

BIELEFELDER ARBEITEN ZUR SOZIALPSYCHOLOGIE

Nr. 49

(Mai 1979)

Dietrich Brackwede:

Das Bogus-Pipeline-Paradigma
und seine Bewertung nach acht Jahren

Zusammenfassung:

Das auf E.E.JONES und Mitarbeiter zurückgehende Bogus-Pipeline-Verfahren zur relativ verzerrungsfreien Erfassung von affektiven Urteilen bzw. Einstellungen wird auf der Grundlage von 17 Experimenten dargestellt und einer Kritik unterzogen.

Einen direkten Zugang zur Seele zu besitzen - zumal zu derjenigen anderer Personen - ist ein Gedanke, der manchen Psychologen, Theologen, Psychoanalytiker oder Kriminalisten wohl verlocken möchte. Der Wunsch danach mag vielleicht heute ein wenig schwächer sein als in früheren Zeiten, und er mag bei Grundlagenforschern weniger ausgeprägt sein als bei solchen Anwendern, die immer wieder feststellen, daß man auch mit Hokuspokus gute praktische Erfolge erzielen kann. Herauszubekommen, was "wirklich in einem Menschen vorgeht", mag damit entschuldbar sein, daß eine "psycho"-hungrige Öffentlichkeit immer wieder danach verlangt, weil sie an die Möglichkeit als solche gern glauben mag. Ist es daher eine Schande, sich eben diesen Glauben experimentell zunutze zu machen, wenn es dem vermeintlichen wissenschaftlichen Fortschritt dient? Im Bogus-Pipeline-Paradigma ist dies geschehen. Ob es einen Fortschritt gebracht hat, soll zu zeigen versucht werden.

Die Idee der "bogus pipeline" (BPL) entsprang dem für die psychologische Einstellungsforschung konstitutiv ungemütlichen Zustand, daß man prinzipiell nicht weiß, wann und in welchem Ausmaß verbale Reaktionen, z.B. Antworten auf Attitude-Statements, durch - wie man es sich erklärt - Reaktionstendenzen der sozialen und/oder personalen Erwünschtheit (social desirability) verfälscht sind. Hätte man einen direkten oder wenigstens besseren Zugang zu dem, was eine Person über einen Meinungsgegenstand tatsächlich denkt, wie sie ihn bewertet und welche Verhaltensintentionen sie ihm gegenüber hegt, so könnten die störenden Reaktionstendenzen vielleicht überwunden werden. Die Amerikaner nennen eine solche Direktleitung "pipeline".

Da die meisten Psychologen - wohl im Gegensatz zu den meisten ihrer Klienten und Angehörigen einiger vorne aufgeführten Fakultäten - an die Existenz einer solchen Pipeline nicht glauben mögen, versuchte E.E.JONES mit verschiedenen Mitarbeitern die Vorspiegelung der Existenz einer derartigen Methode der Wahrheitsfindung methodisch nutz- und fruchtbar zu machen. "Bogus" steht im Amerikanischen für "falsch", "unecht", "geschwindelt".

Könnte man Personen davon überzeugen, daß es mit einer speziellen physiologischen Apparatur möglich sei, Richtung und Ausmaß emotionaler Reaktionen festzustellen und würde sodann von ihnen verlangen, die mit dem

Gerät gemessenen Werte vorherzusagen bzw. nachträglich zu schätzen, so erhalte man Antworten, die vermutlich relativ frei von den Papier-Bleistift-Verfahren gewöhnlich verfälschenden Reaktionstendenzen wären. Dieser, gleichsam die Reliabilität und Validität eines Lügendetektors vortäuschenden Manipulation liegt die Annahme zugrunde, daß man sich nicht gern von einem Gerät taxieren lassen möchte - "subjects do not want to be second-guessed by a machine" (JONES & SIGALL 1971, p.349). Das BPL-Paradigma schließt sich damit an die Arbeiten der GERARD, BRAMEL, VALINS u.a.m. mit der Applikation falschen physiologischen Feedbacks im sozialpsychologischen Experiment an.

Die grundlegende, im folgenden beschriebene, von SIGALL & PAGE (1971) applizierte BPL-Versuchsanordnung kann somit auf eine Reihe von Vorläufern, insbesondere in den Experimenten von JONES, BELL & ARONSON (1972) und COOPER (1971) zurückgreifen. Da sie im wesentlichen unverändert in allen nachfolgend referierten Untersuchungen angewendet wurde, soll sie im Detail beschrieben werden.

Die Vp erhält zunächst einen aus mehreren Items bestehenden Einstellungsfragebogen (der zur späteren Überzeugung, daß die Maschine tatsächlich zutreffende Messungen liefert, dient). Um Verfälschungstendenzen hier gering zu halten, beschäftigen sich die Items in der Regel mit relativ unverfänglichen Gegenständen (Film, Musik, Autos, Sport etc.). Die Antworten können vom V1 nicht eingesehen werden. Wenn der V1 jedoch anschließend die Vp in den eigentlichen Versuchsraum bringt, kann ein Gehilfe die Antworten der Vp unbemerkt kopieren. Der V1 erläutert nun der Vp, daß die Verwendung von Fragebögen in der psychologischen Forschung eine ganze Reihe von Problemen mit sich bringe und man sich daher mit einem kürzlich (z.B. von Prof. X an der Y-Universität entwickelten) "adaptierten Elektromyograph" bzw. "EMG" beschäftigen wolle, der direkte und genaue physiologische Einstellungsmessungen ohne die Möglichkeit von willkürlichen oder unwillkürlichen Reaktionsverzerrungen ermögliche. Die Vp wird vor eine Konsole gesetzt, auf der eine halbkreisförmige Sieben-Punkte-Skala (von -3 bis +3) angebracht ist. Mittels eines Steuerrades kann sie einen Zeiger in die entsprechende Position auf dem Halbkreis bringen. Längsseits der Konsole an einem Metallkasten befindet sich eine ganze Reihe von Anzeigevorrichtungen, Lämpchen, Verbindungskabeln usw., ferner die Aufschrift "EMG". Der Kasten enthält eine entsprechende, halbkreisförmige Skala mit den sieben Punkten, einem Zeiger und mit der Aufschrift "EMG-Output". Von dem Kasten gehen Verbindungskabel zu einem eindrucksvoll aussehenden Gerät, das als Kleincomputer kenntlich ist, sowie zu zwei Elektroden, die sich auf der Hautoberfläche der Vp befestigen lassen.

Der V1 erläutert der Vp, daß das EMG die elektrischen Potentiale in Muskeln messe und dann, wenn die Vp das Steuerrad fest auf den Skalenpunkt 0 (die Mitte) gerichtet halte, die "ersten, unverfälschten Reaktionstendenzen" gegenüber dem zu beurteilenden Meinungsgegenstand (z.B. Attitude-Statement) erfasse. Das EMG werte lediglich die unwillkürlichen Muskelbewegungen aus und sei unanfällig gegenüber den groben Muskelbewegungen.

Der angeschlossene Kleincomputer werte die betreffenden Reaktionen aus und gebe über den EMG-Output das Ergebnis an. Der VI weist noch auf gewisse Ähnlichkeiten zwischen EMG und Lügendetektor hin, meint aber, das EMG habe gewisse Vorteile vor allem dadurch, daß nicht nur die Intensität, sondern auch die Richtung der Reaktionen gemessen werde.

Der Vp werden nun die Elektroden an den Unterarm angelegt, und es folgt eine "Validierung" des EMG, um die Vp von der Leistungsfähigkeit des Gerätes zu überzeugen. Der VI erläutert, daß die Maschine zunächst an das individuell meist unterschiedliche Grundniveau, d.h. die Baseline der Reaktionen jeder Person angepaßt werden müsse. Dazu sollten die von der Person ganz zu Anfang bereits beantworteten Attitude-Statements dienen. Der VI versieht die beiden Seiten der halbkreisförmigen Skala nun mit den Aufschriften "Zustimmung" und "Ablehnung" und bittet die Vp, das Steuerrad ganz fest in die Hände zu nehmen und sich auf die Meinungsskala zu konzentrieren. Er liest das erste Item vor und knipst einen Schalter am Computer an. Der Computer "arbeitet" einen Moment lang und läßt dann den Zeiger des "EMG-Output" eine bestimmte Position einnehmen. Nach dem dritten Item ermuntert der VI die Vp, das EMG ruhig einmal "zum Narren zu halten", indem sie bewußt in die falsche Richtung denken solle, so daß ihre unwillkürlichen Muskelbewegungen ebenfalls in die Gegenrichtung ihrer Einstellung gehen würden. Nachdem alle Items beantwortet sind (wobei der Gehilfe des VI vom Nebenraum aus den EMG-Output entsprechend den heimlich abgelesenen Papier-Bleistift-Reaktionen manipuliert hatte), vergleichen VI und Vp die EMG-Maßwerte ganz offiziell mit den Papier-Bleistift-Werten und stellen fest, daß das EMG tatsächlich korrekte Messungen geliefert hatte und in diesem Falle keiner besonderen Anpassung an das individuelle Grundniveau bedürfe.

Die Bezeichnungen der Skala werden nun gegen die Beschriftungen "charakteristisch" (+3) und "nicht charakteristisch" (-3) ausgetauscht, und die Vp wird instruiert, daß sie zu einer Reihe von Eigenschaften angeben solle, ob diese für eine ganz bestimmte Gruppe (z.B. Farbige) mehr oder weniger typisch seien. Dabei ginge es in dem Versuch darum, festzustellen, "in welchem Maße jemand mit seinen wirklichen Einstellungen übereinstimmt". Die Vp soll demgemäß die Werte, die das EMG liefern würde, schätzen bzw. vorhersagen. Dazu verbaut der VI der Vp den Blick auf die Output-Skala, bittet sie um ihre unverfälschte Reaktion, während sie das Rad festhält und sich auf die Skala konzentriert und notiert diese Reaktion ebenso wie die angebliche Anzeige auf der Output-Skala. - Im Anschluß an die BPL-Messung werden die Vpn vollständig aufgeklärt.

SIGALL & PAGE (1971): A Little Fading, A Little Faking

Den Nachweis der Überlegenheit des BPL-Paradigmas über herkömmliche Papier-Bleistift-Verfahren wie gewöhnliche Rating-Prozeduren hofften SIGALL & PAGE mit einer Untersuchung zu einem Gegenstand zu erbringen, der gegen Antworttendenzen der sozialen Erwünschtheit als besonders anfällig gelten kann: Vorurteile gegenüber Farbigen. 60 weiße Amerikaner wurden nach Zufall auf eine von vier Versuchsbedingungen aufgeteilt. Sie sollten angeben, wie charakteristisch jede von 22 Eigenschaften

entweder für "Amerikaner" oder für "Neger" sei. Die eine Hälfte der Vpn antwortete auf normalen Ratingskalen ohne besonderen Versuch der Kontrolle von Social-Desirability-Tendenzen, die andere Hälfte antwortete unter BPL-Bedingungen. Die 2x2-Varianzanalyse lieferte einen signifikanten Haupteffekt für den Faktor "Ethnische Gruppe" und - wie erwartet - eine signifikante Wechselwirkung zwischen den Faktoren "Ethnische Gruppe" und "Meßmethode". Die Effektivität der BPL-Versuchsanordnung, wie sie sich in solchen signifikanten Wechselwirkungen zeigt, ergab sich bei sechs der 22 Eigenschaftsbegriffe: ehrenhaft, intelligent, dumm, schmutzig, sensibel, faul; bei vieren dieser Begriffe hatte es keinen Haupteffekt gegeben. Weitere deskriptive Analysen zeigten, daß der mittlere Grad an Erwünschtheit der den beiden Gruppen zugeordneten Stereotypen unter Rating-Bedingungen fast gleich war (0.53 für Amerikaner, 0.49 für Neger), während unter BPL-Bedingungen nur die Amerikaner sozial erwünscht stereotypisiert wurden (0.84 für Amerikaner, -0.03 für Neger). SIGALL & PAGE halten nach diesen Ergebnissen das BPL-Paradigma für eine aussichtsreiche Möglichkeit, wahrheitsgemäßen Antworten näherzukommen, schließen jedoch alternative Erklärungen für die gefundenen Resultate - näher diskutiert bei JONES & SIGALL (1971)-für eine Interpretation nicht aus.

JONES & SIGALL (1971): A New Paradigm

In ihrem Aufsatz im Psychological Bulletin, mit dem die Autoren das BPL-Paradigma als neue Methode der relativ verzerrungsfreien Einstellungsmessung vorstellen und über die bisherigen Resultate mit dem "EMG"-Gerät und seinen Vorläufern informieren, diskutieren sie u.a. mögliche alternative Erklärungen für den BPL-Effekt: Einmal könnte man davon ausgehen, daß es für die Vp schwierig ist, Einstellungsurteile bzw. Bewertungen von Meinungsgegenständen auf eine, und nur eine Art und Weise zu geben, d.h. Meinungen und Bewertungen mögen von Zeit zu Zeit und von Situation zu Situation schwanken; dies könnte dazu führen, daß normalerweise eher zu streng als zu mild geurteilt wird, d.h. daß eher "negative" als "positive" Fehler gemacht werden. Wird man also irgendwann einmal unter Druck gesetzt, die "wirkliche" Wahrheit zu sagen, so urteilt man womöglich eher etwas zu schlecht über den Meinungsgegenstand. Mit dieser Antworttendenz ließen sich eventuelle die stark negativen Urteile über Farbige unter BPL-Bedingungen erklären, wenn man miteinbezieht, daß die günstigere Einschät-

zung von Amerikanern zugleich eine negativere Beurteilung von "Nicht-Amerikanern" bedeutet. Eine zweite alternative Deutung des BPL-Effekts zielt auf den "Mehrkomponentenansatz" der Einstellung ab: Geht man davon aus, daß Einstellungen aus affektiven und kognitiven Komponenten bestehen, so könnte es sein, daß das BPL-Gerät die affektive Einstellungskomponente gegenüber der kognitiven in den Vordergrund rückt bzw. "salient" macht. Verweist man nämlich die Vp auf Reaktionen ihres autonomen Nervensystems, so appelliert man ja gewissermaßen an ihre primitivere, affektive Seite. Daher könnte es sich bei BPL-Antworten stärker um Reaktionen unter Ausschaltung differenzierender kognitiver Gesichtspunkte handeln. - Gegenüber solchen alternativen Überlegungen bevorzugen die Autoren jedoch die Interpretation der Bogus Pipeline als einer Art Lügendetektion mit dem Ergebnis verringerter Effekte der sozialen Erwünschtheit.

SIGALL & PAGE (1972): Interpersonal Affekt Via the Bogus Pipeline

Weibliche Vpn wurden einzeln, jeweils zusammen mit einer angeblich weiteren, männlichen Vp in den Warteraum eines psychologischen Laboratoriums gesetzt; in der Hälfte der Fälle war die männliche Person durch Fußschiene deutlich als behindert zu erkennen. Als kurze Zeit später das Telefon in dem Raum klingelte, griff sich jeweils die männliche Person den Hörer und antwortete, wobei sie sich gegenüber dem Fernsprechteilnehmer, der sich offensichtlich verwählt hatte, äußerst gehässig benahm. Im Anschluß an diese Szene bat der V1 die Vpn, in einem kleinen Nebenraum einen Einstellungsfragebogen zu fünf unverfänglichen Themen (Kino, Musik etc.) auszufüllen; geantwortet wurde auf einer Sieben-Punkte-Skala. Die konföderierte Vp kopierte dabei heimlich die Antworten der echten Vp. Sodann wurde die Vp an die EMG-Apparatur gesetzt und bekam die übliche BPL-Insturktion. Zur angeblichen Festlegung ihrer Grundrate wurde sie gebeten, die Items auf ihrem Meinungsfragebogen durchzugehen, und der V1 stellte die Übereinstimmung zwischen Fragebogen und "EMG-Output" fest. Dabei verwendete er die heimlich mitgeschriebenen Vp-Antworten, und so gab das EMG in allen Fällen akkurate Antworten. Bei dem folgenden eigentlichen Versuch wurden einige Items beurteilt, unter denen sich auch eine Beurteilung der anderen "Vp" befand; die Vp sollte sagen, wie sehr sie ihren Partner persönlich mochte, und sie sollte das Ausmaß ihrer Übereinstimmung mit verschiedenen, ihm zuschreibbaren Eigenschaften nennen.

Eine Kontrollgruppe von Vpn unterzog sich den gleichen Aufgaben bei abgeschalteter EMG-Apparatur, indem sie das Steuerrad lediglich als Mittel zur Anzeige ihrer Meinung auf der Ratingskala verwendete. Der Rest der Versuchsanordnung wurde als Datenverarbeitungsgerät deklariert; Elektroden waren nicht zu sehen.

Zunächst ergab sich, daß alle Vpn von den Fähigkeiten des "EMG" überzeugt waren. Wie erwartet urteilten sie unter BPL-Instruktion über ihren Versuchspartner ungünstiger als beim normalen Rating; die mittleren Werte der Eigenschafts-Ratings unterschieden sich von den mittleren BPL-Reaktionen im Vorzeichen. Auch wurden unter BPL-Bedingungen dem männlichen Partner ausgesprochen negative Eigenschaften zugesprochen, nicht aber beim Rating. Die Erwartung, daß man Hemmungen habe, eine offensichtlich behinderte Person negativ einzuschätzen, traf jedoch nur der Tendenz nach ein.

Ein weiteres Experiment sollte zeigen, daß die BPL mehr als nur negative Reaktionen ans Licht zu bringen vermag. Weibliche Studienanfänger wurden wiederum auf die Zellen eines 2x2-Versuchsplans aufgeteilt. Man erklärte ihnen, daß sie mit einer Polaroid-Kamera fotografiert würden und daß das Foto zusammen mit einem von ihr ausgefüllten Fragebogen an eine zweite Person des anderen Geschlechts geleitet würde. Sie selbst würde die entsprechenden Daten der anderen Person erhalten und solle diese dann beurteilen. Danach würden beide dann zu einer tatsächlichen Interaktion zusammengebracht und man wolle sehen, wie sich ihre dyadische Beziehung entwickle. Variiert wurden die beiden Faktoren "Meßmethode" (BPL vs. Rating) und "Attraktivität der männlichen Person" (äußerlich attraktiv und dem Fragebogen nach kompetent vs. unattraktiv und stupide), abhängige Variable war das geäußerte Interesse der weiblichen Person, mit der männlichen Person zu interagieren. Dieses erwies sich als am größten in der Attraktiv/BPL-Bedingung (1.67), und am geringsten in der Unattraktiv/BPL-Bedingung (-1.78). Beim einfachen Rating gab es dagegen in jeder Hinsicht niedrigere bzw. moderatere Werte. Die statistische Wechselwirkung zwischen beiden Faktoren erwies sich als signifikant. Entsprechende Resultate ergaben sich auch für das Ausmaß des Mögens bei der direkten Beurteilung der männlichen Person. Auch hinsichtlich weiterer Personbeurteilungen gab es die extremsten Urteile unter BPL-Instruktion. Die Bogus Pipeline erwies sich damit gegenüber gewöhnlichen Ratingverfahren als überlegen, wenn es darum ging, extreme Urteile zu erhalten, seien diese nun positiv oder negativ.

JONES & WEIN (1972): Attitude Similarity, Expectancy Violation, and Attraction

Zur Erforschung von Reihenfolge-Effekten bei der Personwahrnehmung wandten die Autoren sowohl das BPL-Paradigma als auch eine übliche Rating-Technik an, und zwar im Anschluß an die Untersuchungen von JONES, BELL & ARONSON (1972), die mit einer BPL-ähnlichen Versuchsanordnung arbeiteten. In einem ersten Experiment sollten weibliche Vpn auf 20 heterogene Meinungsgegenstände antworten, indem sie ihr Urteil auf einer 21-Punkte-Skala abgaben und sodann jedes Item mit ihrer Antwort auf ein Tonband sprachen; die Tonbandansagen wurden jedoch über ein Intercom mitgehört und zu zwei fingierten Antwortmustern verarbeitet: Die eine Antwortserie bestand zunächst aus Antworten, die denjenigen der Vp sehr ähnlich waren, so daß die Vp, wenn sie sie hören würde, annehmen müßte, es handle sich um Antworten einer Person mit den ihren sehr ähnlichen Einstellungen. Nach der ersten Hälfte der Items wichen die Antworten jedoch von denjenigen der Person mehr und mehr ab, so daß die Ähnlichkeits-Erwartung verletzt wurde. Bei dem zweiten Antwortmuster handelte es sich um eine Umkehrung des ersten, d.h. zunächst vorliegende "Unähnlichkeit" wurde von "Ähnlichkeit" abgelöst. Nachdem die Vp alle Items beantwortet hatte, brachte ihr der VI das manipulierte Tonband der angeblichen anderen Person, spielte ihr fünf Items mit Antworten vor und bat sie, vorherzusagen, wie diese Person auf die nächsten fünf Items antworten würde. Dann wurden die nächsten fünf Items mit Antworten vorgespielt, wiederum wurde eine Vorhersage erbeten, usw. Nachdem alle Vorhersagen gemacht worden waren, befestigte der VI die BPL-Elektroden am Arm der Vp und bat sie, sich sorgfältig auf ihre Gefühle gegenüber der anderen Person zu konzentrieren, während er ihr einige Fragen stellen würde. Die erste und wichtigste Frage lautete: "Wie sehr mögen Sie persönlich diese Person?" Geantwortet wurde wieder auf einer 21-Punkte-Skala; das "EMG"-Gerät war mit einer entsprechenden Skala versehen worden. Die Hypothese der Untersuchung, wonach Vpn solche Personen stärker mögen, die zunächst unähnliche, dann aber mehr und mehr ähnliche Einstellungen zeigen, ließ sich eindeutig bestätigen. Damit wurde unter BPL-Bedingungen eine Hypothese auf dem Gebiet der interpersonalen Attraktionsforschung bestätigt, die z.B. BYRNE & LONDON (1966) mit Standard-Rating-Methoden nicht verifizieren konnten.

In einem zweiten Experiment ohne BPL-Instruktion wurden - wie bei BYRNE & LONDON - keinerlei Unterschiede zwischen den verschiedenen Reihenfolge-

Bedingungen der Erwartungs-Verletzung in bezug auf ähnliche/unähnliche Einstellungen gefunden.

In einem dritten Experiment, das dem direkten Vergleich beider Urteilsbedingungen diene, wurden weibliche Vpn gemäß einem 2x2-Design auf die Bedingungen der Faktoren "Reihenfolge" und "Meßmethode" aufgeteilt. Es zeigte sich, daß Zielpersonen der Unähnlich/Ähnlich-Sequenz nur unter BPL-Bedingungen stärker positiv eingeschätzt wurden. Die statistische Wechselwirkung zwischen Reihenfolge und Meßmethode entsprach einer Restwahrscheinlichkeit von 0.07. Innerhalb der BPL-Bedingung war der sequentielle Effekt auf dem 5%-Niveau gesichert. Als Erklärungen der Resultate wurden angeboten: Entweder zwingt die BPL-Situation die Vp dazu, sich stärker auf ihre Gefühle als auf ihren Verstand zu konzentrieren, oder unter BPL-Bedingungen werden grundsätzlich stärker negative Gefühle geäußert, oder die BPL-Versuchsanordnung bringt die Vpn dazu, das gesamte Experiment ernster zu nehmen und insofern die Merkmale der Stimulusperson bei der Eindrucksbildung stärker zu beobachten.

OSTROM (1973): A New "ignis fatuus"?

In seinem grundsätzlich kritischen Artikel zur Bogus Pipeline im Psychological Bulletin fragt OSTROM, ob es sich beim BPL-Paradigma nicht um ein neues ignis fatuus handele:

"At night over swampy ground there occasionally appears a phosphorescent light that behaves in a mysterious manner. Sometimes it hovers and sometimes it flits about; and when one approaches it, the light just seems to disappear" (p.252).

In der Geschichte der Einstellungsmessung habe es oft solche Irrlichter gegeben: Immer dann, wenn man glaubte, eine zufriedenstellende Meßtechnik zu besitzen, sei sie einem entglitten. Bei der Enttarnung des BPL-Paradigmas als ein solches Irrlicht bzw. bei seiner kritischen Analyse der von JONES und Mitarbeitern bis dahin erbrachten Belege für die Leistungsfähigkeit bzw. Überlegenheit des BPL-Paradigmas als einer allgemeinen Rahmensituation für die Erfassung subjektiver Urteile ist OSTROM gezwungen, sich auf einen Vergleich mit gewöhnlichen Ratingverfahren zu beschränken, denn nur hierzu waren bislang Belege vorgebracht worden. Der Kritiker weist darauf hin, daß es in dem Experiment von JONES, BELL & ARONSON (1972) Korrelationen zwischen Pipeline-Situation und Papier-Bleistift-Ratings zwischen 0.66 und 0.81 gegeben habe, die man als zufriedenstellend ansehen könne. Weiteren Experimenten von JONES und SIGALL

mit 2x2-Versuchsplänen schreibt OSTROM wenig Bedeutung zu, weil bei der Prüfung der statistischen Wechselwirkung die 5%-Marke knapp verfehlt worden sei. Auch den nicht aufgetretenen Haupteffekt in SIGALL & PAGEs (1972) Experiment mit der stigmatisierten und unverschämten Vp legt der Kritiker zuungunsten der BPL-Versuchsanordnung aus.

Nach OSTROMs Analyse sind Ratings unter BPL-Instruktion und solche unter normalen Papier-Bleistift-Bedingungen funktional durchaus vergleichbar, und die mögliche größere Sensibilität der BPL fiel danach nicht sonderlich ins Gewicht. Eine vorgenommene Feinanalyse des Experiments von SIGALL & PAGE (1971) zur Erwünschtheit von Persönlichkeitseigenschaften, die man Schwarzen und Weißen zuschreibt, ergab, daß die BPL-Ergebnisse den bereits früher mit ordinären Ratings gefundenen ähnlicher sind, als dies die BPL-Befürworter wahrhaben möchten.

Ein zweiter Einwand gegen das BPL-Paradigma bezieht sich auf ethische und praktische Aspekte der Angelegenheit. OSTROM unterstreicht die ethische Verpflichtung, in Humanexperimenten dann, wenn mehrere vergleichbare Meßmethoden vorliegen, diejenige zu wählen, die mit dem geringeren Maß an Täuschung der Vp auszukommen vermag. Sowohl bei der Anwendung der BPL im Zusammenhang mit klinisch relevanten (z.B. MMPI-) Daten als auch bei gewöhnlichen experimentalpsychologischen Untersuchungen werde die Selbstachtung der untersuchten Person angegriffen, wenn sie - spätestens bei der Enttarnung (debriefing) am Schluß des Versuchs - erkennen müsse, daß ihr persönlicher Schutzschild durchbrochen worden sei. OSTROM räumt allerdings ein, daß eine sorgfältig durchgeführte Enttarnungsprozedur auch positive Effekte zeitigen könne. Auf der anderen Seite weist er auf die enorme Unpraktikabilität der gesamten Versuchsanordnung hin. Es ergäben sich Probleme bei der Konstruktion des Gerätes, bei seinem Transport und bei der Untersuchung von mehr als einer Vp zum gleichen Zeitpunkt; auch die Bemühungen, die Apparatur erst einmal glaubwürdig zu machen, seien sehr zeitaufwendig. Selbst wenn man aber einmal den Aufwand toleriere, bliebe doch das Problem, daß der Täuschungscharakter des BPL-Paradigmas in kurzer Zeit weithin bekannt werde und es dann schwierig sei, noch naive Vpn zu finden. Man müsse sich nur ausmalen, was geschehe, wenn nicht primär wissenschaftliche Magazine wie - hier einmal übertragen - "Der Spiegel" oder "psychologie heute" auch ein solches Verfahren durch Publikation zunichte machten.

Schließlich wird den BPL-Protagonisten vorgeworfen, ihre Methode stets mit relativ anfechtbaren, keineswegs aber den bestmöglichen Ratingverfahren verglichen zu haben. Ein alternatives Verfahren zur Meinungs- und Einstel-

lungsmessung sollte aber stets mit dem besten existierenden Verfahren verglichen werden.

Ungeachtet also der möglichen Überlegenheit der BPL in bezug auf die Sensibilität, gefühlsmäßige Bewertungen von Vpn zu erfassen, schneidet das Verfahren nach OSTROM in bezug auf funktionale Vergleichbarkeit, ethische Aspekte und Praktikabilität zu schlecht ab, um eine methodische Alternative zur herkömmlichen Einstellungserfassung abzugeben.

JONES & SIGALL (1973): Where There Is Ignis, There May Be Fire

In ihrer Antwort auf OSTROM verteidigen JONES & SIGALL die Bogus Pipeline nicht als Allzwecksubstitut für konventionelle Ratingskalen, sondern als Verfahren, das in ganz bestimmten Situationen Effekte der sozialen Erwünschtheit (social desirability) vermindere bzw. die affektive Einstellungskomponente stärker zum Vorschein bringe. Es liege nicht in ihrer Absicht, so betonen sie, die BPL als routinemäßig in der Einstellungsforschung anzuwendendes Instrument anzupreisen. Sie stimmen auch mit OSTROM darin überein, daß sich die aufgewiesenen Unterschiede zwischen herkömmlichem Rating und BPL-Rating nicht rigoros interpretieren lassen. Andererseits bezweifeln sie, daß man zwischen der Sensibilität eines Verfahrens und seiner funktionalen Vergleichbarkeit mit anderen Verfahren scharf trennen könne. Daß das BPL-Paradigma gewisse ethische und praktische Probleme aufwerfe, wird auch von JONES & SIGALL anerkannt. Das Eindringen in die Privatsphäre der Vp sei jedoch nicht erheblich größer als etwa bei der Anwendung projektiver Testverfahren oder bei anderen Persönlichkeitsmessungen, deren Interpretation der Vp nicht zugänglich gemacht werde. Die Erfahrung aus ihren eigenen Versuchen zeige im übrigen, daß sich nur sehr wenige Vpn bei der Aufklärung bzw. Enttarnung am Schluß des Versuchs geärgert hätten. Auch die von OSTROM geäußerte Befürchtung, Vpn könnten trotz gegenteiliger Aufklärung an die Fähigkeit der Maschine, ihre wahren Einstellungen zu erfassen, weiterhin glauben und die Aufklärung als solche nicht akzeptieren, wird von JONES & SIGALL zurückgewiesen. Es sei viel leichter einzusehen, daß eine entsprechende Fähigkeit eines psychologischen Gerätes in Wirklichkeit gar nicht existiere als das Gegenteil zu glauben. Das Argument der Unpraktikabilität könne nur dann greifen, wenn beispielsweise jeder Sozialpsychologe sein eigenes EMG konstruiere und Untersuchungen mit dieser Versuchsanordnung die Universitäten überschwemmen. Stattdessen bekräftigen JONES & SIGALL ihre Auffassung, daß die BPL nur dann angewendet werden solle, wenn es um die Untersuchung besonders wichtiger theoretischer Gegenstände gehe, wenn die Einstellungen

von Vpn sich ihrer Messung auf andere Art und Weise widersetzen oder wenn der Versuchsleiter explizit auf unterschiedliche Einstellungskomponenten abhebe. Eine Überbetonung des Aspekts der Praktikabilität von Versuchsanordnungen würde nach Auffassung der die Pipeline verteidigenden Autoren ebenfalls Untersuchungen zur Konformität, zum Self-Esteem oder zum Gehorsamsverhalten verhindert haben. Schließlich stimmen die Autoren mit OSTROM darin überein, daß weitere Untersuchungen mit der BPL notwendig seien. In der Aufforderung zu "weiterer Forschung" durch den Kritiker erblicken sie jedoch einen Widerspruch zu dessen Hervorhebung ethischer und praktischer Unzulänglichkeiten der BPL-Methode. Insgesamt sind sie der Auffassung, daß bereits in ihren bisherigen Untersuchungen wichtige und interpretierbare Unterschiede zwischen Ratingskalen und BPL-Schätzungen nachgewiesen werden konnten.

PAVLOS (1973): Self-Esteem-Effekte

Zwar hatten sich bis dahin unter BPL-Bedingungen beispielsweise extremere Einstellungsäußerungen gegenüber Personen unterschiedlicher Hautfarbe aufweisen und im Sinne verminderter Reaktionstendenzen der sozialen Erwünschtheit interpretieren lassen. Dennoch fehlte bislang eine direkte Prüfung des Anspruchs daß die BPL Social-desirability-Einflüsse verringert. PAVLOS prüfte daher die Annahme, daß SD-Effekte die rassistischen Einstellungsreaktionen von Vpn dann nicht beeinflussen, wenn diese unter BPL-Bedingungen erhoben werden. Der Autor nahm an, daß Personen, die experimentell einer Verringerung ihres Self-Esteem ausgesetzt würden, weniger Vorurteile gegenüber Schwarzen äußern würden, verglichen mit Personen, deren Self-Esteem man künstlich erhöht hat. Unter BPL-Bedingungen dürften solche Unterschiede dagegen nicht zutage treten. Die Prüfung erfolgte in einem 2x2-Versuchsplan mit den Faktoren "Self-Esteem" und "Meßmethode", wobei die Self-Esteem-Manipulation über angebliche Computer-Persönlichkeitsanalysen vorgenommen wurde; abhängige Variable war das Ausmaß der Ablehnung von Farbigen. Ein postexperimentell applizierter Fragebogen erwies zunächst die vollständige Glaubwürdigkeit der EMG-Apparatur. Unter BPL-Bedingungen wurden insgesamt negativere Einstellungen geäußert als im normalen Rating. Es ergab sich ein Haupteffekt für "Self-Esteem" sowie eine auf dem 5%-Niveau gesicherte Interaktion zwischen Selbstwertschätzung und Meßmethode. Die Ergebnisse ließen den Schluß zu, daß vermehrt positive rassistische Einstellungsäußerungen bei Personen mit verringertem Self-Esteem nur unter normalen Rating-Bedingungen auftreten; unter BPL-Bedingungen setzt sich

dieser Social-desirability-Effekt dagegen nicht ins Spiel. Der Autor interpretiert dieses Resultat als Unterstützung von JONES & SIGALL gegen die Gegenkritik von OSTRUM (1973).

BRIGHAM, BLOOM, GUNN & TOROK (1974): Attitude Measurement via the Bogus Pipeline: A Dry Well?

BRIGHAM et al. fassen die bei Einstellungsmessungen auftretenden Reaktions-tendenzen, die JONES & SIGALL mithilfe des BPL-Paradigmas schwächen möch-ten, als zwei Spielarten der "evaluation apprehension" (ROSENBERG 1965) auf: response constraint (Erlebnis eines Reaktions-Druckes, auf sozial erwünschte Art und Weise zu antworten) und accuracy constraint (Erlebnis eines Exaktheits-Druckes, beim Antworten korrekt und glaubwürdig zu er-scheinen). Während sich erstere Tendenz den Autoren zufolge sowohl beim normalen Rating als auch beim Rating unter BPL-Bedingungen auswirkt, er-scheint die zweite nur unter BPL-Bedingungen als zusätzlich erhöht. Drei Untersuchungen am Gegenstand der von JONES & SIGALL eingeführten Einstel-lungen gegenüber Farbigen sollten Unterschiede zwischen verschiedenen Meß-verfahren hinsichtlich beider Arten erlebten Drucks oder Zwanges aufwei-sen:

Experiment I prüfte die Hypothese, daß bei erhöhtem Korrektheitsdruck das Ausmaß positiver Einschätzung von Farbigen durch Weiße abnehmen müsse. Zusätzlich wurde ein vermittelnder Einfluß der Untersuchungssituation er-wartet (Klassenraum vs. individuelle Untersuchung). "Geringer" Genauig-keitsdruck wurde hergestellt, indem die Vp ihre Ratings durch Drehen des Zeigers auf der 7-Punkte-Anzeigetafel und lautes Ansagen der Zeigerposi-tion gegenüber dem VI abgab; dies entsprach der Nicht-EMG-Bedingung bei SIGALL & PAGE. Ein "mittlerer" Genauigkeitsdruck wurde erzeugt, indem die übliche BPL-Bedingung repliziert wurde (Anlegen der Elektroden, Erläu-terung des "EMG", "Validierung" des Gerätes, usw.). "Hoher" Genauigkeits-druck wurde hergestellt, indem zusätzlich zur "mittleren" Bedingung die Wichtigkeit exakter Urteile besonders betont und nach tatsächlich erreich-ter Genauigkeit Prämien vergeben wurden (maximal etwas über 1 Dollar bei "optimaler", d.h. übereinstimmender Beurteilung aller 22 Eigenschaften). Tatsächlich wurden allen Vpn 95 Cents ausgezahlt. Zusätzlich wurde eine Papier-Bleistift-Bedingung in Gruppensituation (im Klassenraum) einge-führt, wie sie den meisten üblichen Einstellungsmeßvorgängen entsprechen dürfte. Wie erwartet ergab sich, daß bei "hohem" Genauigkeitsdruck die ne-gativsten, bei "mittlerem" mittlere, und bei "niedrigem" Korrektheitsdruck

die positivsten rassistischen Einstellungsurteile abgegeben wurden. Die Differenz zwischen den beiden letzteren Versuchsbedingungen, also denjenigen von SIGALL & PAGE (1971), war jedoch nicht signifikant - ansonsten stimmten die Ergebnisse mit der Richtung der Überlegungen dieser Autoren überein. In der Gruppensituation urteilten die Vpn ähnlich wie unter den beiden letzten Bedingungen. Herausgehoben war also lediglich die Bedingung "erhöhten" Genauigkeitsdruckes, der durch zusätzliche Belohnung erzeugt worden war.

Experiment II sollte mehr Klarheit in das Zusammenspiel von Genauigkeitsdruck (accuracy constraint) und Druck in Richtung auf soziale Erwünschtheit (response constraint) bringen. In einem 2x4-Versuchsplan wurden die Faktoren "VI-Kenntnis der Vp-Anworten" (Reaktionsdruck) und "Angenommene Erfahrbarkeit der Antwortgenauigkeit" (Exaktheitsdruck) variiert. Varianzanalysen für jede der betrachteten 22 Eigenschaftsbegriffe erbrachten nur in drei Fällen interpretierbare Resultate und wurden daher nicht weiter diskutiert. Einflüsse der beiden untersuchten Arten erlebten Antwort-Druckes konnten daher nicht festgestellt werden. Daraus wurde geschlossen, daß BPL-Effekte, wenn sie tatsächlich existieren, nicht in allen BPL-Situationen auftreten.

In Experiment III sollte nun die BPL-Situation selbst näher erforscht werden. Sieben Situations-Bedingungen wurden unterschieden: 1) eine anonyme Gruppensituation, ohne Beachtung der EMG-Versuchsanordnung 2) eine anonyme Testsituation gleicher Art, jedoch bei individueller Untersuchung jeder Vp im Labor 3) eine entsprechende Situation, jedoch mit Einführung der Erwartung der Vpn, daß sie später noch einmal auf die gleichen Items zu antworten hätten 4) Beantwortung des Fragebogens, nachdem die Vpn nur von der Existenz des BPL-Gerätes gehört haben 5) entsprechende Situation, wobei die EMG-Apparatur jedoch gezeigt wird, ohne daß sie in Betrieb genommen wird 6) vollständige BPL- bzw. EMG-Situation, wobei die Vp ihre Ratings im Papier-Bleistift-Verfahren abgibt 7) vollständige BPL- bzw. EMG-Situation, mit Ratings auf der EMG-Anzeigetafel, d.h. Drehen des Steuerrades. In dieser Untersuchung wurden 35 Statements auf 60-Punkte-Skalen beantwortet; sie betrafen außer Stellungnahmen zu Farbigen alle möglichen persönlichen, teilweise auch sehr persönliche Gegenstände. Die Varianzanalyse ergab zunächst keinen Effekt für den Unterschied zwischen allen sieben experimentellen Bedingungen. Dies galt sowohl für die einzelnen Items getrennt als auch für faktoranalytisierte Itemgruppen. Teilte man die Instruktionsbedingungen jedoch in solche ohne EMG-Manipulation (1 bis 3) und solche mit EMG-Manipulation (4 bis 7), so ergab sich für die ersten 20 Items ein signifikanter

Unterschied zwischen den experimentellen Bedingungen. Trennte man die Items nach verschiedenen Inhaltsgruppen, so ergaben sich allein für den Bereich "sexuelle Phantasien" signifikante Unterschiede zwischen BPL- und Nicht-BPL-Bedingungen - unter BPL-Bedingungen gaben die Vpn eher zu, ungewöhnliche sexuelle Phantasien zu haben.

Die Resultate der drei Experimente werden von BRIGHAM et al. als "bemerkenswert konsistent" hinsichtlich des Fehlens bedeutsamer BPL-Effekte angesehen. Als eine mögliche Erklärung nennen sie die relative Unglaubwürdigkeit der Versuchsanordnung bei den heutigen "sophisticated" Collegestudenten, ein Argument, das auch OSTROM (1973) verwendet hatte. Gegen diese Interpretation sprechen jedoch sämtliche postexperimentellen Befragungsergebnisse, die immer wieder - auch im vorliegenden Falle - die Glaubwürdigkeit der Apparatur erhärten. Eine weitere Erklärungsmöglichkeit wird darin gesehen, daß die eingeführten Konzepte, nämlich Reaktions- und Genauigkeitsdruck, für Einstellungsmessungen irrelevant sein können. Dies wird aber als sehr unwahrscheinlich angesehen; für Gebiete, auf denen starke soziale Normen existieren, kann man nach Auffassung der Autoren davon ausgehen, daß beide Tendenzen - in entgegengesetzter Richtung - wirksam seien. Schließlich wird darauf hingewiesen, daß der zu erwartende Reaktionsdruck, d.h. der erlebte Zwang, sozial erwünschte Antworten zu geben, in den Untersuchungen vielleicht nicht kräftig genug gewesen sei. Vielleicht sei auch die Tendenz sich möglichst nicht von einer Apparatur überwachen zu lassen, nicht so stark wie gewöhnlich angenommen werde. Denn nur durch finanzielle Zuwendungen (Bedingung 3 in Experiment I) konnte man Personen dazu bringen, signifikant "genauer" zu urteilen. Als Resümé schließen sich BRIGHAM und Mitarbeiter also eher der Kritik OSTROMs als den Auffassungen der BPL-Verfechter an und schätzen die Vorteile der neuen Technik gegenüber den herkömmlichen Verfahren eher als gering ein.

PAGE & MOSS (1975): Attitude Similarity and Attraction

Bekanntlich wurde wiederholt gefunden, daß wahrgenommene Einstellungsähnlichkeit zwischen Personen zu hoher interpersonaler Attraktion führt, wahrgenommene Einstellungs-Unähnlichkeit jedoch nur zu einem mäßigen Grad an Nicht-Attraktivität bzw. Nicht-Mögen einer Person. Diese Asymmetrie könnte das Ergebnis eines Social-Desirability-Effekts sein: Da es sozial unerwünscht ist, jemanden negativ zu bewerten, treten bei wahrgenommener Einstellungsunähnlichkeit nur mäßige negative Bewertungen auf. PAGE & MOSS

variierten in einem 2x2-Versuchsplan die Faktoren "Ähnlichkeit einer fremden Person" (0% bzw. 100% Ähnlichkeit) und "Meßmethode" (BPL bzw. normales Rating). Es gab zwei abhängige Variablen, nämlich die fremde Person "zu mögen" und mit ihr "zusammen arbeiten" zu wollen. Ermittelte man ein gemeinsames Maß dieser beiden Personbeurteilungen, so ergab die Varianzanalyse nur einen Haupteffekt für den Faktor "Ähnlichkeit". Untersuchte man jedoch die beiden Aspekte getrennt, so ergab sich für das Maß des "Mögens" außerdem noch eine signifikante Interaktion zwischen "Ähnlichkeit" und "Meßmethode". Analysierte man die Daten näher, so zeigte sich, daß die Personen unter BPL-Bedingungen einen absolut unähnlichen Fremden signifikant negativer als unter Rating-Bedingungen beurteilten, während die Vpn einen zu 100% ähnlichen Fremden unter BPL-Bedingungen und beim normalen Rating nicht unterschiedlich einschätzten. In bezug auf das "Mögen" einer Person traten damit in erwarteter Weise unter Einsatz des "EMG" symmetrischere Attraktionsverhältnisse auf. Die Untersuchung kann daher als partieller Beleg für die Effektivität der Pipeline beim Durchbrechen von Effekten der sozialen Erwünschtheit angesehen werden.

CHERRY, BYRNE & MITCHELL (1976): Clogs in the Bogus Pipeline

Auch diese Arbeit verglich die BPL mit der herkömmlichen Ratingmethode im Rahmen eines etablierten experimentellen Paradigmas, wobei die Autoren jedoch von vornherein nicht von der Brauchbarkeit der BPL-Versuchsanordnung überzeugt waren. Wie bei PAGE & MOSS wurde von der als gut belegt anzusehenden funktionalen Beziehung zwischen der Ähnlichkeit von Einstellungen und der interpersonalen Attraktion ausgegangen. Es wurde jedoch hier das Konzept der "experimenter demands" eingeführt, und es sollten entsprechende Anforderungscharakteristika (demand characteristics) manipuliert werden. In einem 2x2x2x3-Design wurden (der Reihenfolge nach) folgende Faktoren untersucht: 1) der Anteil ähnlicher Einstellungen 2) das Meßverfahren (BPL bzw. Rating) 3) der Grad der sozialen Erwünschtheit 4) experimentell manipulierte Anforderungscharakteristika. Die Treatments der beiden ersten Faktoren wurden wie in entsprechenden Untersuchungen üblich hergestellt; die SD-Maße wurden durch die Antworten auf der SD-Skala von CROWN/MARLOWE erhoben, und das Maß an Demand Characteristics wurde durch die Antworten auf die Frage bestimmt, welchen Zweck das Experiment wohl habe. Es zeigte sich ein signifikanter Haupteffekt für den Faktor "Einstellungsähnlichkeit" sowie eine signifikante vierfache Wechselwirkung. Eine nähere Analyse dieser Wechselwirkung ergab, daß die Attraktions-Reaktionen der Vpn mit hohem Ausmaß an sozialer Erwünschtheit nur unter BPL-Bedingungen vom Grad der

Anforderungscharakteristika abhängen; die BPL-Methode wird daher als anfällig gegenüber starken Demand Characteristics bei solchen Personen angesehen, welche die Tendenz haben, sozial erwünschte Antworten zu geben. CHERRY et al. weisen aufgrund dieser Ergebnisse alle Annahmen, die auf eine Überlegenheit der BPL gegenüber normalen Rating-techniken bei der Erfassung interpersoneller Attraktion hinauslaufen, zurück. Auch hinsichtlich Sensibilität sehen sie keine besonderen Vorteile der BPL. Zudem sei eine besondere Schwäche des BPL-Paradigmas deutlich geworden: Macht man Anforderungscharakteristika deutlich (salient), indem man Fragen über mögliche der Untersuchung zugrundeliegende Hypothesen einführt, so erweist sich das EMG bei Personen mit hohen SD-Werten sogar als weniger geeignet zur Erfassung von Gefühlen als normale Ratingprozeduren. Die Autoren warnen daher vor einer Anwendung der BPL-Methode in all denjenigen Fällen, in denen ihre Überlegenheit über herkömmliche Verfahren nicht explizit empirisch nachgewiesen ist.

SCHLENKER, BONOMA, HUTCHINSON & BURNS (1976): A Second Researcher's Pipedream

In einer weiteren Replikation der Untersuchung von SIGALL & PAGE, die 22 Eigenschafts-Ratings von "Farbigen" betreffend, wurde ein differenzierter BPL-Effekt gefunden. Angewendet wurde ein 3x2x2-Versuchsplan mit den Faktoren "Meßmethode" (normale BPL-Bedingung / BPL-Bedingung mit Schwächung der Social-Desirability-Komponente / Kontrollbedingung bzw. Rating), "Geschlecht der Vp" und "Geschlecht des V1". Eine Schwächung der SD-Komponente in einer modifizierten BPL-Situation sollte dadurch erreicht werden, daß den Vpn suggeriert wurde, unbewußte Einstellungen und Gefühle hätten keinerlei Einfluß auf den Ausschlag des Gerätes: "So, just report what you believe your true feelings are and don't try to outguess yourself" (p.324). Es ergaben sich jedoch nur in drei von 22 Fällen ("musical", "ostentatious", "aggressive") signifikant methodenabhängige Ratings von Farbigen. Faktoranalytisierte man die Eigenschaften, so daß sich drei Faktoren von Farbigen-Stereotypen ergaben (ein "30er-Jahre-Stereotyp: Schwarze als ignorant und schmutzig", ein "Sambo-Stereotyp: Schwarze als musikalisch und happy-go-lucky", und ein "modisches Stereotyp: Schwarze als ehrbar und intelligent"), so zeigten sich bezüglich des zweiten ("Sambo"-) Faktors interpretierbare BPL-Effekte. Vpn der Kontrollbedingung schrieben den Sambo-Faktor Farbigen weniger zu als Vpn, die unter den beiden BPL-Bedingungen urteilten; zwischen letzteren gab es keine Unterschiede.

Damit wurde abermals eine Teilbestätigung für die Ergebnisse von SIGALL & PAGE vorgelegt. Da die Validität der Bogus Pipeline jedoch nur partiell und nur relativ zu sein scheint, folgern SCHLENKER et al., daß man sich gut überlegen müsse, ob die Anwendung der BPL-Versuchsanordnung den Aufwand lohne. "For most problems, relative differences between groups of people are all that are required, and the bogus pipeline does not appear to affect these. One researcher's bogus pipeline could well be a second researcher's pipedream" (p.328).

GAES, QUIGLEY-FERNANDEZ & TEDESCHI (1978): Unclogging the Bogus Pipeline

In dieser Arbeit wird die methodische Verfahrensweise von CHERRY et al. (1976), die eine Invalidität der BPL konstatierten, rigoros angegriffen. Es wird aufzuzeigen versucht, daß eine Post-hoc-Aufteilung von Vpn gemäß Social-Desirability-Fragebogenscores zu keiner schlüssigen Aussage über die Wirkweise von sozialen Erwünschtheitstendenzen führen könne. Der Verdacht, der Versuchsplan von CHERRY et al. sei nachträglich manipuliert worden, da eine Post-hoc-Aufteilung nach SD, die in insgesamt 12 Zellen des Versuchsplanes zur genauen Teilung der Vpn führte, eine Unwahrscheinlichkeit von astronomischen Ausmaßen besäße, konnte durch eine nachträgliche Information des Herausgebers der betreffenden Zeitschrift, wonach CHERRY und Mitarbeiter ungleiche Zellen in Kauf genommen hätten, ausgeräumt werden. Dennoch halten GAES et al. die SD-Aufteilung für den Schlüssel der seinerzeit vorgefundenen vierfachen Wechselwirkung und deren weiterer Interpretation: "Any strong inference must be based on a preexperimental partition and random assignment procedure" (p.191). Eine Sekundäranalyse der Daten von CHERRY et al. durch die Autoren ergab außerdem eine Ineffektivität des Faktors "Anforderungsmerkmale" (demand characteristics).

Um nicht nur negative Kritik an den Kritikern der Bogus Pipeline zu üben, berichten GAES et al. schließlich noch über zwei unpublizierte eigene Arbeiten, in denen es darum ging, die Brauchbarkeit der BPL beim Vergleich einander widersprechender Theorien aufzuweisen: Da die Autoren die in einer Forced-Compliance-Situation auftretenden Einstellungsänderungen eher für das Ergebnis einer Taktik von Vpn halten, einen guten Eindruck von sich zu erzeugen oder wiederherzustellen (impression management theory), als für das Ergebnis der Reduktion kognitiver Dissonanz im Sinne der Dissonanztheorie, verglichen sie BPL- und Rating-Messungen in einer solchen Forced-Compliance-Situation. Ein Dissonanzeffekt fand

sich lediglich unter Papier-Bleistift-Bedingungen, nicht jedoch unter BPL-Bedingungen. In der BPL-Situation müßten die Vpn nach Auffassung der Autoren annehmen, ihr Lügen würde entdeckt. Der Impression-Management-Theorie entspricht es daher, daß es zu Meinungsänderungen nur in einer Meß-Situation kommt, in der Lügen unentdeckbar zu sein scheinen (GAES, RIVERA & TEDESCHI). In einer weiteren Untersuchung wurde angenommen, daß Personen, die in nicht-equitabler Weise überbelohnt worden sind, öffentlich Schuldgefühle zeigen, insgeheim aber Freude empfinden. Eine Messung des Ausmaßes an Zufriedenheit und Schuldgefühlen unter BPL- und normalen Rating-Bedingungen ergab, daß Personen im normalen Rating umso mehr Schuld empfanden, je ungerechter bzw. höher ihr Gewinn war, daß sie aber größere Zufriedenheit und weniger Schuldgefühle zeigten, wenn sie unter BPL-Bedingungen antworten mußten (RIVERA & TEDESCHI). GAES et al. sehen danach im Gegensatz zu den Kritikern der Pipeline nicht nur die Validität der BPL als erwiesen an, sondern sie glauben auch, daß das Vertrauen in diese Methodik umso stärker zunehmen werde, je häufiger sie in derartigen theorietestenden Untersuchungen eingesetzt werden könne.

GAES, KALLE & TEDESCHI (1978): Impression Management in the Forced Compliance Situation

Ausgehend von der für diese Autoren nun feststehenden Tatsache, daß die meisten Bedenken OSTROMs (1973) gegen die Validität des BPL-Verfahrens ausgeräumt seien, so daß man die Methode zum Testen kontroverser sozialpsychologischer Theorien verwenden könne und solle, wird ein Experimentum Crucis zwischen Dissonanz- und Impression-Management-Theorie ausgeführt. In einer Forced-Compliance-Situation wurden einmal die Art und Weise des einstellungskonträren Verhaltens (öffentlich oder anonym), zum anderen die Art der Einstellungsmessung (BPL oder Papier-Bleistift-Verfahren) variiert. Kritische Variable des 2x2-Designs war die Antwort zu einem Statement, das die Schädlichkeit des Zähneputzens betonte. Es ergaben sich keinerlei Haupteffekte, jedoch eine signifikante Wechselwirkung zwischen "Öffentlichkeit" und "Meßmethode". Unter derjenigen Bedingungskombination, die nach dissonanztheoretischer Auffassung und der von den Dissonanztheoretikern bevorzugten Vorgehensweise zur Einstellungsänderung führen müßte, nämlich der Kombination "Öffentlich/Papier-Bleistift", wurde tatsächlich die einstellungskonträre Position am stärksten vertreten. Das Nichtauftreten entsprechender Meinungsänderungen unter BPL-Bedingungen, das der Dissonanztheorie zuwiderläuft, ließ sich in ter-

mini der Impression-Management-Theorie erklären. Auch eine Reihe weiterer, hier nicht berichteter Daten entsprach diesem Bild. Sämtliche Ergebnisse und Interpretationen der Autoren erwiesen sich dann als stimmig, wenn man die Annahme macht, daß die Vpn sich am BPL-Gerät stärker genötigt - und insofern auch unfreier - fühlten, die Wahrheit zu sagen, als in einer einfachen Antwortsituation.

In einem zweiten Experiment variierten die Autoren die Abfolge der angewendeten Meßmethoden (BPL bzw. Papier-Bleistift-Verfahren) für den Fall mehrmaliger Einstellungsmessung. Vermutet wurde, daß Personen, die ein erstes Mal eine Einstellungsänderung gezeigt hatten, also "gelogen" hatten, ihre Meinung beibehalten würden, auch wenn bei der zweiten Messung die BPL an der Reihe wäre. Für die Reihenfolge Papier-Bleistift/Bogus Pipeline wurde also in der Forced-Compliance-Situation keine Meinungsdivergenz zwischen beiden Messungen erwartet. Auf der anderen Seite wäre auch für Personen, die anfangs "nicht gelogen" haben, die also zuerst einer BPL-Messung unterzogen wurden, kein Anlaß gegeben, beim zweiten Mal eine Einstellungsänderung zu zeigen. Die Ergebnisse waren die folgenden: Wurde zuerst mit Papier-Bleistift-Verfahren gemessen, so ergab sich tatsächlich eine Einstellungsänderung, die weiter beibehalten wurde; wurde zuerst unter BPL-Bedingungen gemessen, so ergab sich keine Einstellungsänderung, und es ergab sich auch bei der zweiten Messung keine solche. Da die Autoren keine durchschlagenden Belege für eine mangelnde Sensibilität der Bogus Pipeline, Meinungsänderungen aufzudecken, finden mögen, interpretieren sie die Ergebnisse als experimentellen Beleg für das Zutreffen der Impression-Management-Theorie und halten das BPL-Paradigma weiterhin für hilfreich bei der Entscheidung zwischen konkurrierenden Theorien, sofern diese auf subjektive Urteile von Personen aufgebaut sind.

QUIGLEY-FERNANDEZ & TEDESCHI (1978): The Bogus Pipeline as a Lie Detector

Zwar hatten TEDESCHI und seine Mitarbeiter bereits gezeigt, wie "Dissonanz-Effekte" unter BPL-Bedingungen verschwinden, und sie hatten somit aus ihrer Sicht Validitätsbelege für das BPL-Verfahren gefunden, doch strebten sie nach einer noch direkteren Zurückweisung der OSTROMSchen Kritik. So versorgten sie in zwei weiteren Experimenten Vpn mit illegitimen Informationen, aus denen die Vpn im Experiment Vorteile beziehen konnten, und sie befragten die Vpn später darüber, ob sie solche Informationen gehabt hätten. Wenn dann unter BPL-Bedingungen mehr "Geständ-

nisse" aufträten als bei einer normalen Befragung mit einem Papier-Bleistift-Verfahren, so wäre darin eine unmittelbare Validierung der BPL im Sinne der Lügendetektion gegeben. In der Tat gaben unter BPL-Bedingungen insgesamt 37%, unter Kontrollbedingungen insgesamt 8% der Personen zu, über illegitime Informationen verfügt zu haben. Bei den Kontrollbedingungen handelte es sich außer schriftlicher Befragung noch um eine Face-to-face-Anfrage. Im Rahmen der bei TEDESCHI und Mitarbeitern üblichen postexperimentellen Befragungen gaben mehr Personen an, unter BPL-Bedingungen verärgert gewesen zu sein; entweder ruft also die BPL-Versuchsanordnung mehr Ärgernisse hervor als dies bei alternativen Verfahrensweisen der Fall ist, oder die BPL ist eher in der Lage, Verärgerung seitens der Vpn, wie sie bei allen psychologischen Untersuchungen solcher und ähnlicher Art auftritt, tatsächlich ans Licht zu bringen.

In einem zweiten Experiment erhöhte man das Ausmaß der Abhängigkeit der Vp von der illegitimen Information, die durch einen VI-Gehilfen verschafft wurde. Von dieser und einigen weiteren kleineren Korrekturen gegenüber dem ersten Experiment erwarteten die Autoren eine höhere "Geständnis"-Rate als beim ersten Experiment. Wiederum zeigten sich unter BPL-Bedingungen deutlich mehr Geständnisse als unter der vergleichsweise anspruchsvollen Bedingung der Face-to-face-Befragung. Je direkter die Nachfrage bezüglich der Verfügung über illegitime Informationen war, desto mehr Geständnisse unter BPL-Bedingungen gab es. Zusammengefaßt waren es 65% der Vpn, die unter BPL-Bedingungen gestanden, nur 5% dagegen bei einfacher Befragung. Unter den verschiedenen, zusätzlich kontrollierten Bedingungen ist vielleicht noch das Geschlecht von VI und Vp von Interesse: Es hatte keinen erheblichen Einfluß auf das Ausmaß "geständigen" Verhaltens - allenfalls neigten Frauen eher dazu, nichts zuzugeben. Den Autoren erscheint als wichtigstes Resultat, daß die Überwindung sozial erwünschter Antworttendenzen durch die Anwendung der BPL-Methode von der Direktheit bzw. Spezifität der gestellten Fragen abhängt; mehrdeutige und verschwommene Fragen hätten auch am "EMG" eine Chance, verfälscht beantwortet zu werden.

Acht Jahre Bogus Pipeline

Eine Würdigung des BPL-Paradigmas acht Jahre nach den Veröffentlichungen von SIGALL & PAGE sowie JONES & SIGALL kann sich auf insgesamt 17 bekanntgewordene experimentelle Untersuchungen, publiziert in insgesamt 11 Arbeiten, stützen, wenn man vorbereitende Studien wie diejenige von JONES et al. (1972) nicht mit berücksichtigt. Von den 17 Experimenten können zwölf als erfolgreich in dem Bemühen, die Fähigkeit der BPL zur Schwächung von Social-Desirability-Effekten nachzuweisen, angesehen werden. Zwei Experimente (Experiment II bei BRIGHAM et al. sowie das Experiment von CHERRY et al.) zeigen dagegen die Untauglichkeit der BPL, während in drei Fällen (Experimente I und III von BRIGHAM et al. sowie dasjenige von SCHLENKER et al.) partielle Validitätshinweise konstatiert werden konnten. Es überwiegen also zahlenmäßig stark solche Untersuchungen, deren Ergebnisse für eine Brauchbarkeit der Pipeline sprechen. Die vorliegenden Resultate sollen unter folgenden Gesichtspunkten bewertet werden: 1) unter dem Aspekt situativer Geltungsbereiche 2) hinsichtlich der zum Vergleich herangezogenen Meßverfahren 3) unter dem Gesichtspunkt der Praktikabilität 4) mit Blick auf die möglicherweise noch verborgenen Qualitäten der BPL als experimentelle Rahmenbedingung.

1) Mittlerweile sind mehrere verschiedenen Themenbereiche sozialpsychologischer Forschung mithilfe der BPL-Versuchsanordnung bearbeitet worden. Eine Zeitlang konnte man den Eindruck gewinnen, als wollten die Forscher sich an der Frage der Zuschreibung von 22 Eigenschaften an farbige bzw. weiße Amerikaner festbeißen. Es wurden aber darüber hinaus direkte Personbeurteilungen untersucht, es wurde die Beziehung zwischen Einstellungsähnlichkeit und Attraktivität fremder Personen erforscht, es wurden Einstellungsänderungen in Force-Compliance-Situationen, Self-Esteem-Effekte, Reaktionen auf Ungerechtigkeit sowie direkte Geständnisse illegitimen Verhaltens untersucht. Die in der Mehrzahl positiven Ergebnisse geben den BPL-Befürwortern ein gewisses Recht, von der relativen Brauchbarkeit der Pipeline in einer Mehrzahl unterschiedlicher Situationen zu sprechen. Daß sämtliche referierten Untersuchungen an Studenten vorgenommen wurden, wird hier nicht erwähnt, weil dies etwa ungewöhnlich wäre, sondern deswegen, weil es das Ausmaß der Generalität der BPL-Validität zusätzlich einschränkt.

2) Auf der anderen Seite wurden fast ausschließlich Vergleiche zwischen BPL-Bedingungen und der Messung von Einstellungen und Meinungen mit gewöhnlichen Papier/Bleistift-Ratingverfahren angestellt, und es wurden

wenig Versuche sichtbar, die BPL-Methode mit gezielt verbesserten Ratings, z.B. solche unter verschärften oder immunisierenden Instruktionen - etwa in der Art der Experimente von SCHMIDT (1970) - zum Vergleich zu bringen. Von einer Überlegenheit der Pipeline kann demnach erst im Vergleich mit relativ eingrenzbaeren psychologischen Antwortverfahren gesprochen werden.

3) Daß in acht Jahren nicht mehr als die hier berichteten Untersuchungen bekannt geworden sind, möchte man als Bestätigung OSTROMS ansehen, der auf die relative Unpraktikabilität der Methode hingewiesen hatte. Auf der anderen Seite erfordert beispielsweise auch der BUSSsche Aggressionssimulator einen gewissen apparativen Aufwand und eine aufwendige Cover Story, ohne daß dies ein Hindernis gewesen wäre, in wenigen Jahren hunderte von Experimenten gemäß dem BUSS-Paradigma zu publizieren. Es verwundert in diesem Zusammenhang vielmehr, daß nicht Versuche gemacht (oder veröffentlicht) worden sind, die BPL-Versuchsordnung bzw. -Instruktion wesentlich unaufwendiger zu gestalten. Dies könnte z.B. in der Weise geschehen, daß die Prozedur der Überzeugung der Vp von der Leistungsfähigkeit des Gerätes stärker auf Suggestion als auf praktischer Demonstration aufzubauen wäre. Es würde einen oder mehrere Versuche lohnen.

4) Der eigentliche Wert der Bogus Pipeline scheint dem Referenten darin zu bestehen, daß die BPL-Versuchsordnung eine Rahmensituation konstituiert, innerhalb deren prinzipiell eine Vielzahl möglicher menschlicher Reaktionen erforschbar bzw. meßbar ist. Es wäre zu kritisieren, daß die BPL-Befürworter das Verfahren bislang nicht wesentlich über einfache Meinungs- und Einstellungsfeststellungen hinaus bei psychologischen Untersuchungen eingesetzt haben. Es ließe sich denken, daß die eigentliche BPL-Dramaturgie, also die Überzeugung der Vp von einer wirksamen Lügendetektion und darauffolgend die Vorhersage der Meßdaten des Gerätes durch die Vp, im Rahmen weiterer experimenteller und diagnostischer Forschungsprozesse einsetzbar wäre. Reaktionstendenzen vom "sozial erwünschten" Typus lassen sich ja keineswegs nur hinter Selbsteinschätzungen und sonstigen, herkömmlicherweise mit Ratings oder weiteren Papier-Bleistift-Verfahren meßbaren subjektiven Bewertungen annehmen, sondern ebenfalls bei der Erfassung offenen Verhaltens, wenn dieses Verhalten soziale Normen berührt. Aggressives

Verhalten im Experiment ließe sich vielleicht wieder leichter als Schädigung einer anderen Person, die auch vom Schädigenden selbst als Schädigung wahrgenommen und anerkannt wird, untersuchen, (so daß das Verhalten nicht umständlich verkleidet und undefiniert werden muß, daß es, um erforschbar zu werden, für die Vp alles mögliche andere, nur nicht aggressives Verhalten, bedeutet), wenn zugleich erreicht werden könnte, daß die Vp von der Meßbarkeit aggressiven Verhaltens unter BPL-Bedingungen überzeugt werden könnte und dementsprechend zum vorübergehenden Ablegen ihres Schutzschildes - zum Wohle der Grundlagenforschung selbstverständlich - bereit wäre. Further research is needed.

Literatur

- BRIGHAM, J.C., BLOOM, L.M., GUNN, S.P. & TOROK, T. Attitude measurement via the bogus pipeline: A dry well? *Representative Research in Social Psychology*, 5, 1974, 97-114.
- BYRNE, D. & LONDON, O. Primacy and recency and the sequential presentation of attitudinal stimuli. *Psychonomic Science*, 6, 1966, 193-194.
- CHERRY, F., BYRNE, D. & MITCHELL, H.E. Clogs in the bogus pipeline: Demand characteristics and social desirability. *Journal of Research in Personality*, 10, 1976, 69-75.
- COOPER, J. Personal responsibility and dissonance: The role of foreseen consequences. *Journal of Personality and Social Psychology*, 18, 1971, 354-363.
- GAES, G.G., KALLE, R.J. & TEDESCHI, J.T. Impression management in the forced compliance situation. Two studies using the bogus pipeline. *Journal of Experimental Social Psychology*, 14, 1978, 493-510.
- GAES, G.G., QUIGLEY-FERNANDEZ, B. & TEDESCHI, J.T. Unclogging the bogus pipeline: A critical reanalysis of the Cherry, Byrne, and Mitchell study. *Journal of Research in Personality*, 12, 1978, 189-192.
- JONES, E.E., BELL, L. & ARONSON, E. The reciprocation of attraction from similar and dissimilar others: A study in person perception and evaluation. In: C.G. McCLINTOCK (Ed.) *Experimental social psychology*. New York: Holt, Rinehart & Winston, 1972, 142-179.
- JONES, E.E. & SIGALL, H. The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76, 1971, 349-364.
- JONES, E.E. & SIGALL, H. Where there is ignis, there may be fire. *Psychological Bulletin*, 79, 1973, 260-262.
- JONES, E.E. & WEIN, G.A. Attitude similarity, expectancy violation, and attraction. *Journal of Experimental Social Psychology*, 8, 222-235.

- OSTROM, T.M. The bogus pipeline: A new ignis fatuus? *Psychological Bulletin*, 79, 1973, 252-259.
- PAGE, R.A. & MOSS, M.K. Attitude similarity and attraction: The effects of the bogus pipeline. *Bulletin of the Psychonomic Society*, 5, 1975, 63-65.
- PAVLOS, A.J. Acute self-esteem effects on racial attitudes measured by rating scale and bogus pipeline. *Proceedings of the 81st Annual Convention, American Psychological Association*, 1973, 165-166.
- QUIGLEY-FERNANDEZ, B. & TEDESCHI, J.T. The bogus pipeline as a lie detector: Two validity studies. *Journal of Personality and Social Psychology*, 36, 1978, 247-256.
- ROSENBERG, M.J. When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, 1, 1965, 28-42.
- SCHLENKER, B.R., BONOMA, T.V., HUTCHINSON, D. & BURNS, L. The bogus pipeline and stereotypes toward blacks. *Journal of Psychology*, 93, 1976, 319-329.
- SCHMIDT, H.D. *Posudzovanie ľudskeho správania rating-škálami*. Bratislava: Slovenské Pedagogické Nakladateľstvo 1970.
- SIGALL, H. & PAGE, R. Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology*, 18, 1971, 247-255.
- SIGALL, H. & PAGE, R. Reducing attenuation in the expression of interpersonal affect via the bogus pipeline. *Sociometry*, 35, 1972, 629-642.