

---

## Die ADV-SKALA

### Ein Instrument zur Messung von Einstellungen gegenüber der ADV\*)

Detlef B. Müller

Seminar für Allgemeine Betriebswirtschaftslehre und Betriebswirtschaftliche Planung der Universität zu Köln,  
5 Köln 41, Albert-Magnus-Platz

---

**Stichworte:** ADV-SKALA, ADV-Wert, Fragebogen, Einstellung zur ADV, Skalenkonstruktion

**Zusammenfassung:** Aus der Notwendigkeit heraus, sowohl für allgemeine empirische Untersuchungen als auch für die Planung und Implementierung computergestützter Informationssysteme Benutzereinstellungen gegenüber der ADV operational zu definieren, wurde ein Meßinstrument in Fragebogenform mit 28 Items im LIKERT-Format entwickelt (ADV-SKALA). Diese 28 Items wurden aus einem Pool von 149 Items nach folgenden Verfahren ausgewählt: – 1. Pretest (Unverständlichkeit, Schwierigkeit und Trennschärfe) – Expertenrating („Eignung“, ADV-Einstellungen zu messen) – 2. Pretest (Schwierigkeit und Trennschärfe) und an 439 Benutzern von Informationssystemen in 11 Betrieben getestet. In Anbetracht der breitgefächerten Zielgruppe weist die Skala angemessene Trennschärfe und eine hohe Stabilität auf. Mit Hilfe einer Faktorenanalyse wurden 4 Subskalen ermittelt, die inhaltlich bedeutende Aspekte der Einstellung zur ADV widerspiegeln. Durch eine Standardisierung und Normierung der ADV-SKALA und der Subskalen wurden die statistische Verwertbarkeit und die Interpretationsfähigkeit der Skala wesentlich erhöht. Die ADV-SKALA scheint somit bestens geeignet, innerhalb empirischer Untersuchungen Benutzer nach deren Einstellungsgrad zur ADV zu differenzieren.

**Key-words:** EDP-SCALE, EDP-value, Questionnaire, Attitude towards EDP, Construction of Scales

**Abstract:** Regarding the necessity to define attitudes towards EDP in an operational way for general empirical investigations as well as for the planning and implementation of computer-based information systems, a questionnaire of 28 items (EDP-SCALE) was developed applying the LIKERT-technique. The 28 items were chosen from a set of 149 items according to the following procedure: – first pretest – expert rating – second pretest, and then presented to 439 users of information systems in 11 organizations. In spite of the variance of the tested group the scale afforded sufficient correlations of the single items with the sum of all items and a high stability. By using factor analysis four subscales could be developed representing important

aspects of attitudes towards EDP. By standardization and normalization of the EDP-SCALE and its subscales the application of statistical methods and the possibilities of interpretation could be improved to a considerable degree. Thus, the EDP-SCALE seems to be rather suitable for the differentiation of users of EDP-systems according to their attitudes towards EDP.

---

### 1. Grundsätzliche Fragestellung

Menschliche Aspekte gewinnen beim Einsatz moderner Informationssysteme immer mehr an Bedeutung. Letztlich ist es häufig die Einstellung der Mitarbeiter zur ADV, die über Erfolg oder Mißerfolg des Informationssystems und damit über die Rentabilität einer weitreichenden Investition entscheidet. Zum anderen findet die Auffassung von der Zufriedenheit der Mitarbeiter als einem Ziel im Zielsystem der Unternehmung immer größere Verbreitung.

Negative Einstellungen zur ADV produzieren nicht nur Sabotage am System (durch Einsetzen falscher Zahlen, Nichteinhalten von Terminen) oder Absentismus (z. B. Flucht in die Krankheit<sup>1)</sup>), sondern auch allgemeine Arbeitsunzufriedenheit.

Wenn auch der Zusammenhang zwischen verbal geäußerten Einstellungen und tatsächlichem Verhalten umstritten ist<sup>2)</sup>, so erscheint dennoch ein Instrument zur Messung von Einstellungen gegenüber der ADV für die Planung von Informationssystemen von Bedeutung zu sein.

Ein solches Instrument könnte folgende Aufgaben erfüllen:

1. Im Rahmen empirischer Untersuchungen bedarf es zur Überprüfung von Hypothesen (z. B. „ältere Mitarbeiter sind gegenüber der ADV negativer eingestellt als jüngere“) einer Operationalisierung der angesprochenen Variablen (hier: Alter in Jahren, Einstellung wie? ).

---

\*) Für seine überaus wertvollen Anregungen zu einer früheren Fassung dieses Aufsatzes ist der Verfasser Herrn Prof. Dr. Helmut E. Lück von der Gesamthochschule Duisburg zu großem Dank verpflichtet.

2. Die Schulung als ein Teilaspekt der Systemimplementierung kann durch das Wissen über die Einstellung der „Schüler“ bedeutend effizienter gestaltet werden.
3. Der Erfolg der Implementierung in Hinsicht auf Einstellung und Zufriedenheit der Benutzer kann überprüft werden.

Ein Instrument, das in diesen Bereichen einsetzbar wäre, liegt nach Wissen des Verfassers nicht vor. Zwar wurden in einigen deutschen Untersuchungen bereits vereinzelt Fragen nach Einstellungen der Benutzer zur ADV gestellt<sup>3)</sup>, ein einheitliches Konzept ist jedoch nicht entwickelt worden. Auch im amerikanischen Raum sind Einstellungsfragen lediglich an spezielle, fallweise relevante Problembereiche geknüpft. Instrumente in Form von Skalen beziehen sich demgegenüber auf Einstellungen zur Automation allgemein<sup>4)</sup>.

Im folgenden sollen die Entwicklung eines Instruments zur Messung von Einstellungen gegenüber der ADV („ADV-SKALA“) geschildert und einige Hinweise für den Einsatz der Skala in der Praxis gegeben werden.

## 2. Anforderungen an die ADV-SKALA

Wie an jedes sozialwissenschaftliche Meßinstrument, sind auch an die ADV-SKALA eine Reihe von inhaltlichen und formalen Anforderungen zu stellen. In Hinsicht auf die Zielsetzungen ergeben sich folgende Bedingungen, die das Instrument erfüllen sollte:

### 1. Bedingung: Einfache Handhabbarkeit

Die Forderung nach einfacher Handhabbarkeit des Instrumentes impliziert zweierlei. Einmal darf die Beantwortung den Probanden nicht über Gebühr beanspruchen (zeitliche Begrenzung bei der Erhebung). Zum anderen soll das Instrument für den Anwendenden schnell und leicht auswertbar sein (zeitliche und qualitative Begrenzung bei der Auswertung). Diese Forderung wird insbesondere dann relevant, wenn das Instrument im Rahmen betrieblicher Untersuchungen von sozialpsychologischen und teststatistischen „Laien“ ausgewertet werden soll.

### 2. Bedingung: Anwendbarkeit bei Benutzern verschiedener Bereiche

Benutzer moderner Informationssysteme unterscheiden sich nach zahlreichen Aspekten, z. B. nach Fachabteilungen, Branchen, Computeranwendungen. Ein Instrument, das in allen Bereichen einsatzfähig ist, hat zwar den Nachteil, nicht konkret genug (z. B. in bezug auf Einstellungen gegenüber bestimmten Systemänderungen) zu messen, andererseits aber den Vorteil, Vergleiche zwischen Benutzern verschiedener Systeme zu ermöglichen. Es besteht hier für die Konstruktion einer Skala ein Zielkonflikt: je *eingengerter* der Benutzerkreis, desto *konkreter* die Fragestellungen und damit die mit Hilfe des Instruments zu gewinnenden Aussagen. Je konkreter diese Aussagen über die Einstellungen der Benutzer zur ADV aber sind, desto weniger *generalisierbar* sind sie im Vergleich zu Einstellungen von Benutzern anderer Systeme.

Entsprechend den oben angeführten Zielsetzungen soll mit dem Instrument ein möglichst großer Personenkreis angesprochen werden können. Eine Einschränkung muß allerdings in jedem Fall gemacht werden: Zur Zielgruppe gehören nur Benutzer der ADV innerhalb bestimmter Fachabteilungen (Personal-, Rechnungswesen, Verkauf u. ä.); ADV-Spezialisten (Systemanalytiker, Programmierer, Operateure) gehören nicht zum in Frage kommenden Personenkreis.<sup>5)</sup>

Die Forderung nach allgemeiner Anwendbarkeit ist jedoch nicht losgelöst zu sehen von der

### 3. Bedingung: Verständliche Formulierung

Je inhomogener die Zielgruppe in bezug auf Vorbildung, Aufgabengebiete, hierarchische Stellung u. a. ist, desto größere Schwierigkeiten entstehen mit der für das Instrument zu wählenden Sprache. Das größte Problem besteht darin, inwieweit die Probanden die einzelnen Fragestellungen gleichartig interpretieren, d. h. ob Homogenität in der Auslegung der Items durch die Befragten vorliegt. Formulierungen, die von zwei Befragten jeweils unterschiedlich verstanden werden, sind für eine vergleichende Untersuchung ungeeignet.

Eine zweite Schwierigkeit liegt in der unterschiedlichen Sprache von Forschern und Befragten.<sup>6)</sup>

### 4. Bedingung: Berücksichtigung verschiedener Aspekte der ADV

Allgemein kann man davon ausgehen, daß die ADV die Einstellungsstruktur der Benutzer vielschichtig berührt. Das erklärt sich ganz einfach aus der Komplexität und den tatsächlichen oder vermeintlichen Auswirkungen des Einstellungsobjektes – der ADV. Dieses Objekt wird unter den verschiedensten Gesichtspunkten gesehen: als Gefährder des Arbeitsplatzes, als Fortschritt an sich, als Vermittler rationaler Arbeitsabläufe u. a. m.

Ein Instrument zur Einstellungsmessung hat eine Reihe dieser Aspekte zu berücksichtigen und darüber hinaus die Möglichkeit zu schaffen, unterschiedliche Ausprägungen innerhalb dieser Aspekte herauszufiltern.

### 5. Bedingung: Zuverlässigkeit und Gültigkeit

Der letzte Anforderungspunkt umfaßt die beiden wichtigsten Voraussetzungen sozialwissenschaftlicher Meßinstrumente:

- die Zuverlässigkeit als Grad der Genauigkeit, mit dem ein bestimmtes Merkmal gemessen wird, gleichgültig, ob das Instrument dieses Merkmal auch zu messen beansprucht;
- die Gültigkeit als Grad der Genauigkeit, mit dem das Merkmal, das gemessen werden soll, auch tatsächlich gemessen wird.<sup>7)</sup>

Diese beiden Bedingungen sind in der Regel recht schwer zu überprüfen. In Hinsicht auf die Zuverlässigkeit beschränkt man sich meist auf die interne Konsistenz.<sup>8)</sup> Die Gültigkeit ist nicht zuletzt davon abhängig, inwieweit es gelungen ist, Sprache und Verständnis über den Objektgegenstand zwischen Forscher und Befragten in Einklang zu bringen. Insofern hängt die Gültigkeit auch weitgehend von der Verwirklichung der o. g. 3. Bedingung ab.

### 3. Die Entwicklung der ADV-SKALA

#### 3.1. Auswahl der Erhebungsmethode

Aus der sozialwissenschaftlichen Literatur sind eine Reihe von Techniken zur Messung von Einstellungen bekannt.<sup>9)</sup> Wegen der auch für Nichtfachleute leichten Auswertbarkeit (1. Bedingung) wurde auf das Verfahren der summierten Einschätzungen (summated ratings)<sup>10)</sup> zurückgegriffen. Dabei wird der Skalen- bzw. Einstellungswert aus der Addition der Einzelwerte für jede Frage bzw. Aussage (Item) ermittelt. Werden die Antworten auf die einzelnen Items z.B. mit 1 bis 5 Punkten von negativer bis zu positiver Einstellung bewertet, so bedeutet ein hoher Skalenwert eine positive, ein niedriger Skalenwert eine eher negative Gesamteinstellung gegenüber der ADV.

Diese Technik bietet weiterhin den Vorteil, einzelne Aspekte der ADV-Einstellung zu Subskalen zusammenzufassen (4. Bedingung).

Die eigentliche Skalenentwicklung wurde eng angelehnt an die Vorschläge von LIENERT zur Konstruktion von psychologischen Tests.<sup>11)</sup>

In einem ersten Schritt wurden Items gesammelt, die eine positive oder negative Bewertung der ADV bzw. der Automation im Büro wiedergeben. Folgende Quellen standen dabei zur Verfügung: (1) Literatur, (2) Gespräche mit Praktikern aus dem Hard- und Software-Bereich, (3) eigene Überlegungen. Insgesamt wurden auf diese Art und Weise 149 Items ermittelt. Der überwiegende Teil entstammte dabei amerikanischen Untersuchungen. Bei den Übersetzungen wurde insbesondere darauf geachtet, den Inhalt ansprechend zu formulieren (3. Bedingung). Um die endgültige Itemsammlung (= Skala) den o.g. Bedingungen anzunähern, wurden die Items einem mehrstufigen Selektionsverfahren unterworfen.

#### 3.2. Inhaltliche Analyse

In einem ersten Ausleseprozeß wurden Items ausgeschlossen, die inhaltlich zu weit vom Einstellungsobjekt (ADV im Betrieb) entfernt lagen. Beispiel: „Der Computer ist der Wegbereiter auf unserem Marsch ins Weltall.“ In Zweifelsfällen wurde das Item im Pool belassen. Es verblieben vorerst 111 Items.

#### 3.3. Erster Pretest

Mit diesen 111 Items wurde ein erster Pretest durchgeführt. Insgesamt 38 Beschäftigte eines Sozialversicherungsträgers der öffentlichen Hand, die insbesondere mit der Datenaufbereitung beschäftigt waren, beantworteten die Items mit Hilfe von drei Antwortvorgaben (Zustimmung – neutrale Haltung – Ablehnung). Darüber hinaus hatten sie die Möglichkeit, vom Sinn oder der Formulierung her unverständliche Items zu kennzeichnen.

Im Anschluß an diesen Pretest wurden 81 Items nach folgenden Kriterien ausgeschieden:

(1) Unverständlichkeit  $p > 10 \%$

Items, die für mehr als 10 Prozent der Befragten vom Sinn her zweifelhaft oder unverständlich erscheinen, sind für das Instrument ungeeignet. Sie entsprechen nicht der oben aufgestellten 3. Bedingung.

(2) Schwierigkeitsgrad  $15 \% < P > 85 \%$

Der Schwierigkeitsgrad gibt den prozentualen Anteil von Zustimmung- oder Ablehnungswerten für jedes Item wieder und ist damit ein Maß für die Differenzierungsfähigkeit der Frage.<sup>12)</sup> Items, denen alle (oder hier mehr als 85 %) der Befragten zustimmen, sind als Meßinstrument wenig geeignet, da sie nicht die positiv von den negativ eingestellten Befragten trennen. Anders ausgedrückt: Diese Aussage liegt so extrem, daß sowohl positiv als auch negativ Eingestellte sie ablehnen (oder ihr zustimmen).

(3) Trennschärfe  $r < 0,4$  ( $p \leq 0,1$ )

Der Trennschärfe-Koeffizient ist definiert als Korrelationskoeffizient zwischen der Antwort für jedes Item und dem Gesamtpunktwert der Skala.<sup>13)</sup> Unter der Annahme, daß der Gesamtskalenwert die Einstellung wiedergibt, zeigt der Korrelationskoeffizient somit an, inwieweit die Antwort auf ein einzelnes Item mit der Einstellung (= der Summe der Einzelantworten) übereinstimmt. Dabei besagt ein hoher Korrelationskoeffizient, daß das jeweilige Item von den Befragten im Sinne der Gesamtskala beantwortet wurde, ein niedriger (hier  $r < 0,4$ ) bedeutet, daß die Befragten mit hohem Skalenwert (also positiver Einstellung) sich in dieser Frage eher negativ zur ADV geäußert haben oder umgekehrt. Es handelt sich dabei um Items, die von den Befragten entweder nicht verstanden oder nicht im Sinne der überwiegenden Zahl der Items interpretiert wurden.

Die nunmehr verbliebenen Items gewährleisten, daß sie (zumindest von den Befragten des ersten Pretests)

- verstanden werden,
- gleichartig interpretiert werden (Homogenität der Auslegung der Items) und
- in gewissem Umfang ein einheitliches Phänomen ansprechen.

#### 3.4. Expertenrating

Gleichzeitig mit diesem teststatistischen Selektionsverfahren wurde mit den 111 Items ein Expertenrating durchgeführt. Hierzu wurden 9 Experten (5 Praktiker aus dem Bereich von Computeranwendungen und 4 Assistenten der Universität zu Köln) gebeten, die Items in bezug auf ihre „Eignung“ zu beurteilen. Unter „Eignung“ sollte verstanden werden, inwieweit die einzelnen Aussagen tatsächlich Einstellungen gegenüber der ADV wiedergeben, die nach den Erfahrungen der Experten in der Praxis von Bedeutung sind, d. h. häufig implizit oder explizit geäußert werden. Dieses Verfahren diente einer ersten Überprüfung der inhaltlichen Gültigkeit der Skala (5. Bedingung).<sup>14)</sup> Items, die von mehr als 50 % der Experten für ungeeignet gehalten wurden, unterlagen ebenfalls der Aussonderung. Davon waren

allerdings keine „neuen“ Items betroffen. Es verblieben also 30 Items, von denen nunmehr angenommen werden konnte, daß sie in bezug auf Verständlichkeit, Inhalt und formale Anforderungen bereits eine einigermaßen brauchbare Skala bilden würden.

### 3.5. Zweiter Pretest

In einem weiteren Pretest wurden diese 30 Items im Rahmen eines umfassenderen Fragebogens einer größeren Stichprobe ( $n = 58$ ) vorgelegt. Diesmal handelte es sich um Mitarbeiter einer Unternehmung der Elektroindustrie. Die Befragten des ersten Pretests waren in der öffentlichen Verwaltung tätig und hatten lediglich Input-Beziehungen zur ADV. Nunmehr sollten die überwiegend dort selektierten Items bei Mitarbeitern einer Industrieunternehmung mit sowohl Input- als auch Output-Beziehungen überprüft werden. Darüber hinaus waren die Befragten in einer Reihe von verschiedenen Fachabteilungen tätig. So konnte eine Überprüfung der Items in Hinsicht auf die 2. Bedingung durchgeführt werden.

Eine zweite Erweiterung bestand in der Auswahl der Antwortvorgaben. Nach den Erfahrungen des ersten Pretests wurden die Antwortvorgaben vereinheitlicht und ihre Zahl auf 5 erhöht.<sup>15)</sup> Inwieweit diese Antwortvorgaben geeignet sind, sollte ebenfalls Untersuchungsziel des Tests sein.

Die Ergebnisse wurden den gleichen teststatistischen Operationen unterworfen wie beim ersten Pretest. Der überwiegende Teil der Items entsprach den unter 3.3. gestellten Anforderungen in bezug auf Schwierigkeit und Trennschärfe. Das Mittel aller Trennschärfen lag bei  $r = 0,53$ , wobei der unterste Wert  $r = 0,24$ , der oberste  $r = 0,77$  betrug. Daraufhin wurden weitere 5 Items ausgesondert. Um die Skala allerdings nicht zu kurz werden zu lassen (4. Bedingung), wurden für die endgültige Fassung noch drei Items ausgewählt und der Skala hinzugefügt.<sup>16)</sup>

Eine weitere Veränderung wurde aufgrund der Antwortverteilungen notwendig. Es zeigte sich, daß eine eindeutige Kumulierung der Antworten zur Mitte hin erfolgte. Die beiden Extremantworten (stimmt genau – stimmt auf keinen Fall) wurden kaum gewählt. Es empfahl sich daher, die Antwortvorgaben an den Außenseiten in ihrem Intensitätsgrad zu entschärfen. Als neue Antwortvorgaben wurden gewählt:

„stimmt“,  
 „stimmt weitgehend“,  
 „stimmt vielleicht“,  
 „stimmt eher nicht“,  
 „stimmt nicht“.

Darüber hinaus wurden die Befragten in einem Vorspann zum Fragebogen aufgefordert, die Brandbreite der vorgegebenen Antworten voll auszunutzen und auch vor extremen Antworten nicht zurückzuschauen. Diese Maßnahmen erwiesen sich in der Hauptstichprobe als erfolgreich.

Die endgültige Skala umfaßt nunmehr 28 Items, die in erster Linie von Befragten „ausgewählt“ wurden, die zur

Zielgruppe für eine Skala zur Messung von Einstellungen gegenüber der ADV gehören (2. und 3. Bedingung). Eine einfache Auswertung ist über die Methode der summierten Schätzungen gewährleistet (1. Bedingung). Über die Berücksichtigung verschiedener Aspekte (4. Bedingung) sowie die Zuverlässigkeit und Gültigkeit (5. Bedingung) sollen erst nach der Auswertung der Hauptstichprobe endgültige Aussagen gemacht werden.

Die Herkunft der Items teilt sich, wie folgt, auf:

BERV:	11,	
KUBICEK:	5,	
SCHIEFER:	3,	
LOUBSER, FULLAN:	2,	
eigene:	7.	17)

13 Items sind positiv (Zustimmung zu diesen Items bedeutet eine positive Einstellung zur ADV), 15 negativ formuliert, um Forschungsartefakten in Form von Ja-Sager-Effekten entgegenzuwirken.<sup>18)</sup>

## 4. Eigenschaften der ADV-SKALA

In einer breit angelegten Untersuchung wurde diese Skala innerhalb eines umfangreicheren Fragebogens 443 Mitarbeitern in 11 Betrieben vorgelegt.<sup>19)</sup> In bezug auf die ADV-SKALA waren 439 Fragebögen auswertbar.

### 4.1. Der ADV-Rohpunktwert

Die Skalenrohwerte variieren zwischen 28 und 140 (Anzahl der Items multipliziert mit dem minimal und maximal möglichen Punktwert der Antworten). Das theoretische arithmetische Mittel liegt bei 84 Punkten. Für die 439 Befragten liegt der niedrigste ADV-Rohwert bei 33, der höchste bei 140. Das arithmetische Mittel liegt mit 94,9 fast 11 Punkte über dem theoretisch zu erwartenden Wert. Allerdings ist diese Tatsache unerheblich, da das theoretische arithmetische Mittel nicht als „Nullpunkt“ der Skala zu interpretieren ist.<sup>20)</sup> Vielmehr sind als Bezugspunkte die Skalenrohwerte von Vergleichsgruppen heranzuziehen.

### 4.2. Die Trennschärfen

Die Trennschärfekoeffizienten liegen zwischen  $r = 0,27$  (Item 8) und  $r = 0,70$  (Item 25). Alle Trennschärfen sind hochsignifikant auf dem 0,1-Prozent-Niveau ( $p < 0,001$ ). Die Items 5 und 8 haben verhältnismäßig niedrige Werte (vgl. Anhang). Man könnte daran denken, diese Items aus der Skala herauszunehmen. Andererseits weisen sie bei bestimmten Untergruppen recht hohe Werte aus. So erreichen Item 5  $r = 0,37$  und Item 8  $r = 0,42$  im Bereich der öffentlichen Verwaltung. In der Gruppe der Produktionsunternehmungen liegen sie mit  $r = 0,22$  (Item 5) und  $r = 0,20$  (Item 8) dagegen recht tief ( $p < 0,01$ ). Ein Vergleich der Trennschärfenmittelwerte (siehe Tabelle 1) zeigt denn auch für die Produktionsunternehmungen den schlechtesten Durchschnittswert. Das scheint in erster Linie durch unterschiedliche Erhebungsmethoden erklärbar. Während in den meisten Ver-

**Tabelle 1.** Eigenschaften der ADV-SKALA, aufgeteilt nach Subgruppen

	Gesamt	Frauen	Männer	Produktionsbetriebe	Versorgungsbetriebe	Verwaltungsbetriebe
Stichprobe	439	108	331	151	194	94
Trennschärfenmittel	0,56	0,58	0,54	0,48	0,58	0,59
Zuverlässigkeit	0,92	0,92	0,92	0,87	0,94	0,94
ADV-Rohwertmittel	94,9	89,4	96,7	96,2	97,7	86,8
Standardabweichung	21	23	20	17	21	23

**Tabelle 2.** Eigenschaften der ADV-SKALA, aufgeteilt nach Erhebungsmethoden

	Kleingruppen-test	Großgruppen-test
Stichprobe	192	247
Trennschärfenmittel	0,52	0,57
Zuverlässigkeit	0,89	0,94
ADV-Rohwertmittel	97,4	92,9
Standardabweichung	19	22

sorgungs- und Verwaltungsbetrieben eine Zusammenfassung aller Befragten in jeweils einem Raum, und damit ein Ausfüllen der Fragebögen unter jeweils gleichen Bedingungen möglich war, konnten die Befragten in einem Versorgungs- sowie allen Produktionsbetrieben die Skala nur am Arbeitsplatz selbst beantworten. Das führte häufig zu externen Einflüssen durch Unterbrechungen bei Telefonanrufen, eiligen Arbeitsaufträgen u. a. Die Werte für die Trennschärfen und die Zuverlässigkeit scheinen unter diesen Einflüssen zu leiden. Allerdings weist die Skala auch unter diesen Bedingungen mit einem Trennschärfenmittel von  $r = 0,52$  und einer Zuverlässigkeit von  $r_{tt} = 0,89$  noch durchaus akzeptable Werte auf (vgl. Tabelle 2).

**Tabelle 3.** ADV-Werte und ihre Eigenschaften

	Zuverlässigkeit	Standardmeßfehler	Konfidenzintervall, 5%-Niveau	Konfidenzintervall, 1%-Niveau	Min.	Max.
ADV-Wert	0,92	2,8	± 5,5	± 7,3	20	78
ADV-Wert 1	0,86	3,7	± 7,3	± 9,7	25	70
ADV-Wert 2	0,77	4,8	± 9,4	± 12,4	25	73
ADV-Wert 3	0,77	4,8	± 9,4	± 12,4	22	74
ADV-Wert 4	0,83	4,1	± 8,1	± 10,6	27	72

**Tabelle 4.** Korrelation zwischen der ADV-SKALA, den Subskalen und ausgesuchten Variablen

	ADV-Wert	ADV-Wert 1 Bedrohung	ADV-Wert 2 Arbeitsorganisation	ADV-Wert 3 Enttäuschung	ADV-Wert 4 Stellung
ADV-Erfahrung	0,21	0,25	0,13 <sup>1)</sup>	0,09 <sup>1)</sup>	0,19
Gehalt	0,30	0,39	0,24	0,10 <sup>2)</sup>	0,20
allg. Arbeitszufriedenheit	0,48	0,43	0,42	0,30	0,44

<sup>1)</sup>  $p \leq 0,05$

<sup>2)</sup>  $p \leq 0,01$

alle übrigen Werte hochsignifikant:  $p \leq 0,001$

### 4.3. Die Zuverlässigkeit

Die Zuverlässigkeitsuntersuchung bezieht sich lediglich auf die „innere Konsistenz“<sup>21)</sup> der Skala. Test-Retest-Zuverlässigkeiten sind im Rahmen der Einstellungsforschung ohnehin problematisch.

Die Halbierungs-Stabilität ergibt sich aus der Korrelation zwischen zwei Skalenhälften (Items mit gerader Kennziffer gegen Items mit ungerader Kennziffer), korrigiert mit Hilfe der SPEARMAN-BROWN-Formel<sup>22)</sup>. Danach ergeben sich sowohl für die Gesamtpopulation mit  $r_{tt} = 0,92$  als auch für die einzelnen Subgruppen (vgl. Tabelle 1) Werte im oberen Bereich der üblichen Anforderungen. In der Literatur liegen die entsprechenden Zuverlässigkeitskoeffizienten zwischen  $r_{tt} = 0,69$  und  $r_{tt} = 0,93$ .<sup>23)</sup>

Auch die Zuverlässigkeitswerte der (später erläuterten) Subskalen (vgl. Gliederungspunkt 4.4.) sind recht hoch. Hieraus ergeben sich sowohl niedrige Standardmeßfehler als auch, daraus resultierend, verhältnismäßig enge Konfidenzintervalle, d. h. die tatsächlich ermittelten Werte sind mit einem relativ kleinen Fehlerbereich behaftet (vgl. Tabelle 3; die angegebenen Werte beziehen sich auf die ebenfalls später behandelten ADV-Werte, vgl. Gliederungspunkt 4.6.).

#### 4.4. Subskalen

Zur Bildung von Subskalen (4. Bedingung) wurde eine Faktorenanalyse nach der Hauptkomponentenmethode mit anschließender Varimax-Rotation vorgenommen.<sup>24)</sup> Es ergaben sich 6 Faktoren mit einem Eigenwert größer als 1,0. Zur exakten Abgrenzung einer sinnvollen Anzahl von Faktoren wurden in verschiedenen Programmläufen 3, 4, 5 und 6 Faktoren vorgegeben und die Ergebnisse in bezug auf die sich bildenden Faktoren und Itemladungen untersucht. Es zeigte sich, daß ein Faktor unter den vier Bedingungen vollkommen konstant blieb, d.h. jeweils die gleichen Items auf diesem Faktor am höchsten luden. Ein zweiter Faktor blieb ebenfalls nahezu konstant bis auf ein Item, das in der 5-Faktorenstruktur auf dem zweiten Faktor lud, in der 6-Faktorenstruktur dagegen zum ersten Faktor wechselte. Bei Vorgabe von 4 und mehr Faktoren spalteten sich die restlichen, noch verbliebenen Items auf die jeweils zur Verfügung stehenden Faktoren auf. Der Verfasser entschied sich nach einer inhaltlichen Analyse der einzelnen Items für eine 4-Faktoren-Struktur. Die Faktoren haben dabei Eigenwerte von 9,0; 2,2; 1,5 und 1,2 und decken eine Gesamtvarianz von knapp 50 % (49,6).

Entsprechend den Faktoren konnten nunmehr bestimmte Items zu Subskalen zusammengefaßt werden. Aufnahmekriterium für ein Item war eine Ladungszahl von  $\geq 0,40$  auf dem jeweiligen Faktor. Die Faktoren erhielten folgende Benennungen:

##### 1. Faktor: Bedrohung

Auf diesem Faktor laden alle Items hoch, die eine Bedrohung des Aufgabeninhaltes und des Arbeitsplatzes, eine Verschlechterung des Betriebsklimas sowie eine Beeinträchtigung der Arbeitszufriedenheit ansprechen.

##### 2. Faktor: Arbeitsorganisation

Unter diesem Faktor fallen Aussagen über die Transparenz der Arbeitsvorgänge, repetitive Tätigkeiten und Einschätzung der veränderten Arbeit an sich.

##### 3. Faktor: Enttäuschung

In diesem Faktor schlagen sich nicht erfüllte Erwartungen nieder. Parallelen lassen sich zum Aspekt der Resignation („wir sind alle potentielle Opfer der Automation“) bei HOOS ziehen.<sup>25)</sup>

#### 4. Faktor: Stellung

Im letzten und vierten Faktor sind alle Items gesammelt, die eine Aussage zu den persönlichen Auswirkungen der ADV auf den Einsatz der eigenen Fähigkeiten und die Karriere machen. Insofern ist hier die persönliche Stellung des Befragten im Verhältnis zur ADV angesprochen. Die zu den Faktoren gehörenden Items sind dem Anhang zu entnehmen.

Die Tabelle 4 gibt einen Überblick über die Produkt-Moment-Korrelationen nach PEARSON zwischen der ADV-SKALA, den Subskalen und einigen Variablen wie Gehalt, Dauer der ADV-Erfahrung und allgemeiner Arbeitszufriedenheit<sup>26)</sup>. Diese Beziehungen sollen hier inhaltlich nicht weiter diskutiert werden. An den teilweise erheblich differierenden Korrelationen wird jedoch deutlich, daß über die Aufspaltung der ADV-SKALA weitere wichtige Informationen generiert werden können. Die Inter-Skalen-Korrelationen weisen denn auch auf eine angemessene inhaltliche Eigenständigkeit der Subskalen hin (siehe Tabelle 5).

Es scheint somit gelungen, eine Reihe wichtiger Aspekte der Einstellung von Benutzern gegenüber der ADV herauszufiltern und auch meßbar zu machen (4. Bedingung).

#### 4.5. Gültigkeit

Für die Gültigkeit des Meßinstruments liegen dagegen bisher weniger schlüssige „Beweise“ vor. Einige Hinweise lassen allerdings erwarten, daß die Skala den relevanten Bereich von Einstellungen der Benutzer heutiger ADV-Systeme abdeckt.

1. Die Items wurden von Experten als „geeignet“ eingestuft (expert validity).
2. Die Faktoren sprechen inhaltliche Dimensionen an, die auch in der Literatur im Vordergrund stehen (content validity).<sup>27)</sup>

Weitere Aufschlüsse über das Instrument werden Untersuchungen über die Beziehungen der Skala zu anderen Variablen erbringen.<sup>28)</sup>

#### 4.6. Die Verteilung

Die Verteilung der ADV-Rohpunktwerte ist rechtsasymmetrisch bimodal.<sup>29)</sup> Das dürfte in erster Linie

Tabelle 5. Interskalenkorrelationen

	ADV-Wert 1 Bedrohung	ADV-Wert 2 Arbeitsorganisation	ADV-Wert 3 Enttäuschung	ADV-Wert 4 Stellung
ADV-Wert 1 Bedrohung	0,82	0,80	0,80	0,84
ADV-Wert 2 Arbeitsorganisation		0,49	0,59	0,58
ADV-Wert 3 Enttäuschung			0,50	0,67
ADV-Wert 4 Stellung				0,62

alle Werte hochsignifikant ( $p \leq 0,001$ )

auf die heterogene Stichprobe zurückzuführen sein. Es unterscheiden sich sowohl das ADV-Rohwertmittel für Frauen (89,4 Punkte) von dem der Männer (96,7 Punkte) signifikant ( $t = 2,95$ ;  $p < 0,01$ ) als auch der für Verwaltungsangestellte mit 86,6 Punkten gegenüber dem Rest der Population mit 97,1 Punkten ( $t = 3,98$ ;  $p < 0,001$ ). Ein Test auf Normalverteilung (Vergleich der Schiefen und Exzesse) ergab denn auch einen signifikanten Unterschied zwischen der gefundenen Verteilung und der Normalverteilung ( $p < 0,01$ ). Das führt zu erheblichen Nachteilen bei der Verarbeitung der Daten, da ein großer Teil statistischer Methoden auf der Annahme normalverteilter Werte basiert.

Aus diesem Grund wurde mit den ADV-Rohpunktwerten über die Prozenträge eine sogenannte Flächentransformation vorgenommen. Die so erhaltenen Werte wurden auf z-Werte *standardisiert*. Anschließend wurden die z-Werte *normiert*. Die Normierung erfolgte auf T-Werte mit einem Mittelwert von  $\bar{x} = 50$  und einer Standardabweichung von  $\sigma = 10$ .<sup>30)</sup> Ebenso wurde mit den Subskalen verfahren. Die nunmehr erhaltenen Zahlenwerte sollen die Bezeichnungen *ADV-Wert* und für die Subskalen entsprechend den Faktoren *ADV-Wert, 1, 2, 3 und 4* erhalten.

Durch die in der psychologischen Testforschung übliche Normierung sind einmal Untergruppen (z. B. Mitarbeiter von Versorgungsunternehmen mit Mitarbeitern von Produktionsunternehmen) direkt vergleichbar, zum anderen kann aus den Subskalenwerten ebenfalls unmittelbar auf unterschiedliche Einstellungen von einzelnen Befragten geschlossen werden. Erhält ein Befragter zum Beispiel einen ADV-Wert von 50, einen ADV-Wert 1 von 70 sowie einen ADV-Wert 4 von 30, so können daraus direkt folgende Schlüsse gezogen werden: Seine allgemeine Einstellung zur ADV entspricht genau dem Durchschnitt. Während er sich von der ADV in bezug auf Aufgabeninhalte, Arbeitsplatz und Arbeitszufriedenheit nicht bedroht fühlt, schätzt er andererseits die Auswirkungen der ADV auf seine persönliche Karriere auch sehr gering ein.

Die Normierung der Skala auf einen gemeinsamen Mittelwert mit gleich großer Standardabweichung erhöht damit die Handhabbarkeit der Skala außerordentlich.

## 5. Praktischer Einsatz der ADV-SKALA

Auf der Grundlage der hier wiedergegebenen Daten und den bei der Erhebung gemachten Erfahrungen lassen sich für den Einsatz der ADV-SKALA in der betrieblichen Praxis einige Hinweise geben:

- In einer kurzen Einführung werden die Befragten über allgemeine Ziele der Untersuchung, Aufbau und Handhabung des Fragebogens sowie zugesicherte Anonymität unterrichtet.
- Die Beantwortung der Items bereitet keinerlei Schwierigkeiten. Bei den durchgeführten Befragungen kam es zu keiner Antwortverweigerung (lediglich 4 Fragebögen waren nicht auswertbar, da eine Seite überschlagen worden war).

- Die Zeit für das Ausfüllen der Skala ist mit ca. 5 Minuten erfreulich kurz und gestattet eine Einbeziehung der ADV-SKALA auch in umfangreichere empirische Untersuchungen.
- Zur Beantwortung sollten die Befragten einer Unternehmung (Abteilung) möglichst in einem Raum zusammengefaßt werden. Ist das nicht möglich, so liefert die ADV-SKALA dennoch brauchbare Ergebnisse.

## 6. Anhang

Aus Platzgründen werden die Items im folgenden lediglich kurz inhaltlich genannt. Interessenten, die eine ausführliche Dokumentation der ADV-SKALA wünschen, wenden sich bitte an den Verfasser.

	+/-	F	r
1. Unerwünschte Nebeneffekte	-	3	0,47
2. Tätigkeit ohne jede EDV *)	-	4	0,69
3. Persönliche Vorteile	+	4	0,58
4. EDV häufig unwirtschaftlich	-	3	0,55
5. Viele Arbeiten ohne EDV nicht möglich	+	-	0,29
6. Arbeitslosigkeit	-	1	0,49
7. EDVA sollten begrenzt werden	-	3	0,62
8. EDV gleich Fortschritt	+	-	0,27
9. Erleichterung für Angestellte	+	4	0,61
10. EDV schafft kaum Probleme	+	2	0,47
11. Bewahrung vor langweiliger Arbeit	+	2	0,48
12. Eigene Fähigkeiten	+	4	0,60
13. EDV mehr einsetzen	+	2	0,54
14. Entwicklung zu schnell	-	3	0,43
15. Unaufhaltsame Ausbreitung	-	3	0,64
16. Abhängigkeit von Maschinen	-	1	0,59
17. Abwertung von Kenntnissen	-	1	0,63
18. Weniger Befriedigung	-	1	0,69
19. Durchsichtige Arbeitsvorgänge	+	2	0,51
20. Verdrängung vom Arbeitsplatz	-	1	0,61
21. Mehr Ordnung	+	2	0,61
22. Unpersönliches Arbeitsklima	-	1	0,64
23. Einengung des Entscheidungsspielraums	-	1	0,58
24. Interessantere Arbeit	+	2	0,61
25. EDV vermissen	+	4	0,70
26. Höhere Erwartungen	-	3	0,52
27. Arbeiten mit Zahlen	-	-	0,49
28. Verbesserung der Verwaltungsarbeit	+	2	0,66

+/- = positive/negative Formulierung; F = Faktor;  
r = Trennschärfe

\*) Im Fragebogen wird anstelle des exakten Begriffs „ADV“ der allgemein verständlichere „EDV“ benutzt.

## Anmerkungen

- 1) Vgl. z. B. HOOS, Ida R.: Automation im Büro, Frankfurt 1966, S. 22 und 96 ff.; HERSHEY, Gerald L.: „Human“ management in a mechanized office, in: Administrative Management, 34, 1973, 4, S. 81–83.
- 2) Siehe BENNINGHAUS, Hans: Soziale Einstellung und soziales Verhalten. Zur Kritik des Attitüdenkonzepts, in: ALBRECHT, G. u. a. (Hrsg.): Soziologie, Festschrift f. R. König z. 65. Geburtstag, Köln u. Opladen 1972, S. 671–707.
- 3) Z. B. bei SCHIEFER, Friedrich: Elektronische Datenverarbeitung und Angestellte. Das Eindringen der elektronischen Datenverarbeitung in die Büroarbeit mit seinen Auswirkungen auf die Orientierung der Angestellten, Diss. Köln 1968; und KUBICEK, Herbert: Informationstechnologie und organisatorische Regelungen, Berlin 1975.
- 4) ROSENBERG, Jerry M.: Perception of automation issues, worker background, and interpersonal behavior, Ph. D. dissertation, New York University 1962; HOPPE, Ronald A., BERV, Elliot J.: Measurement of attitudes toward automation, in: Personnel Psychology, 20, 1967, S. 65–70.
- 5) Diese Definition von „Benutzer“ folgt der Abgrenzung von GROCHLA, Erwin, MELLER, Friedrich: Datenverarbeitung in der Unternehmung 1, Grundlagen, Reinbeck 1974, S. 90 f.
- 6) Zum Problem und zu seiner Lösung vgl. MAYNTZ, Renate, HOLM, Kurt, HÜBNER, Peter: Einführung in die Methoden der empirischen Soziologie, Köln u. Opladen 1969, S. 107; NOELLE, Elisabeth: Umfragen in der Massengesellschaft, Reinbeck 1971, S. 55 f.; ERBSLÖH, Eberhard: Interview, Stuttgart 1972, S. 39 ff.
- 7) LIENERT, Gustav A., Testaufbau und Testanalyse, 3. Aufl., Weinheim u. a. 1969, S. 14 und 16.
- 8) Typisch hierfür sind die Testhalbierungsmethoden; vgl. LIENERT, Gustav A., a.a.O., S. 219 ff.
- 9) Vgl. z. B. den Überblick bei SELLTITZ, Claire, JAHODA, Marie, DEUTSCH, Morton, COOK, Stuart: Attitude scaling, in: JAHODA, Marie, WARREN, Neil (Ed.): Attitudes, Harmondsworth 1966, S. 305–324.
- 10) Das Verfahren wird allgemein LIKERT, Rensis: A technique for the measurement of attitudes, in: Archives Psychology, 1932, 140, S. 1–55, zugeschrieben, allerdings bereits 1931 von ALLPORT, G. W., VERNON, P.E.: A test for personal values, in: Journal of abnormal and social Psychology, 26, 1931, S. 231–248, erwähnt.
- 11) LIENERT, Gustav A., a.a.O., insbesondere S. 37 ff.
- 12) Vgl. LIENERT, Gustav A., a.a.O., S. 87 ff. Der Begriff entstammt der Testpsychologie und bezieht sich insbesondere auf Leistungsaufgaben, bei denen es in der Tat „leichtere“ und „schwierigere“ Aufgaben gibt. Für Einstellungsfragen läßt sich dieser Begriff nur noch in übertragenem Sinn benutzen.
- 13) LIENERT, Gustav A., a.a.O., S. 93.
- 14) MAYNTZ, Renate, HOLM, Kurt, HÜBNER, Peter, a.a.O., S. 66, nennen dieses Verfahren „expert validity“, schreiben ihm allerdings keine große Beweiskraft zu. Hier wurde es nicht so sehr als Abschlußuntersuchung, sondern vielmehr als Methode zur Verbesserung der Gültigkeit eingesetzt.
- 15) Diese Anzahl entspricht dem üblichen Modus der sog. LIKERT-SKALA. Als Antwortvorgaben fungierten: stimmt genau – stimmt weitgehend – stimmt vielleicht – stimmt eher nicht – stimmt auf keinen Fall.
- 16) Diese Items erwiesen sich – wie die Hauptstichprobe zeigen sollte – als durchaus brauchbar.
- 17) BERV, Elliot J.: The measurement of attitudes toward automation, unpublished Master's Thesis, Miami University, Oxford, Ohio 1965; KUBICEK, Herbert, a.a.O.; SCHIEFER, Friedrich, a.a.O.; LOUBSER, Jan J., FULLAN, Michael: Industrial conversion and workers' attitudes to change in different industries, in: Task Force on Labour Relations Study No. 12, Ottawa 1969.
- 18) BUNGARD, Walter, LÜCK, Helmut E.: Forschungsartefakte und nicht-reaktive Meßverfahren. Stuttgart 1974, S. 61 f.
- 19) Die Untersuchung ist Grundlage für das Dissertationsprojekt des Verfassers.
- 20) SÜLLWOLD, Fritz: Theorie und Methodik der Einstellungsmessung, in: Handbuch der Psychologie, Bd. 7, 1. Halbband, Göttingen 1969, S. 475–514, hier S. 499 f.
- 21) LIENERT, Gustav A., a.a.O., S. 234.
- 22) ebenda, S. 221.
- 23) Vgl. EDWARDS, A. L.: Techniques of attitude scale construction, New York 1957, S. 161; LIENERT bezeichnet den häufig geforderten Wert von  $r_{tt} = 0,95$  als zu hoch; LIENERT, Gustav A., a.a.O., S. 242.
- 24) Alle Rechnungen wurden am Rechenzentrum der Universität zu Köln auf einer CYBER 76 mit Hilfe des Programmpaketes SPSS durchgeführt. NIE, Norman H.; BENT, Dale H.; HULL, C. Hadlai: SPSS. Statistical package for the social sciences, New York u. a. 1970.
- 25) HOOS, Ida R., a.a.O., S. 88 ff.
- 26) gemessen mit der SAZ-Kurzskala; FISCHER, L., LÜCK, Helmut E.: Entwicklung einer Skala zur Messung von Arbeitszufriedenheit (SAZ), in: Psychologie und Praxis, 16, 1972, S. 64–76.
- 27) JÄGGI, Urs, WIEDEMANN, Herbert: Der Angestellte im automatisierten Büro, Stuttgart 1963, S. 73 ff. und 149 ff.; HOOS, Ida R., a.a.O.; S. 48 ff.; HANSEN, Robert: Dafür oder dagegen? Einstellung des Personals gegenüber der elektronischen Datenverarbeitung, in: Bürotechnik und Organisation, 19, 1969, S. 47–48; WIEDEMANN, Herbert: Das Unternehmen in der Evolution, Neuwied und Berlin 1971, S. 95.
- 28) Dafür können einmal Ergebnisse der Hauptstichprobe herangezogen werden, zum anderen sind aber auch zusätzliche Untersuchungen in anderen Gruppen wünschenswert.
- 29) Siehe LIENERT, Gustav A., a.a.O., S. 176.
- 30) Zu den hier nur angedeuteten Verfahren muß auf die Literatur verwiesen werden. Vgl. z. B. GUTJAHR, Walter: Die Messung psychischer Eigenschaften, Berlin 1971, S. 159 ff.; MAGNUSSON, David: Testtheorie, Wien 1969, S. 245 ff.; LIENERT, Gustav A., a.a.O.; S. 339 ff.