

Einige Probleme der Erfassung aggressiven Verhaltens im psychologischen Labor

HANS DIETER MUMMENDEY*

1. Einige Gründe für die Untersuchung aggressiven Verhaltens im Experiment

Im Rahmen psychologischer Forschung, die ihre Aufgabe in der Beschreibung, Erklärung und Vorhersage menschlichen Verhaltens erblickt, wurde aggressives Verhalten in der überwiegenden Mehrzahl der empirischen Studien experimentell erfaßt und untersucht. Wenngleich das Bemühen um feldexperimentelles Arbeiten und der dabei erzielte Nutzen mittlerweile als beachtlich anzusehen ist (vgl. PARKE, BERKOWITZ, LEYENS, WEST & SEBASTIAN, 1977), dominieren insgesamt doch eindeutig laborexperimentelle Untersuchungen (vgl. BARON, 1977; ZILLMANN, 1979; A. MUMMENDEY, 1981).

Die wissenschaftlichen Vorzüge der experimentellen Methode scheinen zwar unbestritten zu sein, können jedoch offensichtlich angesichts sophistischer oder plumper, professioneller oder nur modischer Experimentalkritik nicht oft genug genannt werden: Beschränkung auf wenige, meßbare Verhaltensmerkmale; Möglichkeit systematischer Variation unabhängiger Variablen und der Kontrolle antezedenter, konsequenter, störender usw. Bedingungen; Ökonomie; grundsätzlich bessere Kontrolle ethisch problematischer Verfahrensweisen, wie sie bei jeder wissenschaftlichen Beschäftigung mit Verhaltensweisen, die Personen tatsächlich oder intentional zu schädigen geeignet sind, auftreten.

Den Vorteilen der Erfassung und Untersuchung aggressiven Verhaltens im klassisch-psychologischen Laboratoriumsexperiment stehen als mittlerweile gut bekannt etablierte Nachteile zumindest gegenüber: Möglichkeit der externen Invalidierung des

* Der Verfasser dankt Heinz-Gerd BOLTEN, Manfred BORNEWASSER, Ulrich GOERTZ, Rosemarie MIELKE, Christian G. NENTWIG, Bernd SCHIEBEL, Friedrich-Wilhelm SCHRÖER, Jürgen SCHULTZ-GAMBARD, Rüdiger STALLBERG und Uwe TROSKE für ihre Mitarbeit, insbesondere Amélie MUMMENDEY für wertvolle Hinweise.

Experiments durch Verhaltensmanipulationen seitens der Versuchsperson (Vp) aufgrund ihrer vermeintlichen Kognition von «Anforderungsmerkmalen» (demand characteristics) des Experiments, stellvertretend für alle möglichen «Vpn-Effekte»; prinzipielle Nichtberücksichtigung größerer sozialer Kontexte, die Einflüsse auf das gemessene Verhalten ausüben mögen.

Bei der Abwägung dieser und einer Reihe sicherlich weiterer zu berücksichtigender Gesichtspunkte mag die Entscheidung der meisten Forscher zugunsten des laborexperimentellen «Spatzes», und nicht der letztlich doch stets oberhalb der Dachrinne sitzenbleibenden «Tauben» ausgegangen sein, weil in dem ständigen Konflikt zwischen interner und externer Validität logischerweise zunächst für erstere Partei zu ergreifen ist.

2. Die Erfassung aggressiven Verhaltens mit Aggressionssimulatoren

Unter den verschiedenen klassifizierbaren Methoden der Aggressionsmessung im psychologischen Labor dominiert bisher diejenige der Erfassung des Ausmasses einer vermeintlichen Schädigung eines Interaktionspartners im Rahmen einer standardisierten Interaktion. Man mag (wie z.B. BARON, 1977) die Verfahren danach unterscheiden, ob es sich um verbale Attacken, unschädliche oder schädliche physische Angriffe auf ein Gegenüber handelt – in jedem Falle sind die eine experimentelle Situation vornehmlich strukturierenden und standardisierenden Instruktions-Bedingungen von der Art, daß die vermeintliche Schädigung einer anderen Person als Reaktion auf die Darbietung bestimmter Stimuli möglich ist, d.h. in abgestufter Weise oder absolut erfolgen kann oder muß (sofern die Vp nicht eine aggressive Reaktion vollständig verweigert). Es soll dabei an dieser Stelle nicht diskutiert werden, ob oder inwiefern für die einschlägigen Reaktionen das Attribut «aggressiv» angemessen oder erforderlich ist, da hierzu von weiteren Autoren berechtigterweise immer wieder Diskussionsbeiträge geleistet werden.

Die Art der vermeintlichen oder tatsächlichen Schädigung einer anderen Person im Experiment ist unterschiedlich und hängt beispielsweise davon ab, ob frustrierende Ereignisse in die Analyse einbezogen werden sollen oder nicht. Z.B. konstruierten ULRICH & FAVELL eine Versuchsanordnung (*Abbildung 1*), bei der ein Kind

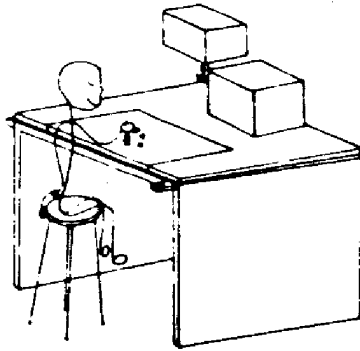


Abbildung 1 Aggressionssimulator von ULRICH & FAVELL

eine Art Schreibtisch-Unterlage, auf der ein anderes Kind Türme aus Gummistöpseln baut, durch Ingangsetzen eines Rüttel-Motors zum Wackeln und damit das Bauwerk des Interaktionspartners zum Einsturz bringen kann. Gemessen wird hier lediglich die Anzahl erfolgter Vibrationen.

Abbildung 2 zeigt das Bedienungspult des Aggressionssimulators von BUSS (1961). Dabei handelt es sich um die wohl bislang am häufigsten in der Forschung eingesetzte «Aggressionsmaschine». Der Vp wird die Rolle eines Lehrenden zugeteilt, der die Leistungen eines im Nebenraum befindlichen Lernenden mittels unterschiedlich starker und unterschiedlich lang andauernder Elektroschocks bewerten, d.h. Fehler bestrafen soll. Da der Versuch vermeintlich der Untersuchung der Wirkung von Strafe beim Lernen dient, kann die Vp nach eigenem Ermessen auf die in Wirklichkeit genau festgelegten Fehler ihres Gegenübers, eines Gehilfen des Versuchsleiters, reagieren. Als Aggressionsindikatoren gelten ge-

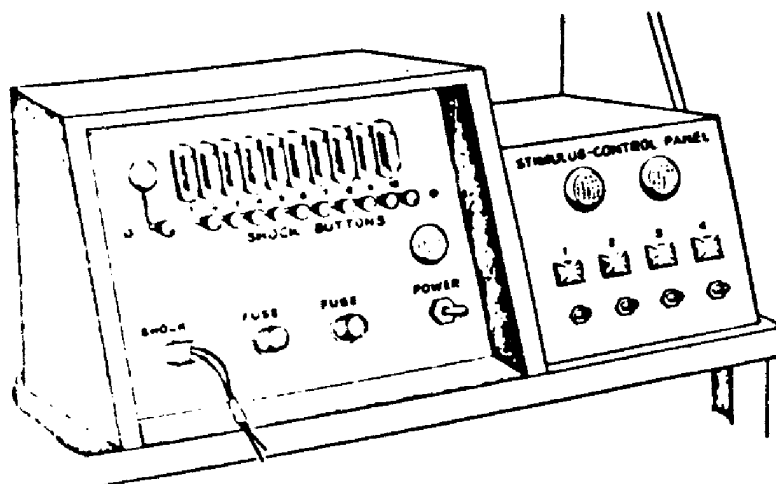


Abbildung 2 Aggressionssimulator von BUSS (Pult der Vp)

wöhnlich die mittlere Intensität (entsprechend den zehn unterschiedlichen Schock-Stärken) und die mittlere Dauer der scheinbar verabreichten Elektroschocks; in insgesamt 60 Durchgängen gibt es 34 Gelegenheiten, Fehler des Lernenden zu bestrafen, d.h. Schocks auszuteilen.

Im Vergleich mit Versuchsanordnungen wie derjenigen von UL-
RICH & FAVELL wird der wohl gewichtigste Vorteil der Verwendung angeblicher Elektroschocks oder verwandter Strafreize in Aggressionssimulatoren deutlich: Schwache Elektroschocks stellen Schädigungen dar, die schmerzhaft, aber ungefährlich sind, die sich ohne großen Aufwand wiederholt erzeugen, leicht abstufen und quantifizieren lassen.

Während die Bussche Aggressionsmaschine eine charakteristische Rollenverteilung zwischen der Vp und ihrem Interaktionspartner voraussetzt, die man in Analogie zu bestimmten pädagogischen Interaktionsformen sehen könnte, betonen die Konstruktionsprinzipien anderer Aggressionssimulatoren stärker die Ausgewogenheit aggressiver Reaktionsmöglichkeiten bzw. die Gleichgewichtigkeit der Interaktionspartner. *Abbildung 3* zeigt als Beispiel das «Interaktions-Aggressometer» von DA GLORIA & DE RIDDER (1977). Hier sitzen beide Versuchspartner vor identischen Bedienungspulten und können wechselweise Elektroschocks vergeben und empfangen, und zwar letzteres über den Griff eines Stiftes, mit dem sie eine sensomotorische Aufgabe auszuführen haben. Die Vpn können zwischen drei Schock-Knöpfen wählen, die mit unterschiedlicher Wahrscheinlichkeit die Blockierung des Ziels des Gegenspielers (nämlich den Stift in die passende Vertiefung einzu-

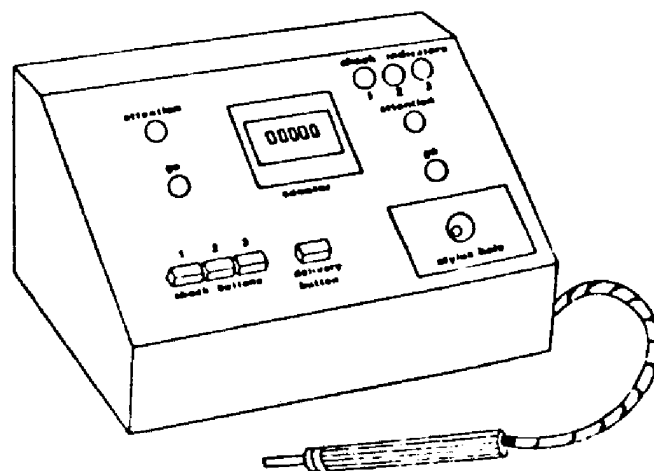


Abbildung 3 Aggressionssimulator von DA GLORIA & DE RIDDER

führen) bewirken, sofern dies auf ein Signal hin schnell genug geschieht. Der vermeintliche Sinn des Experiments liegt in der Erforschung der Störbarkeit motorischer Leistungen durch Elektroschocks. Als Aggressionsindikator dient die mittlere Anzahl der während der letzten 25 von insgesamt 45 Durchgängen verabreichten Schocks.

Der in Aggressionssimulatoren zumeist mindestens suggerierte Interaktionscharakter zwischen zwei Personen wird auch bei der Aggressionsmaschine von SCHMIDT-MUMMENDEY (1972) deutlich (Abbildung 4). Die Vp soll ebenso wie ihr Gegenüber auf einer «Landkarte», die dem «Lastwagen-Spiel» von DEUTSCH & KRAUSS (1960) entlehnt ist, abwechselnd Schritt um Schritt vom Start bis zum Ziel gelangen; gewonnen hat, wer als erster sein Ziel erreicht. Da man nur gewinnen kann, wenn man den kürzesten, allerdings durch einen Engpaß führenden Weg wählt, muß es jeweils in diesem nur durch einen der beiden Spieler passierbaren Engpaß zu einer Konfrontation kommen. An dieser Stelle wird nun das BUSSsche «Erzieher»-Konzept übernommen: Wenn der (durch einen Gehilfen des Versuchsleiters mit vorher festgelegten Reaktionen dargestellte) Gegenspieler signalisiert, daß er nicht zurückweichen will, soll die Vp ihn durch die Vergabe unterschiedlich intensiver und dauerhafter Elektroschocks (elf Stärkegrade) zum Zurückweichen bewegen. Der vermeintliche Zweck des Experiments besteht in der Untersuchung der Wirkung von Strafe. In eigenen Untersuchungen (zusammenfassend MUMMENDEY, 1978) wurde gezeigt, daß vermeintliche Elektroschocks durch das Mittel der «Geldstrafe» (angeblicher Entzug von Vpn-Geldern) ersetzbar sind; das identische Gerät wird dann so interpretiert, daß die elf Straf-Tasten ei-

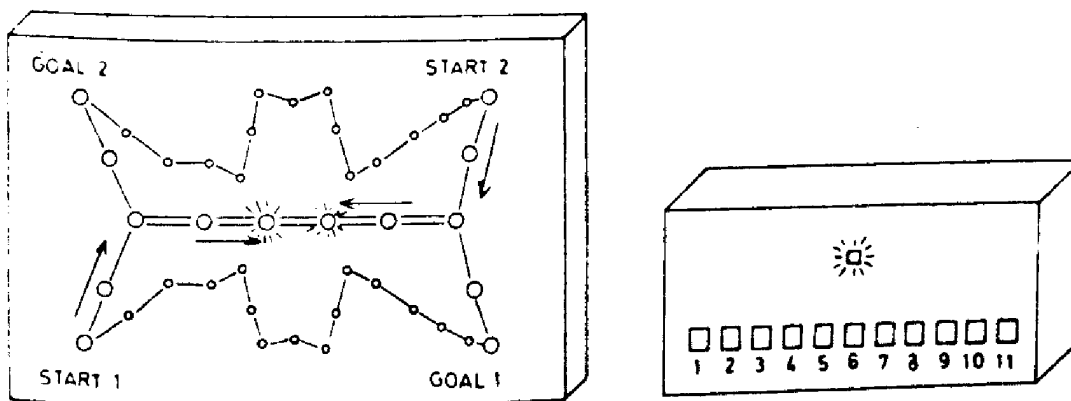


Abbildung 4 Aggressionssimulator von SCHMIDT-MUMMENDEY

nen Geldzählautomaten steuern. Um die offensichtliche Ungleichheit der Interaktionspartner zu kompensieren, startet der Gegenspieler mit einer angeblich doppelt so hohen Bezahlung wie die Vp selbst für ihre Teilnahme am Experiment erhält. Ein kontinuierliches Ansteigen der aggressiven Reaktion (der mittleren Straf-Intensität und -Dauer) wird dadurch begünstigt, daß sich der Gegenspieler im Verlauf der 27 Durchgänge in zunehmendem Maße widersetzt, so daß die Vp – will sie die Instruktion befolgen und nicht selbst aufgeben – gewöhnlich mit zunehmend aggressiven Reaktionen antwortet.

Der aus mehreren Gründen zu bevorzugende Verzicht auf den Einsatz (auch nur vorgetäuschter) Elektroschocks wird auch in solchen Aggressionssimulatoren realisiert, die sich von vornherein auf verbales Verhalten, d.h. Verbalinjurien beschränken. Beispielsweise verwendete DANN (1974) ein einfaches Gerät, das der Vp bzw. ihrem Interaktionspartner durch Knopfdruck bzw. Lichtsignal auf einer Sieben-Punkte-Skala (vgl. *Abbildung 5*) anzeigen soll, wie intensiv entsprechende Urteile des Gegenübers hinsichtlich des Zutreffens einer Reihe von Eigenschaften auf ihre Person ausfallen. Dabei sind die Eigenschaftsbezeichnungen bipolar angeordnet und entsprechend der gleichen Anordnung der Skalenpunkte auf dem Aggressionssimulator. Negative Eigenschaftspole sind z.B. «dumm», «arbeitscheu», «schmutzig», «verlogen» usw. DANN'S Versuchsanordnung entfernt sich wegen der Beschränkung auf verbales Verhalten von der Untersuchung «offenen» Verhaltens in Richtung auf Verhaltens-Intentionen. Verhaltensabsichten stellen zwar ein Kernstück mancher Aggressionsdefinition dar, werden jedoch gewöhnlich eher der Einstellungs-Ebene als der Ebene des Verhaltens zugerechnet.

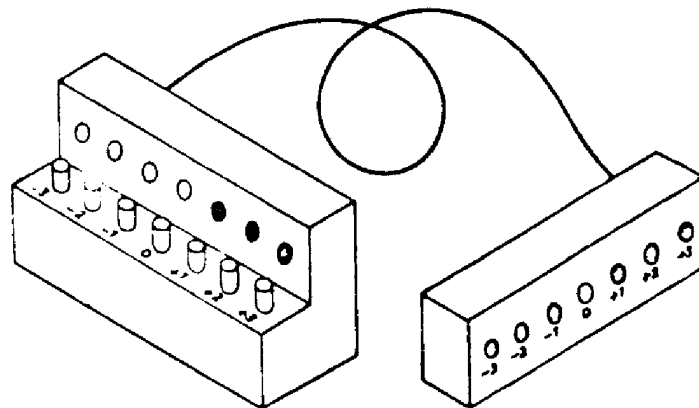


Abbildung 5 Aggressionssimulator von DANN

3. Zur Validität von Aggressionssimulatoren

Die Gültigkeit von mit Aggressionssimulatoren gemessenen Werten aggressiven Verhaltens mag aus verschiedenen grundsätzlichen Erwägungen, die logische und ökologische Argumente einbeziehen, fragwürdig erscheinen. Wie immer, wenn psychologische Konstrukte zu erfassen sind, werden jedoch empirische Hinweise auf den Grad des Zusammenhanges zwischen solchen Maßzahlen und andersartigen Aggressionsmaßen die stärksten Argumente sein.

Bei einer kürzlich vorgenommenen Durchsicht von über 500 empirischen/experimentellen Untersuchungen zum aggressiven Verhalten aus den letzten fünfzehn Jahren (MUMMENDEY, SCHIEBEL, TROSKE & STURM, 1977) ergab sich allerdings, daß nur in sehr seltenen Fällen überhaupt mehr als ein einziger Aggressionsindikator verwendet wurde, wenn aggressives Verhalten gemessen und interpretiert werden sollte. Für diese nur in rund vier Prozent aller Fälle vorgenommenen Mehrfachmessungen ergab sich eine mittlere Interkorrelation verschiedener Aggressionsmaße von knapp unterhalb der 0,30-Grenze, also einem Wert, der aus der empirischen Persönlichkeitsforschung als «magische», weil so schwer zu überwindende Höhe des Zusammenhanges von Verhaltensmerkmalen in verschiedenen Situationen angesehen wird (vgl. z.B. ENDLER, 1973, S. 288). Betrachtet man dabei die Zusammenhänge zwischen an Aggressionssimulatoren und mit anderen Verfahrensarten gemessenen Aggressionsmaßen, so unterschreiten diese im Mittel noch die genannte Schranke – mit Ausnahme der Korrelationen zwischen Aggressionssimulator und Verhaltens-Rating, die BERKOWITZ, PARKE, LEYENS & WEST (1974) bei jugendlichen Delinquenten in der Größenordnung zwischen 0.38 und 0.49 fanden. Überhaupt ergab sich, daß Interkorrelationen umso niedriger waren, je «offener» das betrachtete Verhalten, und daß die Zusammenhänge umso eher interpretierbar waren, je ähnlicher das Erfassungsniveau war (z.B. wenn in beiden Fällen verbale Berichte erhoben wurden).

Faßt man jede Aggressionsmessung für sich genommen als eine *Situation* eigener Art auf, so scheint sich aufgrund unseres Literaturüberblicks zu bestätigen, daß Aggressionsmessungen umso eher miteinander korrelieren, je ähnlicher die Situationen sind, in denen aggressives Verhalten gemessen wird.

In der oben erwähnten Arbeit konnte durch eine Untersuchung von 54 Realschülern der Abgangsklassen, bei denen aggressives Verhalten an der Aggressionsmaschine von SCHMIDT-MUMMENDEY (in der «Geldstrafen»-Version) mit Aggressions-Inventaren, Lehrer-Ratings, Selbst-Ratings und projektiven Verfahren erfaßt wurde, der in der Literaturübersicht erhaltene Eindruck nur bestätigt werden: Während Aggressionsmaße nicht-offenen Verhaltens maximal bis unter die 0.30-Marke miteinander korrelierten, waren die Zusammenhänge mit den Indikatoren des Aggressionssimulators, nämlich mittlerer Intensität und Dauer der verabreichten Strafreize, gleich Null.

Die niedrigen Zusammenhänge zwischen verschiedenen, empirisch oder experimentell gewonnenen Aggressionsmaßen können nicht ausschließlich damit erklärt werden, daß es sich häufig um nicht weiter überprüfte Ad-hoc-Maße mit ungeklärter Reliabilität handelt bzw. dadurch, daß fast alle berichteten Korrelationsmaße nicht minderungskorrigiert sind und die wahre Höhe des Zusammenhanges zu unterschätzen scheinen, ebenso nicht durch mögliche Hinweise auf überhomogene Personenstichproben etc., also diejenigen hier angebrachten Argumente, die auch in der Person-Situation-Diskussion der 70er Jahre eine Rolle gespielt haben. Zu oft wurden reine Nullkorrelationen gefunden, und «dremol null is null bliff null» (Kölner Karnevalsschlager). Auf der anderen Seite ließen sich mit an Aggressionssimulatoren gewonnenen Maßen aggressiven Verhaltens in der Vergangenheit doch psychologisch höchst sinnvolle Resultate erzielen, und es ergaben sich auch respektable Reliabilitätswerte für solche Verfahren (vgl. MUMMENDEY, 1978). Mangelnde Validierung an einem Kriterium kann prinzipiell auch Ausdruck hoher situativer Spezifität sein. Schließlich scheint gerade die oft zu hörende Kritik an der Anwendung von Aggressionsmaschinen zu bestätigen, daß es sich bei experimentellen Situationen, in denen diese Versuchsanordnungen Anwendung finden, um außergewöhnliche Situationen handelt. Es lag daher nahe, weitere Hinweise auf die mögliche Validität von mit Aggressionssimulatoren gewonnenen Meßwerten über die Berücksichtigung des Grades an Spezifität bzw. Generalität von Situationen, in denen ein Individuum aggressiv handeln kann, zu erhalten.

4. Beziehungen zwischen Situationsspezifität und Validität experimentell gewonnener Aggressionsindikatoren

In den meisten an Aggressionssimulatoren erfaßten Reaktionen überwiegt offensichtlich die «instrumentell-aggressive» über die «Wut»- bzw. «Ärger-Aggressions»-Komponente. Zwar enthalten die meisten experimentellen Situationen der geschilderten Art sicherlich auch Momente des Ärgers bzw. der allmählich oder spontan entstehenden Wut über den Gegenspieler, der «nicht lernen will», «widerborstig ist», das zielgerichtete Verhalten der Vp «frustriert» usw., doch erscheint das «sachbezogene», gelegentlich geradezu «versachlichte» Streben nach «Gewinnen», «Vorwärtskommen», «Sich-Durchsetzen», also nach Zielerreichung dominant.

Betrachtet man die in der Fachliteratur zur – eher gelegentlichen – Validierung von Aggressionsmaßen herangezogenen Kriterien, so beziehen sich diese meist nicht nur auf alternative Erhebungsarten von Maßen aggressiven Verhaltens, sondern oft auch auf sehr unterschiedliche Aggressions-Konzepte. Instrumentell-aggressives Verhalten läßt sich beispielsweise schon allein deshalb schwerlich durch Persönlichkeitsinventare oder Ratingsverfahren validieren, weil die meisten Fragebogen nicht den «instrumentellen» Aspekt aggressiven Verhaltens messen und weil den meisten Ratern «instrumentelle» Aggression im wahrsten Sinne des Wortes kein Begriff ist. Schließlich scheint letzterer ja auch widersprüchlich zu sein, da «aggressiv» zu den eher negativ bewerteten Persönlichkeitseigenschaften gehört (vgl. etwa ANDERSON, 1968), während «Sich-Durchsetzen», «Ziele trotz Hindernissen erreichen» etc. nicht nur in unserer, sondern vermutlich in den meisten uns bekannten Gesellschaften positive Konnotationen besitzen dürfte. Eine Validierung instrumentell-aggressiver Verhaltensmaße, die in einer experimentellen Situation gewonnen wurden, durch Fragebogen oder Ratingtechniken hätte vermutlich umso eher Aussicht auf Erfolg, je einander ähnlicher die Verhaltensmuster sind, die im Experiment gemessen und mit Papier-Bleistift-Verfahren erfragt werden. Soll die Gültigkeit im Sinne transsituativer Generalität instrumentell-aggressiven Verhaltens untersucht werden, so sollten Informationen über den Grad der Ähnlichkeit zwischen der experimentellen Standardsituation und der Vergleichssituation in systematischer Weise in die Analyse einbezogen werden.

5. Eine Untersuchung zum Zusammenhang zwischen Spezifität und Validität «instrumentell-aggressiven» Verhaltens

Um instrumentell-aggressive Verhaltensweisen, wie sie im Alltag einer bestimmten Personengruppe auftreten können, mit experimentellen Indikatoren vergleichen zu können, wurden insgesamt 32 Beschreibungen solcher Verhaltensweisen gesammelt. Sie sollten in infinitiver Formulierung mögliches instrumentell-aggressives Verhalten von Studenten bezeichnen (z.B. «bei der Parkplatzsuche alle erlaubten Mittel einsetzen, um noch einen Parkplatz zu bekommen», «im Fahrstuhl unbemerkt das Plakat einer bekämpften politischen Gruppe überkleben», «einen Partner absichtlich eifersüchtig machen, um ihn fester an sich zu binden» usw.). Diese Statements wurden von elf Experten, und zwar ausschließlich Personen, die schon als Versuchsleiter am Aggressionssimulator von SCHMIDT-MUMMENDEY gearbeitet hatten, hinsichtlich ihrer Nähe bzw. Ähnlichkeit zum Verhalten der Vp in der experimentellen Situation an diesem Aggressionssimulator skaliert. Die eindimensionale Skalierung nach der THURSTONEschen Methode der gleicherscheinenden Intervalle (vgl. EDWARDS, 1957) führte zur Auswahl von folgenden 18 Alltagsverhaltensweisen (in Klammern THURSTONE-Skalenwert, der den Grad der Ähnlichkeit der Verhaltensweise mit derjenigen in der experimentellen Situation – mindestens 1, höchstens 11 – ausdrückt):

1. Bei längerem Warten vor einer Telefonzelle die gerade telefonierende Person auffordern, sich kurz zu fassen (5)
2. Sich in öffentlichen Verkehrsmitteln energisch durchsetzen, um einen Sitzplatz zu erreichen (9)
3. Als Autofahrer nachts den Entgegenkommenden zwingen, zuerst abzublenden (10)
4. In einer scharfen Diskussion den Beitrag eines Konkurrenten durch Lächerlichmachen abwerten (7)
5. Jemanden, der in öffentlichen Verkehrsmitteln den gleichen Sitzplatz beansprucht, freundlich aber bestimmt darauf hinweisen, daß man eher da war (4)
6. Einen gefährlich dicht auffahrenden Überholer durch absichtliches Langsamfahren zurechtweisen (8)
7. Auf Anweisung eines Vorgesetzten Handlungen zum Nachteil eines Kollegen ausführen (9)
8. Einen Partner absichtlich eifersüchtig machen, um ihn fester an sich zu binden (4)
9. Einen gefährlich bummelnden Autofahrer durch Auffahren oder Signalgeben dazu bewegen, schneller zu fahren oder den Weg frei zu machen (10)
10. Im Fahrstuhl unbemerkt den Anschlag einer bekämpften politischen Gruppe überkleben (2.5)

11. Bei Konflikten mit Familienangehörigen oder Partnern prinzipiell hart bleiben, wenn man von der Richtigkeit der eigenen Meinung überzeugt ist (6.5)
12. Einen unerlaubt überholenden Wagen durch dichtes Auffahren daran hindern, sich wieder einzuordnen (9)
13. Gute Beziehungen zu einem Dozenten ausnutzen, um auf Kosten anderer einen Praktikumsplatz zu erhalten (9)
14. Auf Drängen eines Beifahrers rücksichtsloser fahren als sonst (8)
15. Als Angehöriger einer politischen Gruppe das Plakat einer zu bekämpfenden Gruppe mit eigenen Parolen überkleben (3)
16. Bei Familienkrach grundsätzlich nicht nachgeben und die anderen zum Einlenken zwingen (9)
17. Bei der Parkplatzsuche alle erlaubten Mittel einsetzen, um noch einen Parkplatz zu bekommen (9)
18. Sich beim Geschäftsführer beschweren, um eigene Wünsche gegenüber einem(r) Kellner(in) oder einem(r) Verkäufer(in) massiv durchzusetzen (7).

Daraufhin absolvierten 50 Fachhochschüler der Fachrichtung Wirtschaft das Spiel an dem Aggressionssimulator, wobei für jede Vp die mittlere Intensität der erteilten Geldstrafen als Indikator instrumentell-aggressiven Verhaltens ermittelt wurde. Anschließend wurden jeder Vp die oben aufgeführten 18 Alltagsverhaltensweisen schriftlich vorgelegt, und sie wurde gebeten anzugeben, welche dieser Verhaltensweisen sie selbst schon einmal ausgeführt habe. Berechnete man nun für jede dieser Verhaltensweisen den Zusammenhang zwischen der Bejahung, d.h. Bestätigung, dieses Alltagsverhalten ausgeführt zu haben und dem Aggressionsmaß der Vp in der experimentellen Situation, so ergab sich lediglich für Verhaltensweise 9 («einen gefährlich bummelnden Autofahrer . . .») eine sowohl mit PEARSONS r_{xy} als auch mit KENDALLS tau signifikante Korrelation (0.31 bzw. 0.24), im übrigen ergaben sich ausschließlich Nullkorrelationen. Die Verhaltensweise 9 weist den relativ höchsten THURSTONE-Skalenwert auf, ähnelt also der Aggressionssimulator-Situation am stärksten.

Berechnete man nun den Zusammenhang zwischen a) dem Skalenwert eines Statements, d.h. seiner Nähe zur Experimentalsituation, und b) der Höhe der Korrelation zwischen der Bejahung eines Statements (Ausführung der entsprechenden Verhaltensweise) und dem Maß aggressiven Verhaltens am Aggressionssimulator, so betrug jener Zusammenhang 0.27 bzw. 0.28. Es resultierte also eine schwach positive Korrelation zwischen der Experiment-Ähnlichkeit aggressiver Alltagsverhaltensweisen und dem Validitätskoeffizienten der Straf-Intensität am Aggressionssimulator, der auf unterschiedlichen Alltagsverhaltensweisen beruhte. Mit anderen

Worten ergaben sich gewisse Hinweise darauf, daß die Kriteriumsvalidität des experimentellen Aggressionsindikators von der Ähnlichkeit des Kriteriums mit der experimentellen Situation abhängt.

Scheint somit die Gültigkeit und damit die mögliche Generalisierbarkeit eines Maßes aggressiven Verhaltens im psychologischen Labor weitgehend auf äußerst spezifische Situationen, z.B. Straßenverkehrssituationen beschränkt zu sein und stellt sich somit als insgesamt recht unbefriedigend dar, so ermuntert doch das tendenziell erwartungsgemäße Teilergebnis zur Suche nach möglichen störenden Bedingungen. Vor allem eine Tendenz bei den Vpn, bei verbaler Befragung über aggressives Verhalten beschönigend zu antworten, d.h. die Reaktionstendenz der «sozialen Erwünschtheit» (social desirability) kann als Störvariable niemals ausgeschlossen werden. Als daher fünf der oben erwähnten Experten jede der 18 Alltagsverhaltensweisen zusätzlich nach dem Grad der sozialen Erwünschtheit einstuften, korrelierte dieses Social Desirability-Maß erwartungsgemäß mit dem THURSTONE-Wert, der die Nähe zum experimentellen Konzept «instrumentelle Aggression» anzeigt, negativ (-0.34). Partialisierte man daraufhin die Erwünschtheits-Variable aus der Korrelation zwischen dem Skalenswert der Alltagsverhaltensweisen und der Aggressionsmaschinen-Validität aus, so erhöhte sich diese von 0.27 auf 0.31. Dieser Zuwachs ist zwar nicht beachtlich, signalisiert aber die mögliche Effektivität einer Kontrolle der Störvariablen «Soziale Erwünschtheit einer aggressiven Verhaltensweise» mit einer verbesserten Methodik.

6. Eine weitere Untersuchung mit Selbsteinschätzungen aggressiven Verhaltens in einer Bogus-Pipeline-Situation

Eine gegen herkömmliche Befragung verbesserte, bei Beschönigungstendenzen provozierenden Themen indizierte Erhebungsmethode stellt das Bogus-Pipeline-Paradigma dar (vgl. SIGALL & PAGE, 1971; BRACKWEDE, 1979). Das Verfahren, das mit falscher physiologischer Rückmeldung über die affektive Einstellungskomponente arbeitet, ist allerdings recht aufwendig.

In einem Experiment wurden 60 Studenten sowohl hinsichtlich ihres aggressiven Verhaltens an dem Aggressionssimulator von SCHMIDT-MUMMENDEY als auch bezüglich ihrer Selbsteinschätzung hinsichtlich der oben aufgeführten 18 instrumentell-aggressi-

ven Alltags-Verhaltensweisen, diesmal jedoch unter Bogus-Pipeline-Bedingungen untersucht. Zusätzlich wurde eine postexperimentelle Befragung bezüglich Erlebnisweisen am Aggressionssimulator vorgenommen. Zur Kontrolle sequentieller Effekte wurde die Hälfte der Vpn zuerst dem Spiel am Aggressionssimulator mit postexperimenteller Befragung, die andere Hälfte dagegen zuerst der Selbstbeurteilung aggressiver Verhaltensweisen unter Bogus-Pipeline-Bedingungen unterzogen.

Die Messung aggressiven Verhaltens am *Aggressionssimulator* erfolgte in verkürzter Form (insgesamt sieben Durchgänge), nachdem sich Korrelationen um 0.90 zwischen der Kurzfassung und der ursprünglichen Fassung des Versuchs hinsichtlich aller relevanten Parameter ergeben hatten. Die Vp wurde wie üblich instruiert, als erste ans Ziel der «Landkarte» (vgl. *Abb. 4*) zu gelangen zu versuchen und im Falle einer Konfrontation in dem Engpaß den Gegenspieler durch Betätigung der Straftasten, die einen Geldautomaten steuerten, zum Zurückweichen zu zwingen. Je höhere Tasten und je länger sie die Tasten drückte, desto mehr Geld werde dem Gegner (maximal sein gesamtes Vpn-Honorar von DM 20) abgezogen. In den sieben Durchgängen des Spiels weicht der Gegenspieler stets nach 1, 0, 2, 2, 4, 4, 4 Strafen zurück, d.h. er verhält sich nach anfänglich einmaligem und anschließend ausbleibendem Widerstand in zunehmendem Maße renitent. Gemessen wurden die mittlere Gesamt-Strafintensität und die mittlere Gesamt-Strafdauer, ferner die entsprechenden Werte für diejenigen drei Phasen des Versuchs, in denen der Gegenspieler einen unterschiedlich starken Widerstand bietet, nämlich den 1., den 3.-4., und den 5.-7. Durchgang.

Bei der *postexperimentellen Selbsteinschätzung* wurde mit Hilfe siebenstufiger Ratingskalen – neben einigen irrelevanten Fragen – erhoben, wie «interessant» das Start-Ziel-Spiel gewesen sei, wie «sympathisch» man den Gegenspieler gefunden habe, in welchem Maße die Vp «ärgerlich oder wütend» geworden sei, inwieweit sie sich habe «gegen den Gegenspieler effektiv durchsetzen» wollen, in welchem Maße sich ihre Reaktionen als «aggressiv» und als «zielstrebig» bezeichnen ließen.

Die *Bogus-Pipeline-Messung aggressiver Alltagsverhaltensweisen* geschah wie folgt: Die Vp wurde in einen kleinen, mit elektronischen Versuchsaapparaturen vollgestopften Raum geführt und in einem Vorversuch zunächst von der Fähigkeit eines EMG-Gerätes zur Messung unwillkürlicher Muskelspannungen, Ausmaß und Richtung von Einstellungen auf einer Sieben-Punkte-Skala zu erfassen, überzeugt. (Anstelle einer detaillierten Beschreibung des Verfahrens wird auf die zitierte Literatur verwiesen.) Mit einer einzigen Ausnahme äußerte keine Vp Zweifel an der offensichtlichen Korrektheit der Messung. Anschließend wurde die Vp gebeten, im folgenden Beurteilungsdurchgang die Meßergebnisse, die das Gerät bezüglich der eigenen Urteile liefern werde, vorherzusagen, damit die subjektiven Vorhersagen den Outputwerten des Rechners, der zuvor die von den Vpn akzeptierten Rückmeldungen geliefert hatte, verglichen werden könnten. Während die Outputwerte der Vp vorenthalten wurden, gab diese daraufhin ihre Selbsteinschätzung zu den 18 oben aufgeführten aggressiven Alltagsverhaltensweisen auf einer Sieben-Punkte-Skala. Von diesen Selbstbeurteilungen kann, ohne daß dies unmittelbar überprüfbar wäre, erwartet werden, daß sie relativ beschönigungsfrei sind.

Es ergaben sich folgende Resultate:

Zwischen den Indikatoren instrumentell-aggressiven Verhaltens am Aggressionssimulator und der Bejahung einiger instrumentell-aggressiver Alltagsverhaltensweisen traten schwach positive Korrelationen (bis zur Höhe von 0.27) auf. Die Korrelation zwischen dem THURSTONE-Skalenwert der Alltagsverhaltensweisen, d.h. ihrer phänomenalen Nähe zu der Experimentalsituation, und dem Zusammenhang zwischen Verhaltensausführung im Alltag und Straf-Intensität im Experiment betrug insgesamt 0.14; sie war in den Versuchsdurchgängen 1 (0.31) und 3 bis 4 (0.27) höher als in den Durchgängen 5 bis 7 (0.10). Nur für die Anfangsphasen des Spiels am Aggressionssimulator werden demnach Zusammenhänge in der von der ersten Untersuchung bekannten Größenordnung erreicht. Finden sich somit keine substantziellen Korrelationen in bezug auf die experimentellen Indikatoren, so ergibt sich hinsichtlich der postexperimentellen Selbsteinschätzung:

Die Korrelationen zwischen der skalierten Experiment-Nähe der aggressiven Alltagsverhaltensweisen und dem Zusammenhang zwischen deren Bejahung und der Selbsteinschätzung des Verhaltens am Aggressionssimulator betragen

- für die Beurteilung, «ärgerlich/wütend» geworden zu sein	0.53
- für die Beurteilung, «aggressive» Reaktionen gezeigt zu haben	0.35
- für die Beurteilung, sich «effektiv durchsetzen» zu wollen	0.45
- für die Beurteilung der eigenen Reaktionen als «zielstrebig»	0.46.

Demnach scheinen es auf den ersten Blick eher die subjektiven, weniger die objektiven Aggressionsindikatoren aus der Experimentalsituation zu sein, die eine Beziehung zu Alltagsverhaltensweisen zeigen. Entsprechend der Konzeption des Spiels am Aggressionssimulator ergeben sich insgesamt und in allen Phasen des Versuchs auf dem 1 %-Niveau gesicherte, positive Korrelationen zwischen objektiver Aggressivitätsmessung (Straf-Intensität) und dem subjektiven, bei postexperimenteller Selbsteinschätzung erhobenen Urteil, man habe sich gegen den Gegenspieler «effektiv durchsetzen» wollen.

Das zunächst unerwartete Resultat, daß sich auch bei Verwendung einer Versuchsanordnung, in der beschönigende Antworttendenzen weitgehend reduziert sein müßten, keine substantziellen Validitätshinweise für experimentelle Aggressionsindikatoren, stattdessen jedoch für postexperimentelle Selbsteinschätzungen ergeben, lenkt den Blick auf offensichtliche Unzulänglichkeiten bisheriger Anwendung von Aggressionssimulatoren. Die alleinige

Auswertung offen beobachtbarer Reaktionen führt vermutlich zu psychologisch unbefriedigenden Daten, weil offen beobachtbare Verhaltensweisen nur einen Teil der mehr oder weniger aggressiven Reaktionen einer Vp ausmachen. Der nicht beobachtbare, möglicherweise jedoch indirekt erfahrbare Teil ihrer Reaktion ist für die Vp ebenso wie für den Forscher vielleicht ebenso wichtig wie der beobachtbare. So machte der Versuchsleiter gelegentlich die Beobachtung, daß sich Vpn, die hohe Straf-Intensitäten gewählt hatten, anschließend als nicht besonders «aggressiv» einschätzten. Hält man die Selbsteinschätzung für relativ glaubwürdig, wozu man bei Selbstbeurteilungen mangels Außenkriterien stets vorläufig gezwungen ist, so kann diese Beobachtung am besten damit erklärt werden, daß verschiedene Vpn das gleiche Verhalten am Aggressionssimulator unterschiedlich definieren oder interpretieren. Z.B. kann für eine Vp die Vergabe intensiver Strafreize in erster Linie Herstellung von Gleichheit zwischen Vp und Gegenspieler, die in dieser wie in den meisten anderen experimentellen Situationen ähnlicher Art verletzt ist, bedeuten.

Akzeptiert man, daß Vpn ihr eigenes Verhalten im Experiment in unterschiedlicher Weise interpretieren und somit objektive experimentelle Aggressionsindikatoren mehrdeutig sind, so kommt der Selbsteinschätzung der Vp der Status einer Moderatorvariable bei der Untersuchung der Validität von Maßen aggressiven Verhaltens am Aggressionssimulator zu. Dementsprechend konnten z.B. SCHEIER, BUSS & BUSS (1978) die moderierende Wirkung von «private self-consciousness» in bezug auf den Zusammenhang zwischen offenen Verhaltensmaßen und Selbstberichtsmaßen aggressiven Verhaltens zeigen. Somit müßten, falls an Aggressionssimulatoren experimentell gewonnene Indikatoren nicht vollständig invalide sind, die Zusammenhänge zwischen der mittleren Straf-Intensität und der unter Bogus-Pipeline-Bedingungen gegebenen Bestätigung, ein bestimmtes aggressives Verhalten im Alltagsleben mehr oder weniger zu zeigen, bei solchen Personen, die ihr Verhalten selbst im Anschluß an die Messung als aggressiv im Sinne des experimentellen Konzeptes von Aggression einschätzen, höher sein als bei solchen Personen, die ihr Verhalten selbst als wenig aggressiv beurteilen.

Zur Prüfung dieser Annahme wurde diejenige postexperimentelle Selbsteinschätzungs-Variable, die das hier untersuchte Konzept «instrumentell-aggressiven» Verhaltens vermutlich am besten repräsentiert, nämlich die Beurteilung, man habe sich gegen den

Gegner «effektiv durchsetzen» wollen, ausgewählt. Die Selbsteinschätzung hinsichtlich «ärgerlich/wütend» schied allein schon deswegen aus, weil sich mehr als 60 % der Vpn postexperimentell als überhaupt nicht ärgerlich/wütend, der Rest auch nur in sehr geringem Ausmaß, bezeichnet hatte. Erwartungsgemäß ergab sich hinsichtlich des Selbstratings, man habe sich «effektiv durchsetzen» wollen, eine symmetrische Verteilung mit zwei Gipfeln. Die Stichprobe wurde am exakten Median getrennt, so daß sich eine Personengruppe mit geringer Durchsetzungstendenz (Median 3.2; n = 33), und eine mit hoher Durchsetzungstendenz (Median 4.7; n = 27) ergaben.

Die Korrelationen zwischen dem Gesamtmaß aggressiven Verhaltens am Aggressionssimulator (mittlere Straf-Intensität) und denjenigen vier Alltags-Verhaltensweisen, hinsichtlich deren interpretierbare Zusammenhänge mit der Selbstbeurteilung, sich «effektiv durchzusetzen» bestanden hatten, stellten sich, getrennt für Personen mit geringer bzw. hoher Durchsetzungstendenz wie folgt dar:

Korrelation zwischen mittlerer Straf-Intensität am Aggressionssimulator und Alltagsverhalten:

	<i>Postexperimentelle Selbsteinschätzung</i> «Es kam mir bei dem Spiel vor allem darauf an, mich gegen den Gegenspieler effektiv durchzusetzen»	
	<i>Low-Scorer</i> (n=33)	<i>High-Scorer</i> (n=27)
3. «Als Autofahrer nachts den Entgegenkommenden zwingen, als erster abzublenden» (10)	0.13	0.24
9. «Einen gefährlich bummelnden Autofahrer durch Auffahren oder Signalgeben . . .» (10)	-0.04	0.34
13. «Gute Beziehungen zu einem Dozenten ausnutzen, um auf Kosten anderer . . .» (9)	-0.13	0.10
18. «Sich beim Geschäftsführer beschweren, um eigene Wünsche gegenüber . . .» (7)	-0.11	0.29

Diesen Korrelationsdifferenzen entspricht im günstigsten Falle zwar nur eine Restwahrscheinlichkeit von rund 10 %, doch bestätigt die auffällige Konsistenz der Richtung der Differenzen, daß nur in denjenigen Fällen, in denen eine Person ihr eigenes Verhalten im Experiment als «instrumentell-aggressiv» interpretiert, positive Zusammenhänge zwischen offen beobachtbarem Verhalten am Aggressionssimulator und im Alltagsleben erwartet werden können.

Die Ergebnisse dieser Untersuchung stützen daher jene Auffassungen über aggressives Verhalten, die die Notwendigkeit der Berücksichtigung der subjektiven Interpretation der Interaktionssituation durch die beteiligten Interaktionspartner betonen (vgl. den Beitrag von Amélie MUMMENDEY, in diesem Buch).

7. Zusammenfassung

Die Erfassung aggressiven, insbesondere instrumentell-aggressiven Verhaltens an Aggressionssimulatoren in experimenteller Situation wird durch die methodischen Vorteile der Kontrollierbarkeit verschiedener Variablengruppen, der Ökonomie und ethische Gesichtspunkte nahegelegt. Bedauerlicherweise mangelt es an Validitätshinweisen für entsprechende Aggressionsmaße. Eigene Untersuchungen mit einem Aggressionssimulator belegen die mangelhafte Validität bzw. hochgradige Situationsspezifität von experimentell gewonnenen Maßen instrumentell-aggressiven Verhaltens. Eine Verbesserung von Zusammenhängen zwischen entsprechenden Meßwerten und Maßen aggressiven Verhaltens in Alltagssituationen ist zu erwarten, wenn die subjektiven Interpretationen des eigenen Verhaltens durch die Versuchspersonen in die Analyse einbezogen werden.

Das Manuskript wurde Januar 1980 abgeschlossen.