche Dauer, mit der die Tasten bei diesen Gelegenheiten

gedrückt wurden.

(Beide Indikatoren, Stärke und Dauer, wurden nur auf der Basis jener Fälle bzw. Durchgänge berechnet, in denen tatsächlich Tasten gedrückt wurden; Verweigerungen wurden nicht berücksichtigt. Bei allen folgenden Ergebnissen zeigen sich keinerlei Unterschiede zwischen den solcherart bereinigten Maßen einerseits, und den rohen Durchschnittswerten andererseits. Eine Inspektion der Verweigerungshäufigkeiten ergab, daß a) Verweigerungen insgesamt äußerst selten waren, und b) Verweigerungen erwartungsgemäß bei der Kontrollgruppe etwas häufiger als bei der Versuchsgruppe auftraten. Bei den 30 Personen und jeweils 27 Bestrafungssituationen der Gesamt-Versuchsgruppe bei der ersten Messung kamen nur insgesamt 9 Verweigerungen vor, bei der um eine Person kleineren Kontrollgruppe nur 16; dies entspricht einem Anteil von rund 1% Verweigerungen bei der Versuchsgruppe, und rund 2% Verweigerungen bei der Kontrollgruppe. Ein ähnliches Bild ergab sich bei den zweiten Messungen.)

Die folgenden Tabellen 1 und 2 geben für alle Personengruppen, die zur jeweils 2. Messung (nach 2 Wochen oder 2 Monaten) erschienen, Mittelwerte, Standardabweichungen und t-Werte für den Mittelwertsvergleich abhängiger Messungen (1. vs. 2. Messung) wieder.

	<u>1. Messung</u>		<u>2. Messung</u> (2 Wochen)		t	df	p
	\bar{X}	s	\bar{X}	s			
<u>Versuchs-</u> <u>gruppe</u> (n = 11)	8.04	2.21	7.51	2.51	1.05	10	>0.05
<u>Kontroll-</u> <u>gruppe</u> (n = 10)	5.64	2.67	5.64	2.74	0.00	9	>0.05
	<u>1. Messung</u>		<u>2. Messung</u> (2 Monate)				
<u>Versuchs-</u> <u>gruppe</u> (n=6)	6.28	2.47	4.06	2.01	1.44	5	>0.05
<u>Kontroll-</u> <u>gruppe</u> (n=9)	7.56	1.75	7.55	2.02	0.02	8	>0.05

Tabelle 1: Durchschnittliche Aggressionswerte (Stärke) von Versuchs- und Kontrollgruppen bei erster und zweiter Messung

	<u>1.Messung</u>		<u>2.Messung (2 Wochen)</u>		t	df	p
	\bar{X}	s	\bar{X}	s			
<u>Versuchs-</u> <u>gruppe</u> (n=11)	22.33	11.53	25.10	28.53	0.36	10	>0.05
<u>Kontroll-</u> <u>gruppe</u> (n=10)	11.37	8.00	11.26	8.48	0.09	9	>0.05
	<u>1.Messung</u>		<u>2.Messung (2 Monate)</u>				
<u>Versuchs-</u> <u>gruppe</u> (n=6)	16.74	13.04	12.18	7.44	1.17	5	>0.05
<u>Kontroll-</u> <u>gruppe</u>	31.83	32.56	23.89	22.14	1.33	8	>0.05

Tabelle 2: Durchschnittliche Aggressionswerte (Dauer) von Versuchs- und Kontrollgruppen bei erster und zweiter Messung

In Bezug auf Straf-Stärke und Straf-Dauer ergeben sich keine signifikanten Unterschiede zwischen erster und zweiter Messung nach 14 Tagen bei Versuchs- oder Kontrollgruppe, während sich bei zweiter Messung nach zwei Monaten nicht-signifikante Tendenzen zur Abnahme aggressiven Verhaltens ergeben; letztere sollten jedoch angesichts der sehr kleinen Stichproben nicht überbewertet werden.

Geben die (im wesentlichen unveränderten) Mittelwerte der Stichproben Aufschluß über eine relative Konstanz der Aggressionswerte in Versuchs- und Kontrollgruppe, so zeigen die Korrelationen zwischen erster und zweiter Messung, daß eine intraindividuelle Konstanz aggressiven Verhaltens nur bei den (unbehandelten) Personen der Kontrollgruppe auftritt (Tabelle 3).

Tabelle 3:

<u>Retest-Korrelation</u>	<u>Stärke</u>	<u>Dauer</u>
(2 Wochen) Versuchsgruppe (n=11)	-0.10	0.43
(2 Wochen) Kontrollgruppe (n=10)	0.85 ⁺⁺	0.89 ⁺⁺
(2 Monate) Versuchsgruppe (n=6)	0.34	0.69
(2 Monate) Kontrollgruppe (n=9)	0.78 ⁺⁺	0.85 ⁺⁺

Retest-Korrelation

(ohne Berücksichtigung der Länge der Zwischenzeit)

	<u>Stärke</u>	<u>Dauer</u>
Versuchsgruppe (n = 17)	0.24	0.45
Kontrollgruppe (n = 19)	0.85 ⁺⁺	0.87 ⁺⁺

(⁺⁺ = Unterschreitung des 0.01-Niveaus)

Eine erhebliche intraindividuelle Verhaltenskonstanz zeigt sich also auch, wenn man von der Größe des Zeitraumes zwischen beiden Messungen absieht. Interpretierbar im Sinne von Meßwertstabilität sind lediglich die Aggressionswerte der unbeeinflussten Kontrollgruppe.

In den Abbildungen 1 bis 3 sind noch einmal die Mittelwerte aggressiven Verhaltens für die abhängige Variable Stärke der Versuchsgruppen (schwarz) und Kontrollgruppen (weiß) zu den drei hier unterschiedenen Zeitpunkten graphisch dargestellt. In Abb.1 (Meßwiederholung nach 14 Tagen) scheint sich ebenso wie in Abb.2 (Meßwiederholung nach 2 Monaten) ein Abfall der Versuchsgruppen-Werte (Nachlassen des Modell-Einflusses) und eine Konstanz

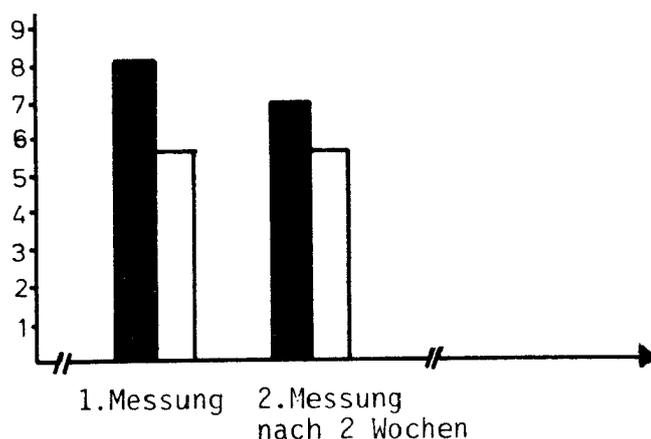


Abb.1: Aggressionswerte (Stärke) von Versuchsgruppe (aggressives Modell; schwarz) und Kontrollgruppe bei Meßwiederholung nach 2 Wochen (ohne Modell)

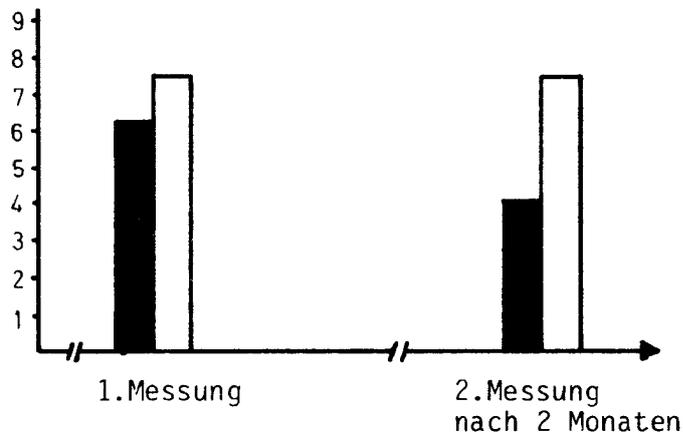


Abb.2: Aggressionswerte (Stärke) von Versuchsgruppe (aggressives Modell; schwarz) und Kontrollgruppe bei Meßwiederholung nach 2 Monaten (ohne Modell)

der Aggressionswerte der Kontrollgruppe zu zeigen. In Abb.3 sind die durchschnittlichen Aggressionswerte für alle drei Zeitpunkte nebeneinandergestellt; als erste Messungen sind hier - abweichend von den Berechnungen in Tab.1 und 2 - die Werte sämtlicher Personen, die zur ersten Messung kamen (also einschließlich derer, die nicht mehr zur zweiten

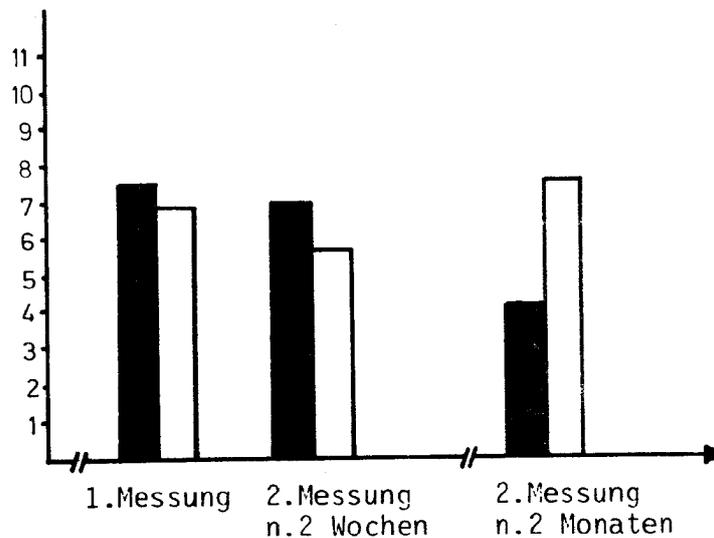


Abb.3: Aggressionswerte (Stärke) der Versuchsgruppe (schwarz) und der Kontrollgruppe (weiß); Meßwiederholungen ohne Modelldarbietung

Untersuchung erschienen) verarbeitet. Die mittleren Stärke-Werte betragen hier für die 30 Vpn der Versuchsgruppe 7.51 und für die 29 Vpn der Kontrollgruppe 6.97. In Abb.3 beruht demgemäß der erste Block auf 59, der zweite auf 51, und der dritte auf 15 Meßwerten. Hier zeigt sich ein nach zwei Wochen schwacher, nach zwei Monaten starker Abfall der Aggressionswerte der Versuchsgruppe (die bei der ersten Untersuchung ein sehr aggressives Modell beobachtet hatte) sowie ein relatives Gleichbleiben der Kontrollgruppenwerte.

Stabilität des Modell-Effekts. Da die Versuchsgruppen bei ihrer jeweils ersten Untersuchung ein hoch-aggressives Modell gesehen hatten, bei der zweiten Untersuchung aber nicht mehr, läßt sich die Persistenz eines möglichen Modell-Effektes durch einen Vergleich der Meßwertdifferenzen zwischen Versuchs- und Kontrollgruppe bei der ersten und zweiten Messung abschätzen. Die Tabellen 4 und 5 geben für alle Personen, die zwei Untersuchungen absolvierten, Mittelwerte, Standardabweichungen und t-Werte für den Mittelwertsvergleich unabhängiger Stichproben (Versuchs- vs. Kontrollgruppe) wieder.

	<u>Versuchsgruppe</u> (n=11)		<u>Kontrollgruppe</u> (n=10)		t	df	p
	X	s	X	s			
<u>1.Messung</u>	8.04	2.21	5.64	2.67	2.26	19	< 0.05
<u>2.Messung</u> <u>n.2 Wochen</u>	6.95	2.21	5.64	2.74	1.52	19	> 0.05
	<u>Versuchsgruppe</u> (n=6)		<u>Kontrollgruppe</u> (n=9)				
<u>1.Messung</u>	6.28	2.47	7.56	1.75	-1.18	13	> 0.05
<u>2.Messung</u> <u>n.2 Monaten</u>	4.06	2.01	7.55	2.02	-3.29	13	< 0.01

Tab.4: Durchschnittliche Aggressionswerte (Stärke) bei erster und zweiter Messung; Mittelwertsdifferenzen zwischen Versuchs- und Kontrollgruppen (Modell-Effekt)

	<u>Versuchsgruppe</u> (n=11)		<u>Kontrollgruppe</u> (n=10)		t	df	p
	\bar{x}	s	\bar{x}	s			
<u>1.Messung</u>	22.33	11.53	11.37	8.00	2.51	19	<0.05
<u>2.Messung</u> <u>n.2 Wochen</u>	25.10	28.53	11.26	8.48	1.54	9	>0.05
	<u>Versuchsgruppe</u> (n=6)		<u>Kontrollgruppe</u> (n=9)				
<u>1.Messung</u>	16.74	13.04	31.83	32.56	-1.24	7	>0.05
<u>2.Messung</u> <u>n.2 Monaten</u>	12.18	7.44	23.89	22.14	-1.47	7	>0.05

Tab.5: Durchschnittliche Aggressionswerte (Dauer) bei erster und zweiter Messung; Mittelwertsdifferenzen zwischen Versuchs- und Kontrollgruppen (Modell-Effekt)

Hier zeigen sich keineswegs eindeutige Ergebnisse. Bei den Wiederholern nach zwei Wochen ergibt sich ein signifikanter Modelleffekt bei der ersten Messung, der sich bei der zweiten Messung nur noch als Tendenz wiederfindet. Dies gilt sowohl für die Stärke- wie für die Dauer-Messung. Bei den Personen, die mit zwei Monaten Abstand untersucht wurden, zeigt sich bei der ersten Messung kein Modell-Effekt, dagegen ergibt sich bei der zweiten Messung ein umgekehrter Effekt für das Stärke-Maß (d.h. die Vpn der Kontrollgruppe verhalten sich aggressiver als die der Versuchsgruppe), und eine entsprechende Tendenz zeigt sich auch beim Dauer-Maß.

Die Wirkung der beiden untersuchten Faktoren "Modell-Aggression" und "Meßwiederholung" und ihre Wechselwirkung läßt sich in einem varianz-analytischen 2x2-Design mit Meßwiederholungen in einem Faktor prüfen (vgl. WINER 1971, 287). Diese Prüfung wurde getrennt für Meßwiederholungen nach zwei Wochen und nach zwei Monaten vorgenommen. Tabelle 6 gibt die Ergebnisse dieser Analyse für die abhängige Variable "Stärke", Tabelle 7 die Resultate für die abhängige Variable "Dauer" wieder.

Abhängige Variable: Stärke; Meßwiederholung nach 2 Wochen

	QUS	df	MQUS	F	p
Zw Vpn	196.61	19			
A (Modell)	32.04	1	32.04	3.51	<10%
Vpn innerh.Gruppen	164.57	18	9.14		
Innerh.Vpn	60.16	20			
B (Meßwiederholung)	1.01	1	1.01	----	
A x B	1.42	1	1.42	----	
B x Vpn innerh.Gruppen	57.73	18	3.21		

Abhängige Variable: Stärke; Meßwiederholung nach 2 Monaten

	QUS	df	MQUS	F	p
Zw Vpn	90.54	11			
A (Modell)	21.41	1	21.41	3.10	10%
Vpn innerh.Gruppen	69.03	10	6.90		
Innerh.Vpn	35.10	12			
B (Meßwiederholung)	6.56	1	6.56	3.40	<10%
A x B	9.27	1	9.27	4.80	<10%
B x Vpn innerh.Gruppen	19.27	10	1.93		

Tabelle 6: Varianzanalyse der Faktoren "Modell-Aggression" und "Meßwiederholung" (Aggressionsindikator: Straf-Stärke)

Abhängige Variable: Dauer; Meßwiederholung nach 2 Wochen

	QUS	df	MQUS	F	p
Zw Vpn	9118.37	19			
A (Modell)	1533.89	1	1533.89	3.64	<10%
Vpn innerh.Gruppen	7584.48	18	421.36		
Innerh.Vpn	3401.78	20			
B (Meßwiederholung)	225.39	1	225.39	1.32	
A x B	92.26	1	92.26	----	
B x Vpn innerh.Gruppen	3084.13	18	171.34		

Abhängige Variable: Dauer; Meßwiederholung nach 2 Monaten

	QUS	df	MQUS	F	p
Zw Vpn	8317.09	11			
A (Modell)	1358.72	1	1358.72	1.95	
Vpn innerh.Gruppen	6958.37	10	695,84		
Innerh.Vpn	513,59	12			
B (Meßwiederholung)	115.81	1	115.81	2.91	
A x B	0.16	1	0.16	----	
B x Vpn innerh.Gruppen	397.62	10	39.76		

Tabelle 7: Varianzanalyse der Faktoren "Modell-Aggression" und "Meßwiederholung" (Aggressionsindikator: Straf-Dauer)

Keiner der geprüften F-Werte erreicht das 5%-Niveau, doch lassen sich einige Tendenzen (p kleiner als 10%) erkennen: Faktor A (Modell-Aggression) zeigt sich tendenziell wirksam (mit Ausnahme beim Dauer-Maß bei den Wiederholern nach zwei Monaten). Faktor B (Meßwiederholung) wirkt tendenziell nur beim Stärke-Maß und bei der Gruppe der Wiederholer nach zwei Monaten; hier ergibt sich auch eine entsprechende Wechselwirkung zwischen A und B, die aber nicht interpretiert werden soll, da sie nur das 10%-Niveau erreicht.

Sequentielle Effekte. Zwei Reihenfolge-Effekte, die möglicherweise zur weiteren Aufhellung der untersuchten Veränderungen beitragen könnten, wurden zusätzlich mit deskriptiven Verfahren untersucht: der Verlauf des Indikators "Stärke" in den verschiedenen Phasen des Experiments, und der mögliche Zusammenhang zwischen der Reihenfolge der untersuchten Personen und deren aggressivem Verhalten (Stärke-Maß).

Stellt man den Verlauf der Straf-Intensität während des 27 Strafsituationen umfassenden Versuchs getrennt für die Durchgänge 1-4, 5-8 und 9-12 dar, also jene Gruppen von experimentellen Durchgängen, in denen der Gegenspieler der Vp jeweils einmal, zweimal, und viermal anzeigt, daß er den Weg nicht freigeben will (vgl. SCHMIDT et al. 1976c, 9 ff.), so ergibt sich das in den Abbildungen 4 und 5 erkennbare Bild.

Aus diesen Darstellungen geht hervor, daß die Versuchsgruppen (durchgezogene Linien) in fast allen Fällen die für unsere Versuchsanordnung charakteristische Verlaufsform zeigen: zögernder Anstieg oder sogar Abfall aggressiven Verhaltens von Phase I zu Phase II, dann schnellerer Anstieg von Phase II zu Phase III, so daß insgesamt ein Anstieg aggressiven Verhaltens von Phase I zu Phase III, also über den gesamten Versuch hinweg zu erkennen ist.

Die Kontrollgruppen, die keinen Film mit einer aggressiven Person gesehen hatten, zeigen den Knick in der Mitte des Versuchs jedoch offensichtlich nicht; bei ihnen erscheint der Anstieg von I nach II in den meisten Fällen ausgeprägter als bei den Versuchsgruppen. Deutlich ist das hohe Aggressionsniveau der Kontrollgruppe der Meßwiederholer nach zwei Monaten zu erkennen.

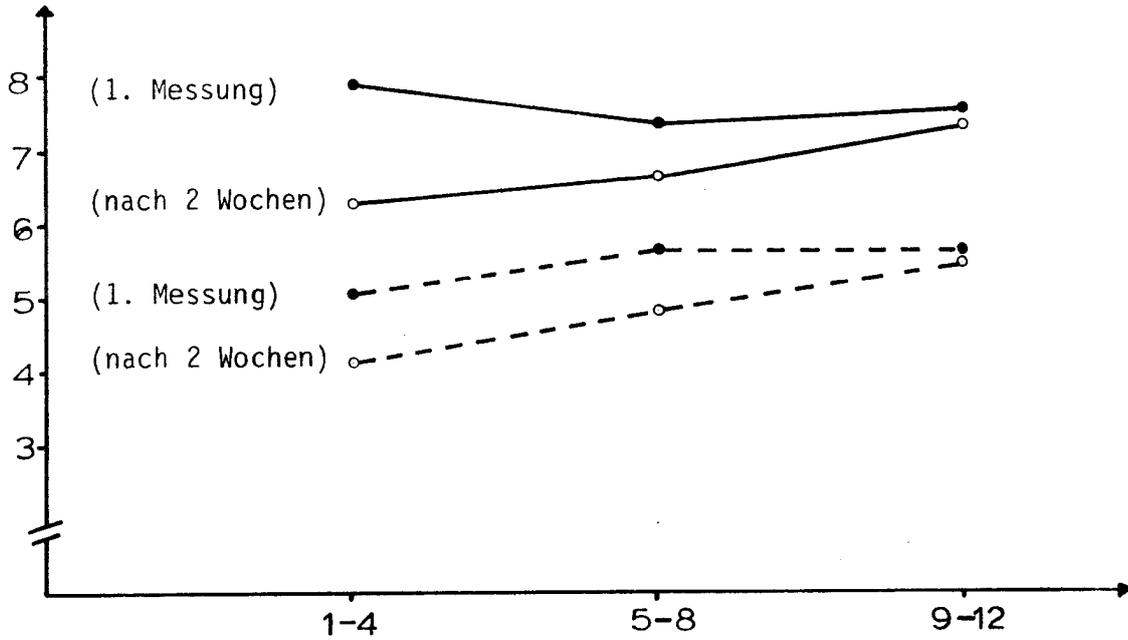


Abb.4: Verläufe des Aggressionsindikators "Stärke" in den drei Phasen des Versuchs, in denen sich der Gegenspieler einmal, zweimal, und viermal widersetzt (Versuchsgruppen durchgezogene, Kontrollgruppen gestrichelte Linien) Versuchswiederholung nach zwei Wochen

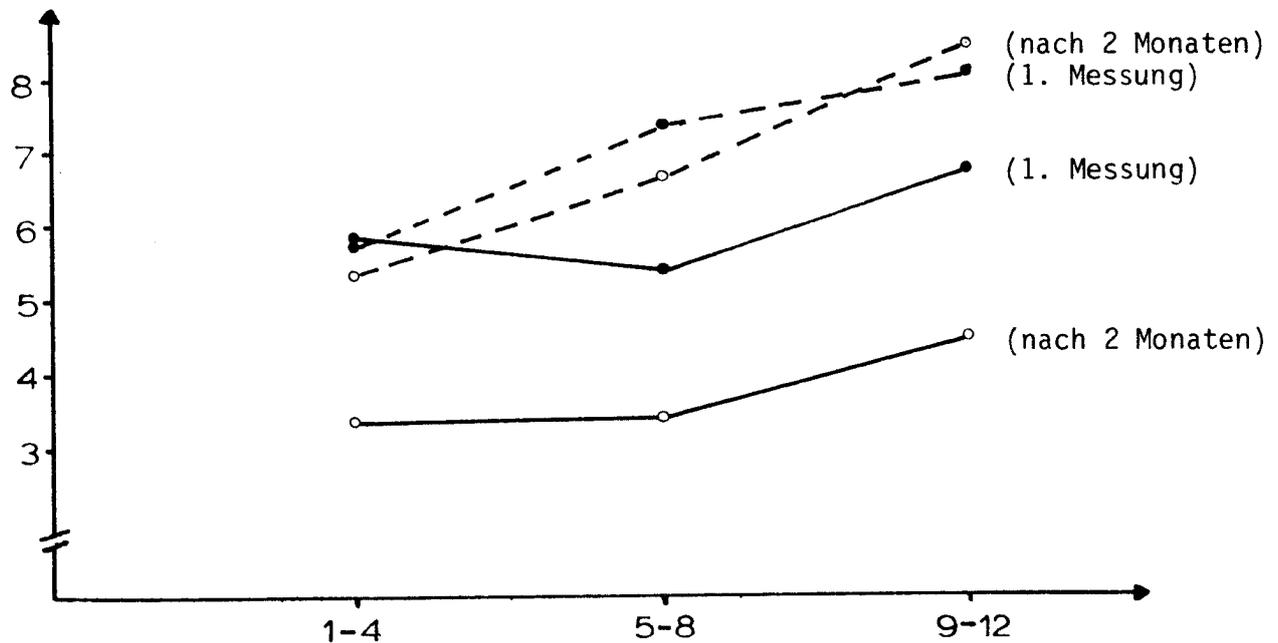


Abb.5: Verläufe des Aggressionsindikators "Stärke" in den drei Phasen des Versuchs, in denen sich der Gegenspieler einmal, zweimal und viermal widersetzt; Meßwiederholung nach zwei Monaten

Da sich bei einer der Gruppen von Vpn (Meßwiederholer nach zwei Wochen), in der Mitglieder von Versuchs- und Kontrollgruppe gemischt waren, tageszeitlich bedingte Störeffekte (s.o.S.7) ergeben konnten, wurden bei der ersten und zweiten Messung dieser Gruppe getrennt bei Versuchs- und Kontrollgruppe Rangkorrelationen zwischen der Reihenfolge der Untersuchung und der Straf-Stärke berechnet. Bei geringen Stichprobengrößen (zwischen sechs und neun) ergaben sich lediglich bei der Kontrollgruppe interpretierbare Reihenfolge-Effekte, die jedoch zu den zwei verschiedenen Zeitpunkten gegenläufig sind (0.88 bei erster, -0.86 bei zweiter Messung) und daher schwer interpretierbar sein dürften. Manchmal schien bei dieser Gruppe aggressives Verhalten also umso stärker, manchmal umso schwächer zu sein, je später die Untersuchung stattfand.

Korrelation der abhängigen Variablen. Die Produkt-Moment-Korrelation zwischen Straf-Stärke und -Dauer beträgt in diesem Experiment bei allen 59 Personen, die an der ersten Untersuchung teilnahmen, 0.43, bei allen 36 Personen, die an zweiten Messungen teilnahmen, 0.49.

Zusammenfassung der Ergebnisse und Diskussion

In einer relativ schwierig zu planenden und auszuführenden experimentellen Untersuchung mit Meßwiederholungen bei nicht-studentischen Versuchspersonen (Lehrlingen) konnten Daten aggressiven Verhaltens zu drei verschiedenen Zeitpunkten erhoben werden: bei 21 Personen (von ursprünglich 30) in einer Wiederholung nach zwei Wochen, und bei 15 Personen (von ursprünglich 29) nach zwei Monaten.

Die Ergebnisse erscheinen in zweierlei Hinsicht überraschend: Einmal zeigt sich eine erhebliche, in diesem Ausmaß kaum zu erwartende Konstanz aggressiven Verhaltens (Straf-Stärke und Dauer am Aggressionssimulator) bei den Vpn der Kontrollgruppen, d.h. nicht-signifikante Differenzen zwischen den Gruppenmittelwerten wiederholter Messungen sowie Retest-Korrelationen um 0.86. Zum anderen zeigt sich kein ausgeprägter Modell-Effekt, d.h. der Unterschied zwischen einer Versuchsgruppe, die vor der ersten Messung einen Film mit einem hoch-aggressiven Modell gesehen hatte, und der unbehandelten

Kontrollgruppe war bereits bei der ersten Messung nur bei der Hälfte der Personen interpretierbar. Bei der anderen Hälfte übertreffen die Aggressionswerte der Kontrollgruppe nach zwei Monaten sogar fast diejenigen der Versuchsgruppe.

Angesichts der niedrigen Personenzahl, die für einen exakten Vergleich unabhängiger und abhängiger Stichproben herangezogen werden können, sollen hier nur diese groben, ins Auge fallenden Verhältnisse interpretiert werden. Die in den Abbildungen 1 bis 3 dargestellten Tendenzen entsprechen im übrigen weitgehend den Erwartungen: relative Konstanz der Kontrollgruppen-Werte, relativer Abfall der Versuchsgruppen-Werte.

Zunächst scheint erwiesen, daß die Indikatoren instrumentell-aggressiven Verhaltens bei der vorliegenden Versuchsanordnung eine erstaunliche Zeitstabilität aufweisen. Sie ist auch unter Berücksichtigung der Zahl der beteiligten Personen nach acht Wochen kaum geringer als nach zwei Wochen. Bei allen 19 Personen der wiederholt gemessenen Kontrollgruppe werden 74% Varianz der zweiten Messung durch die erste Messung determiniert. Die Retest-Stabilität der experimentell gewonnenen Werte erreicht damit annähernd das für Leistungstests, mindestens aber das für Persönlichkeitstests erwartete Maß. Gegenüber einer möglichen Auffassung, spezifisches aggressives Verhalten wie dasjenige im Rahmen einer ganz speziellen experimentellen Umgebung sei weitgehend situationsabhängig auch im zeitlichen Sinne, jede Reaktion auf eine solche Situation sei also gewissermaßen einmalig, wurden hier sowohl prinzipielle Wiederholbarkeit als auch Verhaltensstabilität über längere Zeit hinweg nachgewiesen.

Im Gegensatz zu bisherigen Ergebnissen und unseren Erwartungen scheint zu stehen, daß im vorliegenden Experiment kein deutlicher Modell-Effekt, d.h. kein interpretierbarer Einfluß der im Rahmen der Instruktion beobachteten, aggressiven Modell-Person zutage tritt. Auf diese Weise kann die vorliegende Untersuchung auch nicht viel zur Frage der Stabilität eines solchen Modell-Effekts beitragen. Eine Erklärung für das Ausbleiben der Modell-Wirkung liegt jedoch auf der Hand. Der Effekt der Modell-Aggression wird in Experimenten dieser Art aus den Differenzen der Gruppenmittelwerte von Versuchs- und Kontrollgruppe erschlossen. In allen früheren Experimenten verglichen wir dabei eine Gruppe, deren

Mitglieder ein ausgesprochen nicht-aggressiv reagierendes Modell sahen. Diese zweite, gegenüber der ersteren gegensätzliche Versuchsbedingung war bei der vorliegenden Untersuchung Ökonomie-Erwägungen zum Opfer gefallen, da wir annahmen, der Modell-Effekt "Modell-Aggression" sei kräftig genug, um sich auch im Vergleich einer Versuchsgruppe ("hoch-aggressives Modell") mit einer unbehandelten Kontrollgruppe ("kein Modell") zu zeigen. Die Ergebnisse belehren uns nun eines Besseren und bestätigen theoretische Überlegungen und experimentelle Befunde z.B. von BANDURA & HUSTON (1961) und BERKOWITZ (1965), denen zufolge nicht-aggressive Modelle angeregtes aggressives Verhalten vermindern können. Das durch die Instruktion nahegelegte aggressive Verhalten könnte demnach durch ein ausgesprochen nicht-aggressives Modell nachhaltig gedämpft werden. Bei BARON & KEPNER (1970) schockten z.B. Personen, denen man ein Modell gezeigt hatte, das nur niedrige Elektroschocks an seinen Gegenspieler verteilte, ihr Gegenüber anschließend in geringerem Maße als Personen, die kein Modell beobachtet hatten.

In unseren bisherigen Experimenten stellt sich der resultierende Modell-Effekt demnach offensichtlich als Ergebnis des Zusammenwirkens mindestens zweier Effekte dar: eines aggressionsfördernden und eines aggressionshemmenden bzw. -mindernden Modelleffektes. Der fördernde Effekt, vermittelt durch die Applikation der Versuchsbedingung "Hoch-aggressives Modell", reicht allein nicht immer (z.B. nicht bei allen Personen einer Stichprobe, nicht bei kleineren Stichproben etc.) aus, um deutlich stärker aggressives Verhalten im Vergleich zu nicht-modelliertem aggressivem Verhalten zu gewährleisten - zu stark dürften die in der Versuchsanordnung einschließlich der Instruktion liegenden Anreize zur Vergabe von Strafen an den Gegenspieler im Verlauf des Experiments sein.

Durch einen Mangel des Versuchsdesigns wird damit abermals die große Bedeutung sichtbar, die nicht-aggressiven Modellen bei der Reduktion offen aggressiven Verhaltens der hier untersuchten Art zuzukommen scheint.

Literatur

- BANDURA, A. & HUSTON, A.C. 1961. Identification as a process of incidental learning. Journal of Abnormal and Social Psychology, 63, 311-318.
- BARON, R.A. & KEPNER, C.R. 1970. Model's behavior and attraction toward the model as determinants of adult aggressive behavior. Journal of Personality and Social Psychology, 14, 335-344.
- BERKOWITZ, L. 1965. Some aspects of observed aggression. Journal of Personality and Social Psychology, 2, 359-365.
- CAMPBELL, D.T. & STANLEY, J.C. 1963. Experimental and quasi-experimental designs for research. In: GAGE, N.L. (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally, 171-246.
- HEISE, D.R. 1969. Separating reliability and stability in test-retest correlation. In: BLALOCK, H.M. (Ed.) Causal Models in the Social Sciences. New York: Aldine, 348-363.
- SCHMIDT, H.D. & SCHMIDT-MUMMENDEY, A. 1974. Waffen als aggressionsbahnende Hinweisreize: eine kritische Betrachtung experimenteller Ergebnisse. Zeitschrift für Sozialpsychologie, 5, 201-218.
- SCHMIDT, H.D., SCHMIDT-MUMMENDEY, A., SCHRÖER, F.-W. & STALLBERG, R. 1976a. Der Einfluß aggressiven Modellverhaltens, Status des Modells und Selbsteinschätzung auf offen aggressives Verhalten. Bielefelder Arbeiten zur Sozialpsychologie, Nr.2 (Februar 1976).
- SCHMIDT, H.D., SCHMIDT-MUMMENDEY, A., GOERTZ, U., MAUS, G., SCHRÖER, F.-W. & TROSKE, U. 1976b. Der Einfluß von Modell-Aggression und Modell-Status auf das Erteilen von Geldstrafen als instrumentell-aggressives Verhalten. Bielefelder Arbeiten zur Sozialpsychologie, Nr.3 (März 1976).
- SCHMIDT, H.D., SCHIEBEL, B. & TROSKE, U. 1976c. Untersuchung der Stabilität von Modell-Effekten aggressiven Verhaltens gegenüber Populations- und Situationsunterschieden. Bielefelder Arbeiten zur Sozialpsychologie, Nr.10 (August 1976).
- SCHMIDT-MUMMENDEY, A. 1972. Bedingungen aggressiven Verhaltens. Bern und Stuttgart: Huber.
- WINER, B.J. 1971. Statistical principles in experimental design. New York: Wiley, 2nd ed.

BIELEFELDER ARBEITEN ZUR SOZIALPSYCHOLOGIE

Bis April 1977 erschienen:

- Nr.14 (1/77): Hans Dieter Mummendey, Rosemarie Mielke, Günter Maus und Bernhard Hesener: Untersuchungen mit einem mehrdimensionalen Selbsteinschätzungsverfahren
- Nr.15 (2/77): Rosemarie Mielke: Zur Erfassung von Umweltmerkmalen im Ausbildungsbereich
- Nr.16 (2/77): Manfred Bornwasser: Die Konsistenzbeziehung zwischen Einstellungen und offenem Verhalten
- Nr.17 (3/77): Hans Dieter Mummendey: Einstellungen (sets) bei der Erforschung der Beziehung zwischen Einstellungen (attitudes) und offenem Verhalten
- Nr.18 (4/77): Hans Dieter Mummendey, Bernd Schiebel und Uwe Troske: Experimentelle Untersuchung der Stabilität instrumentell-aggressiven Verhaltens bei Meßwiederholungen in verschiedenen Zeitabständen

BIELEFELDER ARBEITEN ZUR SOZIALPSYCHOLOGIE

1976 erschienen:

- Nr.1 (2/76): Hans Dieter Schmidt: Die Förderung nationalistischer Einstellungen durch internationale Sportwettkämpfe in sozial-psychologischer Sicht
- Nr.2 (2/76): Hans Dieter Schmidt, Amélie Schmidt-Mummendey, Friedrich-Wilhelm Schröer und Rüdiger Stallberg: Der Einfluß aggressiven Modellverhaltens, Status des Modells und Selbsteinschätzung auf offen aggressives Verhalten
- Nr.3 (3/76): Hans Dieter Schmidt, Amélie Schmidt-Mummendey, Ulrich Goertz, Günter Maus, Friedrich-Wilhelm Schröer und Uwe Troske: Der Einfluß von Modell-Aggression und Modell-Status auf das Erteilen von Geldstrafen als instrumentell-aggressives Verhalten
- Nr.4 (4/76): Manfred Bornewasser: Naturwissenschaftliche und verhaltenstheoretische Orientierungen in der Sozialpsychologie (I)
- Nr.5 (4/76): Rosemarie Mielke und Hans Dieter Schmidt: Zur Psychologie politischen Verhaltens - eine empirische Analyse von Verhaltensmustern beim Kampf um Gleichberechtigung der Frau
- Nr.6 (4/76): Hans Dieter Schmidt: Selbstwahrnehmung und Beeinflussung
- Nr.7 (4/76): Lothar P. Schardt: Wem nützt die Eignungsdiagnostik? Zum konservativen Charakter der klassischen Testtheorie und ihrer Anwendung in der betrieblichen Praxis
- Nr.8 (5/76): Lothar P. Schardt: Ansätze zu einer arbeitsorientierten Eignungsdiagnostik
- Nr.9 (7/76): Manfred Bornewasser, Fritz Köster, Rosemarie Mielke und Cornelia Pawelcik: Erstellung und erste Überprüfung eines Fragebogens zur instrumentellen Arbeitsorientierung bei Jugendlichen
- Nr.10 (8/76): Hans Dieter Schmidt, Bernd Schiebel und Uwe Troske: Untersuchung der Stabilität von Modell-Effekten aggressiven Verhaltens gegenüber Populations- und Situationsunterschieden
- Nr.11 (9/76): Hans Dieter Schmidt: Selbstwahrnehmung als interpersonelle Wahrnehmung: Hypothesen
- Nr.12 (10/76): Hans Dieter Schmidt, Lothar P. Schardt und Rosemarie Mielke: Ein- und mehrdimensionale Skalierung der gewerkschaftlichen Kampfbereitschaft von Belegschaften
- Nr.13 (Herbst 1976): Hans Dieter Schmidt: Auf dem Wege zu einer sozialen Psychophysiologie des akuten Vampirismus (vergriffen)