

# Planung und Durchführung von schriftpsychologischen Untersuchungen

Teut Wallner

Methodenentwicklung und Kontrolluntersuchungen im Bereich der Schriftpsychologie können nur dann zu brauchbaren Ergebnissen führen, wenn eine Reihe von Voraussetzungen erfüllt ist. Diese variieren jedoch mit der Fragestellung. Es können folglich keine Untersuchungsmodelle vorgelegt werden, die sich generell anwenden lassen. Solche Modelle wären nicht einmal wünschenswert, da sie leicht zu eingleisiger Forschung führen könnten. Es wäre weiterhin ein hoffnungsloses Unterfangen, wollte man alle denkbaren Schwächen und möglichen Fehlerquellen bei Vorbereitung, Durchführung und Darstellung von schriftpsychologischen Untersuchungen aufzählen und analysieren. Es können daher nur allgemeine Feststellungen zu kritischen Untersuchungsmomenten getroffen werden. So banal diese manchmal erscheinen mögen: Ein Blick in die Fachliteratur zeigt, dass sie immer noch nicht oder nicht hinreichend betrachtet werden. Es wird sich ferner zeigen, dass bei vielen Vorschlägen nicht eindeutig von *š*kannö, *š*sollö oder *š*mußö gesprochen werden kann. Was bei der einen Fragestellung eine unabdingbare Forderung ist, kann bei einer anderen zu einem relativ belanglosen Detail werden und vice versa. Auch die Reihenfolge der hier dargestellten Punkte muss als unverbindlich betrachtet werden, da auch hier die jeweilige Aufgabenstellung verschiedene Gewichte setzt.

In der gewählten Dreiteilung der Probleminventur (Vorbereitung, Durchführung, Berichterstattung) nimmt das Kapitel über die Vorbereitung den grössten Raum ein. Das ist kein Zufall: Umsichtige und systematische Planung erleichtert nicht nur die Durchführung einer Untersuchung, sie kann auch wesentlich zu ihrem Gelingen beitragen.

Die Literaturangaben im Text weisen nur auf wenige Beispiele hin. Wegen des Umfangs der Fragestellung muss auf Vollständigkeit verzichtet werden. Deshalb konnte auch nicht jede Feststellung durch Literaturhinweise belegt werden. Nähere Auskünfte bietet die Spezialliteratur.

Vorbeugend sei schliesslich zu bedenken gegeben, dass die folgenden Anmerkungen, Vor- und Ratschläge auf Beobachtungen und Erfahrungen des Verfassers beruhen, und dass man in einigen Fragen durchaus verschiedener Meinung sein kann. Über den Wert einer Reihe der vorgetragenen Empfehlungen kann letztlich nur Experiment und Kontrollexperiment entscheiden. Immerhin könnte die kritische Beachtung der folgenden Gesichtspunkte zu einer in der Schriftpsychologie dringend notwendigen methodologischen Präzisierung führen.

## 1. Vorbereitung

### 1.1. Fragestellungen und Untersuchungsziele

Wer Rückschlüsse aus der Handschrift auf die Persönlichkeit ihres Urhebers ziehen will, muss auf folgende drei Fragen wissenschaftlich begründete, exakt beweisbare und überprüfbare Antworten geben können:

- Ist die Handschrift ein hinreichend kontantes Phänomen und typisch für den Schreiber?
- Können die Elemente der Handschrift ö auf die die Persönlichkeitsdiagnose aufbaut ö mit Sicherheit registriert werden und variieren diese Elemente oder Variablen hinreichend unabhängig voneinander?
- Welche diagnostische Valenz haben die Variablen der Handschrift?

Die meisten schriftpsychologischen Untersuchungen befassen sich dementsprechend mit einer oder mehreren der drei Problemkategorien: Stabilität, Reliabilität und Validität.

#### 1.1.1. Stabilität der graphischen Variablen oder der schriftpsychologischen Aussagen

Die Fragen, auf die man Antworten heischt, können beispielsweise lauten\_

- Welchen Einfluss haben verschiedene Schreibgeräte, Drogen usw. auf die Handschrift ein und desselben Schreibers?
- Welche Handschriftenvariablen variieren mit der Schreibsituation (Schönschrift, forciertes Diktat, Aufzeichnung für den Hausgebrauch), welche Variablen sind relativ konstant?
- Wie stabil sind die Beurteilungen ein und desselben Gutachters an verschiedenen, gleichzeitig angefertigten Schriftproben ein und desselben Schreibers?

Als Beispiel für derartige Untersuchungen neuern Datums sei *auf Hofsommer, Holdsworth und Seifert (1965)* verwiesen.

Stabilitätsuntersuchungen graphischer Variablen, für die keine anderen Kriterien zur Verfügung stehen, setzen stets die Ermittlung der Reliabilität (Zuverlässigkeit) voraus.

#### 1.1.2. Reliabilität der graphischen Variablen oder der schriftpsychologischen Aussagen

1.1.2.1 Reliabilitätsuntersuchungen werden vorwiegend durchgeführt, wenn man wissen will, mit welcher *Sicherheit graphische Variablen* identifiziert und eingestuft werden können.

Es wurde ö und wird leider immer noch ö a priori vorausgesetzt, dass sich Handschriftenvariablen mit absoluter Sicherheit erfassen lassen. Alle bisher durchgeführten Kontrolluntersuchungen haben jedoch ergeben, dass dies durchaus nicht der Fall ist. Eine aufschlussreiche Zusammenstellung von Reliabilitätsdaten findet sich bei *Schneevoigt (1968)*.

Als Daumenregel gilt nach den bisher vorliegenden Daten, dass die Reliabilität mit der Komplexität einer Variablen sinkt. Je unzuverlässiger die Erfassung einer graphischen Variable ist, umso unzuverlässiger wird damit auch die auf sie aufbauende psychologische Ausdeutung.

Es gibt aber immer viele graphische Variablen, die noch niemals auf ihre Reliabilität untersucht worden sind, denen aber ex cathedra zentrale diagnostische Bedeutung zugesprochen wird, und die deshalb in der Praxis vielfach vorbehaltlos ausgewertet werden. Aussagen über die diagnostische Valenz von graphischen Variablen, deren Reliabilität nicht bekannt ist, sind jedoch reine Hypothesen. Es kann daher mit gutem Recht gefordert werden, dass bei jeder Untersuchung, die mit graphischen Variablen durchgeführt

wird, stets die Reliabilität dieser Variablen ermittelt oder zumindest auf früher ermittelte Werte hingewiesen wird.

Über die Feststellung der Reliabilität hinaus muss untersucht werden, ob die graphischen Variablen unabhängig voneinander variieren. Bestehen mehr als zufällige Zusammenhänge, gilt es, die der Zielsetzung angemessenen Schritte (z. B. Eliminierung von Variablen aus der Untersuchung) zu unternehmen. §Synonyme§ Variablen etwa würden die Ergebnisse anschliessender Untersuchungen oft eher verwirren.

1.1.2.2 Nach der *Übereinstimmung von schriftpsychologischen Aussagen* wurde bisher selten gefragt. Auch sie wurde stets als selbstverständlich vorausgesetzt. Wiederum haben Kontrollexperimente gezeigt, dass dies durchaus nicht in dem Masse der Fall ist, wie man behauptet hat (Wallner, 1965, 1969[a]). Der Wert derartiger Reliabilitätsuntersuchungen ist jedoch recht begrenzt. Sie zeigen nur, ob und in welchem Ausmass Gutachter bei Beurteilungen des gleichen Schriftmaterials zu gleichen Ergebnissen kommen. Etwaige Differenzen können bei genauer Kenntnis der Gutachterqualifikation gedeutet werden. Über die Gültigkeit der Aussagen lassen sich selbst sei sehr grossen Übereinstimmungen keinerlei Schlüsse ziehen. Hohe Übereinstimmungen beweisen nur, dass mehrere Gutachter die gleiche Methode auf die gleiche Weise angewendet haben.

### 1.1.3. Validität der schriftpsychologischen Aussagen

Die Herleitung eines Zusammenhangs zwischen Handschriftenvariable oder Variabelsyndrom und Persönlichkeit des Schreibers bleibt so lange Arbeitshypothese, wie ihre Gültigkeit nicht durch das statistische Experiment sichergestellt ist. Die Ermittlung der Validität (Gültigkeit) schriftpsychologischer Methoden kann also nur unter Anwendung statistischer Verfahren auf Gutachter- oder §Schriftvariabelebene§ erfolgen.

1.1.3.1 *Validitätsuntersuchungen auf Gutachterebene* sind relativ leicht durchzuführen: Schriftpsychologische Aussagen werden direkt mit bestimmten Kriterien (s. d.) verglichen (korreliert).

Die Nachteile solcher kurzgeschlossenen Untersuchungen überwiegen jedoch etwaige Vorteile:

Bei negativen Ergebnissen bleibt offen, ob sie auf mangelnden Fähigkeiten der Gutachters, Material- oder Kriterienmängeln oder auf der Unzulänglichkeit der Methode beruhen.

Positive Ergebnisse dagegen dürfen nicht verallgemeinert werden. Sie gelten nur für den beteiligten Gutachter. Dieser Gutachter müsste zudem im Kontrollexperiment bei gleichbleibender Fragestellung mindestens noch einmal gleiche oder ähnliche Ergebnisse erzielen, wenn man den seltenen, aber statistisch zu erwartenden glücklichen Zufall ausschliessen möchte.

Da die graphischen Variablen bei derartigen Untersuchungen ausgeklammert bleiben, weiss man auch nicht über das Zustandekommen der richtigen ó d.h. vom Zufall signifikant abweichenden ó Aussagen. Gerade an diesem Wissen wäre den Schriftpsychologen aber sehr gelegen. Erklärungen des Gutachters haben jedoch nur informativischen Wert. Sie müssten letztlich doch durch eine Untersuchung auf der Schriftvariabelebene ergänzt und bestätigt werden.

Validitätsuntersuchungen auf der Gutachterebene testen folglich nur den gegenwärtigen Stand der schriftpsychologischen Aussagemöglichkeiten eines bestimmten Gutachters.

### 1.1.3.2 *Validitätsuntersuchungen auf Schriftvariabelebene*

Der Begriff §Schriftvariabelebene§ ist bekannter unter der Bezeichnung §Merkmalsebene§. Die Bezeichnung §Merkmal§ wird oft unlogischer Weise sowohl für alle Komponenten der Handschrift als auch für eine spezielle Gruppe gebraucht. Der Begriff §Variabel§ umfasst alle Elemente der Handschrift einschliesslich der Merkmale. Da bei Validitätsuntersuchungen nicht nur Merkmale (im engeren Sinne) überprüft werden, ist der Begriff §Schriftvariabelebene§ (oder abgekürzt: Variabelebene) angemessener.

Als Kriterien für Validierungsversuche auf der Variabelebene kann man sowohl schriftpsychologische Begutachtungen als auch unabhängige Aussenkriterien verwenden. Im ersten Falle testet man trotz Einschaltung von graphischen Variablen wiederum nur den einzelnen Gutachter respektive tradierte schriftpsychologische Theorie. Im zweiten Falle dagegen erhält man von der Theorie unabhängige Daten darüber, welche diagnostische Valenz bei verschiedenen graphischen Variablen erwartet werden kann. Das aber ist für die wissenschaftliche Konsolidierung der Disziplin weitaus angemessener und auch nützlicher und sollte deshalb mit Vorrang angestrebt werden.

## 1.2 Wahl der statistischen Methode

Eine exakte Bedeutungsfindung für graphische Variablen und/oder deren Syndrome sowie Kontrollen der psychodiagnostischen Ergiebigkeit schriftpsychologischer Methoden können nur mit Hilfe wissenschaftlich exakter, d. h. hier statistischer Verfahren erfolgen. Wer die Verwendung solcher Methoden in der Schriftpsychologie ablehnt, verlässt den Bereich wissenschaftlichen Denkens. Diskutabel ist nur die Frage, welche statistische Methode zur Anwendung kommen soll. Verschiedene Fragestellungen erfordern nämlich die Verwendung verschiedener statistischer Methoden oder lassen die Wahl zwischen verschiedenen Methoden zu. Andererseits erfordert jedes statistische Verfahren die Erfüllung bestimmter Bedingungen. Werden diese nicht berücksichtigt, so besteht das Risiko, dass sich die erzielten Ergebnisse nachträglich als ganz oder teilweise unbrauchbar erweisen. Dem Nicht-Statistiker ist daher die Konsultation eines Fachmannes dringend anzuraten. Diese Konsultation sollte grundsätzlich im Planungsstadium stattfinden. Bereits gesammelte Daten nachträglich auf ein statistisches Verfahren hin zu trimmen ist ein schwieriges, meist hoffnungsloses Unterfangen.

Das statistische Verfahren sollte möglichst ein in der Psychologie gebräuchliches Verfahren sein, um direkte Vergleiche mit den Ergebnissen anderer Untersuchungen anstellen zu können. Aus dieser Verfahrenswahl ergeben sich fast automatisch auch die Daten der Ergebnisse ( $r$ ,  $R$ ,  $\phi$ ,  $\chi^2$  usw.), die ebenfalls allgemein gebräuchlich sein sollten. Ergebnisse, gewonnen an ausgepichteten Modellen, dargestellt in ungewöhnlichen Werten, hängen ó bildlich gesprochen ó im leeren Raum.

Über Einzelheiten der Methoden informieren die Lehrbücher der psychologischen Statistik und Spezialaufsätze.

### 1.2.1 Stabilitäts- und Reliabilitätsuntersuchungen

Für die Prüfung des Zusammenhanges zwischen zwei Beobachtungsserien (z. B. Schätzungen einer graphischen Variable durch zwei Gutachter) und der Kontrolle gegenseitiger Abhängigkeit von Variablen hat sich die Anwendung einfacher Korrelationsmethoden bewährt.

### 1.2.2 Validitätsuntersuchungen

Bei Gültigkeitsuntersuchungen werden im Allgemeinen drei Untersuchungsmodelle verwendet:

Validierung einzelner Variablen

Faktorenanalyse

Multiple Regressionsanalyse

Validierungsexperimente lassen sich jedoch auch mit anderen Verfahren durchführen. So hat z. B. *Schneevoigt* (1968) die Zuverlässigkeit der schriftpsychologischen Intelligenzdiagnose auf Gutachterebene mit der sogenannten Matching-Methode untersucht. Dieser Versuch ist für die Schriftpsychologie neuartig, lässt aber nur die Bearbeitung zahlenmässig kleiner Versuchsreihen zu.

1.2.2.1 Die Validierung einzelner Variablen ist relativ ungebräuchlich. Sie wird meist nur dann vorgenommen, wenn man etwa eine bestimmte Behauptung testen will. Im Verhältnis zum Arbeitsaufwand ist das Verfahren unergiebig, denn man erhält auf eine Frage nur eine bejahende oder verneinende Antwort. Es wäre daher völlig unökonomisch, jede einzelne Arbeitshypothese der Schriftpsychologie auf diese Weise zu prüfen.

Dass eine einzelne graphische Variable hoch mit einem Kriterium korreliert, ist zwar kaum bei Merkmalen zu erwarten, könnte aber bei hoch komplexen Variablen durchaus der Fall sein.

Gebräuchlicher ist schon das Korrelieren einer Reihe von graphischen Variablen mit psychologischen Aussagen oder anderen Kriterien (z. B. Wallner, 1966). Derartige Untersuchungen haben insofern einen Wert, als sie voraussetzungslose Ermittlungen in Form von Vorstudien darstellen.

Die in diesem Zusammenhang zu erstellenden Korrelationsmatrizen erhält man allerdings bei Faktoren- und Regressionsanalysen bereits als Nebenprodukte. Alle bisher genannten Untersuchungsverfahren lassen sich mit Hilfe einfacher Rechenmaschinen bewältigen. Für qualifizierte Analysen jedoch ist es wegen der anfallenden Rechenarbeit nur ratsam, einen Computer zur Hilfe zu nehmen.

1.2.2.2 Die Faktorenanalyse ist gegenwärtig die gebräuchlichste Methode für Validierungsversuche auf Variabelebene. Sie gibt viel differenziertere Aufschlüsse über die Zusammenhänge zwischen allen in sie eingehenden Variablen und nicht nur zwischen jeweils zwei von ihnen.

Man kann die Faktorenanalyse auf die Untersuchung der Zusammenhänge zwischen rein graphischen Variablen beschränken. Das Ergebnis gibt Aufschluss über die Varia-

belstruktur der Handschrift (Seifert, 1964; Wallner, 1968). Über die Validität der eventuell gefundenen Faktoren sind keine Aussagen möglich. (Durch Paralleluntersuchungen könnte man aber z. B. die Strukturen der Handschrift bei verschiedenen Schriftvorlagen, Kulturgruppen, Geschlechtern, Berufsgruppen usw. untersuchen ó was bisher noch nicht geschehen ist. Es könnte sich dabei ergeben, dass diese verschiedenen Gruppen von Schreibern ganz verschiedene Faktorenstrukturen ó also Syndromgruppen ó aufweisen.)

Bezieht man dagegen Kriterien in die Analyse ein, dann erhält man Faktoren, die neben graphischen Variablen auch Kriterien enthalten können. Die Faktorenladung eines solchen Kriteriums gibt die Korrelation mit dem Faktor an (factor validity).

Die bisher vorgelegten Faktorenanalysen (z. B. *Fahrenberg*, 1961; *Adolfs*, 1963; *Fischer*, 1964; *Lockowandt*, 1966) haben allerdings noch keine für die Praxis entscheidenden Ergebnisse erbracht. (Der Grund dafür ist m. E. vor allem in einer verfehlten Auswahl der Kriterien [s. 1.4.1.4] und in einer einseitigen Bevorzugung messbarer graphischer Variablen [s. 1.3.2 und 1.3.4] zu suchen).

Der Wert der Faktorenanalyse liegt jedenfalls in der Möglichkeit, schriftpsychologische Hypothesen vielseitig zu prüfen und auf Grund der Ergebnisse neue oder korrigierte Hypothesen aufzustellen. Entscheiden für das Gelingen eines Projektes ist jedoch, wie weit der Forscher die Möglichkeiten der Faktorenanalyse auszuschöpfen vermag. Das setzt vor allem grösste Umsicht und Sachkenntnis bei der Auswahl der zu verarbeitenden Variablen und erst in zweiter Linie statistische Kenntnisse voraus.

1.2.2.3 Die multiple Regressionsanalyse (nach *Doolittle*) ist in der schriftpsychologischen Forschung bisher kaum angewendet worden (*Timm*, 1965). Das mag daran liegen, dass die Errechnung multipler Korrelationen (R) mit einfachem Rechengertät sehr umständlich und daher zeitraubend ist. I, Zeitalter der Grossrechner entfällt dieses Problem.

Im Gegensatz zur Faktorenanalyse, die nur die Korrelation einer Variable mit ihrem Faktor angibt, erhält man durch die multiple Regressionsanalyse die Korrelationen zwischen abhängigen und unabhängigen Variablen. Konkret bedeutet dies für schriftpsychologische Untersuchungen, dass man errechnen kann, wie hoch die multiple Korrelation z. B. zwischen einem Kriterium und graphischen Variablen oder umgekehrt zwischen einem Handschriftelement und einer Reihe verschiedener Kriterien ist. Dieses Verfahren ermöglicht somit eine von Spekulationen unabhängige Bedeutungsfindung.

Durch die entsprechende Programmierung der Berechnungen kann man sogar feststellen, welche Variablen im Einzelnen mit welchem Gewicht zur Diagnose eines Kriteriums beitragen können und auf welchem Zuverlässigkeitsniveau die jeweilige Variablenkombination liegt. Eine auf Grund derartig ermittelter Variabelsyndrome gestellte Diagnose kann im Kontrollexperiment durch Korrelation mit dem Kriterium überprüft werden.

Selbstverständlich hängt auch hier der wissenschaftliche Wert der Ermittlungen sowohl von der Qualität der verarbeiteten Daten als auch von der Qualifikation des Untersuchungsleiters ab.

Die multiple Regressionsanalyse lässt u. a. auch einen direkten Vergleich der Ergiebigkeit verschiedener Methoden zu. Will man etwa feststellen, ob und in welchem Ausmass schriftpsychologische Methoden zur Verbesserung einer testpsychologischen Untersuchung beitragen können, so errechnet man zuerst die Ergiebigkeit der Tests (Korrelation: Test x Kriterium). In einer zweiten Analyse berechnet man dann die multiplen Korrelationen des Kriteriums mit Tests und graphischen Variablen. Die Differenz zwischen dem R der ersten und der zweiten Analyse lässt dann Rückschlüsse darüber zu, wie weit die graphischen Variablen zur Verbesserung der Diagnose beitragen können. Auf diese Weise lässt sich also ein direkte und noch dazu praxisnaher §Brauchbarkeitsnachweisō für die schriftpsychologischen Methoden erbringen. Ein derartiger Versuch wurde kürzlich von *Wallner* (1969, deutsche Fassung in Druck) mit guten Ergebnissen abgeschlossen.

### 1.3 Auswahl und Einsatz graphischer Variablen

#### 1.3.1 Anzahl der graphischen Variablen

Falls die Fragestellung die Anzahl der zu untersuchenden graphischen Variablen nicht beschränkt, ist es immer ratsam, so viele Variablen wie nur irgend möglich zu erfassen. Bei Validierungsversuchen ist dies sogar eine Notwendigkeit. Eine obere Grenze wird heute nur durch die Aufnahmefähigkeit der Computer gesetzt. Diese ist unterschiedlich bei verschiedenen Typen und Programmen und sollte daher vorher festgestellt werden. Die Möglichkeit der gleichzeitigen Erfassung sämtlicher bereits definierter Variablen muss allerdings wegen ihrer unüberschaubaren Anzahl zumindest vorläufig noch als Utopie betrachtet werden.

#### 1.3.2 Auswahl der graphischen Variablen

Bei Stabilitäts- und Reliabilitätsuntersuchungen steuert die Fragestellung weitgehend die Auswahl der Variablen. Für die Ergebnisse von Validierungsexperimenten ist es von nicht abschätzbarer Bedeutung, ob verschiedene Definitionen ein und derselben graphischen Variable zu gleichen oder abweichenden Ergebnissen führen. Derartige Variablen sollten vor der endgültigen Validierung auf Übereinstimmung geprüft werden. Sind die Differenzen so gering, dass sie als zufällig betrachtet werden können, dann sollte die jeweils weniger reliable, bei fehlenden Unterschieden die weniger präzise definierte Variable einschliesslich der nach ihr ermittelten Werte fallen gelassen werden. Andernfalls entstehen Schwierigkeiten bei der Deutung der Resultate.

Bei Validierungsuntersuchungen sollte die Bevorzugung bestimmter Arten von graphischen Variablen unbedingt unterbleiben. Hier sei vor allem auf die Favorisierung von messbaren Variablen hingewiesen. Neben der Untersuchung von metrisch messbaren Merkmalen sollten auch Eindruckscharaktere, Eindruckscharaktergruppen und feste Zeichen berücksichtigt werden, auch wenn dies die statistische Bearbeitung erschweren sollte. Die am leichtesten zu definierenden und zu messenden Merkmale könnten sich nämlich als die diagnostisch am wenigsten ergiebigen Variablen erweisen. Bisherige Ergebnisse lassen diese Annahme mehr als gerechtfertigt erscheinen.

#### 1.3.3 Definition der graphischen Variablen

Grösste Sorgfalt ist auf die Definierung der Variablen zu verwenden. Unklare, verschwommene oder zu allgemein gehaltene Definitionen können zu erheblichen Abweichungen zwischen mehreren Beurteilern bei der Ermittlung der Ausprägungsgrade von Handschriftenvariablen führen. Je komplexer eine Variable zusammengesetzt ist (etwa Rhythmen und Versteifungsbefunde), um so minutiöser muss auch die Definition sein. Definitionen sollten stets in fixierter Form vor Arbeitsaufnahme vorliegen. Um negative Korrelationen zu vermeiden, sollte bei Skalierungen stets auch die Skalierungsrichtung angegeben werden.

Meist werden Definitionen verbal geboten. Bei komplexeren Variablen können aber z. B. Schwerpunktsverschiebungen in der Auslegung des Textes vorkommen, wenn nicht jeder einzelne Bestandteil der Definition mit Gewichten versehen wird. Dies ist jedoch niemals konsequent durchgeführt worden und auch kaum mit Erfolg durchführbar. Stattdessen sollten für alle Skalenstufen einer Variable über die verbale Definition hinaus Beispiele geboten werden. In vielen Fällen wird es sich sogar als notwendig erweisen, mehrere Beispiele pro Skalenstufe vorzulegen. Es ist selbstverständlich, dass diese Beispiele in Voruntersuchungen unter Mitarbeit hoch qualifizierter Beurteiler statistisch-experimentell geeicht werden müssen. (Derartige Eichungsuntersuchungen wären eines oder gar mehrerer spezieller Forschungsunternehmen würdig). Viele Missverständnisse oder grobe Fehlerquellen liessen sich auf diese Weise aus der Welt schaffen, und die von vielen Fachleuten ersehnte Erstellung eines exakten Schriftatlas würde der Verwirklichung um einen grossen Schritt näher kommen.

#### 1.3.4 Ermittlungsmodus

Graphische Variablen können mitunter sowohl gemessen als auch geschätzt (gestuft) werden.

In den Untersuchungen der letzten Jahre zeigt sich ein extremer Hang zur Messung und zur Errechnung von allerlei Relationen zwischen diesen Messungen. Eine reiche Flora neuer Variablen und Subvariablen ist das Ergebnis. Gegen dieses Vorgehen ist einzuwenden,

- dass für die meisten messbaren Variablen sehr hohe, nur zufällig abweichende Übereinstimmungen mit Schätzungen ermittelt wurden (Wallner, 1961), was zeitraubende Messungen an diesen Variablen überflüssig macht, und
- dass da, was da an Quotienten und Relationen mit grösster Akribie errechnet wird, keinesfalls auch diagnostische Valenz haben muss.

Tatsächlich sind die Ergebnisse dieser mit grossem Arbeitsaufwand betriebenen Untersuchungen äusserst mager. Die Erklärung dürfte im folgendem liegen:

Was an der Handschrift mit heutigen Methoden messbar ist, das ist vorwiegend der manipulierbare formale Anteil der Schrift. Was dagegen diagnostisch relevant sein dürfte, liegt vermutlich hauptsächlich in der Bewegungsführung und den aus ihr resultierenden Komponenten. Eine Messung der ursprünglichen, die Handschrift prägenden Bewegung ist sicher möglich. Aber die bislang erprobten Verfahren sind entweder §laboratoriumsgebundenō oder samt und sonders zu grob, um den praktischen Anforderungen in diesem Bereich der graphischen Bestandesaufnahme gerecht zu werden.

Diese Feststellungen dürfen nicht als Herabwertung der umfassenden Messexperimente aufgefasst werden. Die Experimente waren vielmehr von grösster Bedeutung für die weitere Forschung, da sie den relativ geringen Nutzen von exakten Messungen über weite Bereiche hin ermittelt haben. Weitere Versuche in dieser Richtung erscheinen aber nach den Gegebenheiten so lange überflüssig, wie keine adäquateren Messmethoden zur Verfügung stehen. Die Forschung sollte sich stattdessen mehr der Weiterentwicklung von Registrierungsmethoden für Eindruckscharaktere und Eindruckscharaktergruppen (Ganzheitseindrücke) zuwenden. Dieses Gebiet ist zwar viel schwerer zugänglich und voller technischer und statistischer Probleme, erscheint aber lohnenswert zu bearbeiten.

Ein Versuch in dieser Richtung ist die Vereinigung von Eindruckscharakteren zu Gruppen. Er geht von der Hypothese aus, dass gewisse Eindruckscharaktere und entsprechende feste Zeichen nur verschiedene Seiten ein und desselben Sachverhaltes sind. Derartige Zusammenfassungen mögen gegenüber den feinsinnigen Deutungshypothesen der Graphologie grob erscheinen. Die bisher gewonnenen Forschungsergebnisse sprechen aber offensichtlich für ein derartiges Vorgehen (Wallner, 1968).

### 1.3.5 Registrierungsverfahren

Werden an graphischen Variablen Messungen ausgeführt, so sollten bei der Datenverarbeitung stets die Rohwerte benutzt werden.

Bei der Schätzung (Skalierung) von Ausprägtheitsgraden treten mehrere methodologische Probleme auf.

Die Wahl der Skalenstufen muss der jeweiligen Variable angemessen sein. Bei weit streuenden Variablen ist  $\sigma$  auch aus statistischen Gründen  $\sigma$  eine neungradige Skala (stanine) zu empfehlen. Lässt die Variable eine so differenzierte Skalierung nicht zu, ist die grösste noch brauchbare Skala experimentell zu ermitteln. In einigen Fällen muss man sich erfahrungsgemäss mit dreistufigen Skalen oder gar mit einer einfachen Dichotomie (vorhanden/nicht vorhanden) begnügen.

Eine weitere Frage ist, ob man eine Normalverteilung der Werte anstreben soll oder nicht. Die beobachtbaren Werte sind nämlich nicht immer mit den nach der Normalverteilung zu erwartenden Werten kongruent. (Im Zweifelsfalle gibt ein einfacher  $\chi^2$ -Test Auskunft). Klassifiziert man bei signifikant von der Normalverteilung abweichenden Werten nach den beobachteten Werten, ergeben sich Schwierigkeiten bei der weiteren statistischen Verarbeitung. Forciert man dagegen die Werte in eine Normalverteilung, treten Verzerrungen in der Verteilung der Werte auf.

Einen Ausweg aus diesem Dilemma bietet eine andersartige Auswertung der Daten: Die einzelnen, genau definierten Skalenstufen werden als unabhängige Variablen behandelt. (Eine Faktorenanalyse auf dieser Basis wurde kürzlich von Stein Lewinson [1968] durchgeführt).

Ein derartiges Verfahren kann sich auf Validierungsexperimente positiv auswirken. Bisher hat man nämlich Kriteriendaten mit den Werten graphischer Variablen über ihre gesamte Variationsbreite hinweg korreliert. Einer solchen Berechnung wird aber automatisch unterstellt, dass die miteinander korrelierten Beobachtungsserien über die gesamte Variationsbreite parallele Bedeutungen haben (z. B. kleine Schrift bedeutet das

Gegenteil einer grossen, Zwischenwerte hier entsprechen Zwischenwerten dort). Derartige Hypothesen sind aber in der Schriftpsychologie sehr selten. Vielmehr wurde seit eh und je angenommen, dass nur bestimmte Ausprägtheitsgrade einer graphischen Variable mit bestimmten Ausprägtheitsgraden eines Kriteriums korrespondieren.

Abb. 1 mag dies verdeutlichen:

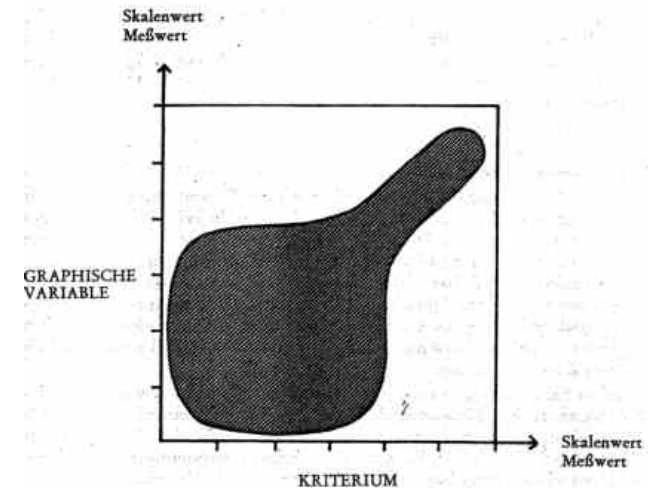


Abb. 1. Verteilung einer graphischen Variable und eines Kriteriums

Obwohl in Abb. 1 eindeutige Zusammenhänge zwischen den höchsten Ausprägtheitsgraden der graphischen Variable und des Kriteriums bestehen, dürfte die Korrelation zwischen den beiden Serien insgesamt nicht-signifikante Werte ergeben. Dagegen würden Validitätsberechnungen, die sich auf einzelne Skalenstufen beschränken, äusserst hohe Werte erbringen.

Nähere Untersuchungen der Validität von Handschriftenvariablen nach diesem Verfahren erscheinen demnach äusserst zweckmässig und erfolgversprechend. (Bisher sind meines Wissens noch niemals Berichte über Versuche in dieser Richtung veröffentlicht worden). Man muss dann allerdings mit nicht-parametrischen Berechnungsverfahren arbeiten, die gegenüber parametrischen Methoden einige Nachteile aufzuweisen haben.

Bei der Registrierung von graphischen Variablen wie von sämtlichen Daten über einen Probanden überhaupt, ist auch die arbeitstechnische Seite zu beachten. Ist die Bearbeitung des Materials durch einen Computer vorgesehen, so empfiehlt es sich, ein Arbeitsblatt pro Proband oder (bei nicht allzu vielen Daten) pro Untersuchungsgruppe auszuarbeiten, auf dem die Daten unmittelbar und endgültig eingetragen werden können. Die Anordnung der Daten soll so erfolgen, dass das Blankett sich direkt als Unterlage für die Stanzung von Lochkarten eignet. Ein mehrmaliges manuelles Überführen von Daten führt nämlich erfahrungsgemäss zu Übertragungsfehlern, deren Konsequenzen nicht abzusehen sind.

Bei der Ausarbeitung von Arbeitsblättern sollte von vornherein auf bereits vorhandene Programme für die Datenbearbeitung durch Computer Rücksicht genommen werden. Meist wird eine bestimmte Reihenfolge der Daten gewünscht:

- Daten zur Person
- Prediktoren (Test oder graphische Variablen)
- Kriterien

### 1.3.6 Zusammenfassung der Daten

Wirken mehrere Gutachter oder Beurteiler in einer Untersuchung mit, so stellt sich automatisch die Frage, ob deren Ergebnisse einzeln behandelt oder zusammengefasst werden sollen. Die Ergebnisse jedes einzelnen Gutachters können nämlich manchmal zu niedriger Reliabilität haben. In derartigen Fällen sollte man unbedingt mit Zusammenfassungen arbeiten, da die Reliabilität  $\hat{o}$  vor allem bei zwei bis drei Gutachtern - merklich ansteigen pflegt vorausgesetzt, dass überhaupt ein gemeinsamer Trend in den Beurteilungen gegeben ist. Selbst wenn die Aussagen aller Gutachter hoch reliabel sind, empfiehlt sich eine Zusammenfassung.

In jedem Falle sollte man ohne Rücksicht auf die Reliabilität sowohl mit Einzeldaten als auch mit Zusammenfassungen arbeiten. Es kann sich nämlich später erweisen, dass nur einer von mehreren Gutachtern im Alleingang gute Werte erzielt. Für die für die weitere Bearbeitung notwendigen Entscheidungen bietet diese Mehrdeutigkeit also oft sehr wertvolle Unterlagen.

Die Summierung und Aufbereitung von Daten mehrerer Gutachter kann mit Vorteil dem Computer überlassen werden.

## 1.4 Auswahl der Kriterien und Einsammlung von Kriteriendaten

Die Wahl der Kriterien sollte stets im Hinblick auf die Fragestellung erfolgen. In der Praxis ist es aber oft so bestellt, dass der Forscher mit dem Vorlieb nehmen muss, was ihm geboten oder gar aufgezwungen wird oder was allenfalls ohne grösseren Arbeitsaufwand beschafft werden kann. Oft genug sind also die Arbeitsbedingungen so geartet, dass der Untersuchungsleiter viele Abstriche von den von ihm als ideal oder zumindest wünschenswert angesehenen Bedingungen machen muss. Auf jeden Fall sollte aber die Erlangung so objektiver und valider Kriterien wie möglich angestrebt werden.

Die gebräuchlichsten Arten von Kriterien sollen im Folgenden kurz skizziert und kommentiert werden:

### 1.4.1 Verschiedene Arten von Kriterien

#### 1.4.1.1 Persönliche Daten

Je nach Fragestellung können Alter, Geschlecht, Art und Dauer der Schulausbildung, erlernter Beruf, ausgeübter Beruf, Dauer der Berufserfahrung usw. verwendet werden. Diese Daten sind zwar relativ objektiv, dürften aber für Validitätsuntersuchungen nur von begrenztem Wert sein.

#### 1.4.1.2 Klinische Daten

Medizinische und psychiatrische Daten können ebenfalls als Kriterien verwendet werden. Am objektivsten sind Laboratorienbefunde. Klinische und vor allem psychiatrische Diagnosen können nur  $\hat{o}$  wie alle anderen subjektiven Beurteilungen (s. 1.4.1.7)  $\hat{o}$  mit entsprechenden Vorbehalten benutzt werden.

#### 1.4.1.3 Intelligenztests

Bei Verwendung von Intelligenztests als Kriterien sollten zu Berechnungen stets die Rohwerte und nicht die Stanine-, T- oder andere Werte (IQ) verwendet werden.

Es ist ausserdem zu empfehlen, Reliabilität und Validität der Kriterien an der selben Versuchsgruppe zu überprüfen, die auch Gegenstand der schriftpsychologischen Erhebung ist. (Derartige Daten gelten nämlich nicht generell, sondern nur unter bestimmten Bedingungen für die untersuchte Population und auch dann oft nur für eine bestimmte Zeitspanne):

#### 1.4.1.4 Projektive Tests

Projektive Tests werden mit dem Ziel konstruiert, bestimmte Aufschlüsse über die Persönlichkeit der Testanden zu erbringen. Es gibt jedoch bis heute kein einziges auch nur annähernd allgemein anerkanntes Verfahren dieser Art. Nicht einmal der Rorschach-Test mit seiner weltweiten Verbreitung bildet hier eine Ausnahme.

Dass man dennoch weitaus umstrittenere Methoden als Kriterien für schriftpsychologische Validitätsuntersuchungen wählt, beruht auf der Annahme, dass gleichartige Resultate, gewonnen an völlig verschiedenen Untersuchungsverfahren, die hinter den Verfahren liegenden Theorien stützen. Konkret ausgedrückt: Korrelieren die Aussagen des projektiven Tests A über die Persönlichkeitsvariable X signifikant mit den schriftpsychologischen Begutachtungen der Persönlichkeitsvariable X, so messen beide Verfahren innerhalb gewisser Grenzen tatsächlich die Persönlichkeitsvariable X. (Diese Validitätsform wird als construct validity bezeichnet. Über andere Formen geben die einschlägigen Lehrbücher Auskunft).

Fassen wir die Kritik an dieser Arbeitsweise und die daraus resultierende Empfehlung in ein Gleichnis: Es ist zwar menschlich gesehen höchst lobenswert, wenn ein Gebrechlicher den andern stützt. Viel besser wäre es aber zweifellos für beide, wenn jeder eine eigene, stabile und verlässliche Stütze fände.

#### 1.4.1.5 Leistungen

Produktionszahlen, Schulzeugnisse in bestimmten Fächern, Verkaufserfolge, ergometrische Ergebnisse usw. können in direkten Messungen erfasst oder auf Grund von erzielten Resultaten beurteilt werden. Selbstverständlich stellen Leistungsmessungen und Leistungsbeurteilungen keine absoluten Werte dar, denn auch sie werden von anderen Faktoren (z. B. Motivation) manchmal sogar sehr stark beeinflusst. Immerhin kommen Leistungen dem Idealkriterium zumindest näher als die anderen Kriterien. Wo immer sich Leistungen als Kriterien anbieten, sollten sie auch verwendet werden. Ein Nachteil ist in vielen Fällen nur, dass Leistungen oft komplexe Variablen sind und selten mit reinen Persönlichkeitsvariablen zusammenfallen. Letztere können immerhin erschlossen und in anschliessenden Spezialstudien näher untersucht werden.

#### 1.4.1.6 Fragebögen

Haben Fragebögen nur die Aufgabe, die Einsammlung von objektiven Daten systematischer zu gestalten, so ist gegen ihre Verwendung kaum etwas einzuwenden. Sollten aber die Antworten Aufschluss über die Persönlichkeit des Beantworters geben, so können alle Argumente gegen die Verwendung angeführt werden, die sich in der Spezialliteratur finden.

In bestimmten Situationen kann man allerdings mit relativ unverfälschten Antworten rechnen, selbst wenn die Fragen ganz offen formuliert sind. Rat- und Hilfesuchende sind beispielsweise mehr als andere bereit, Schwächen (oder was sie als solche ansehen) preiszugeben. In Bewerbungssituationen würden die gleichen Personen auf die gleichen Fragen mit grösster Wahrscheinlichkeit völlig anders reagieren.

#### 1.4.1.7 Subjektive Beurteilungen

In vielen Untersuchungen werden Persönlichkeitsbeurteilungen von Personen verwendet, die den Probanden kennen oder bei denen man eine solche Kenntnis voraussetzt. Gebräuchlich sind die Einsammlungen von Lehrer-, Vorgesetzten-, Kameraden- und Verwandtenurteilen. Auch die Beurteilungen durch Psychologen, Ärzte und Sozialarbeiter sowie Selbstbeurteilungen gehören in diese Kategorie.

Verwendet man derartige Beurteilungen für Forschungszwecke, so muss man die Reliabilität dieser Beurteilungen am vorliegenden Material ermitteln. Man benötigt also mindestens zwei voneinander unabhängige Beurteiler. Durch den Nachweis der Reliabilität der Aussagen werden die subjektiven Beurteilungen objektiviert. Die Bezeichnung *sub*subjektiv ist also ein terminus technicus und kein Werturteil.

Da die Reliabilität aus mannigfachen Gründen, auf die hier nicht näher eingegangen werden kann, oft recht niedrig ausfällt, versucht man sie durch eine statistische Manipulation aufzubessern: Man verwendet statt zwei bis drei Gutachtern deren 10 bis 12. Wenn die Beurteilungen dieser Gutachter nur einigermaßen in gleicher Richtung tendieren, erhält man durch Summierung der Aussagen am Ende statistisch gesehen hohe Reliabilitätswerte. Bei der Zusammenfassung von Laienbeurteilungen (z. B. Kameradenbeurteilungen) mag dieses Verfahren immerhin noch vertretbar sein, da man Schwächen in der Exaktheit der Einzelbeurteilungen auf mangelnde Schulung, Erfahrung und Befähigung zurückführen kann. Eine derartige Hochrechnung von Begutachtungen von Fachleuten muss dagegen als dubiose betrachtet werden.

Die statistische Bearbeitung von subjektiven Beurteilungen setzt den Einsatz von regulierenden Methoden voraus, da freie Persönlichkeitsbeschreibungen nicht exakt verglichen werden können und nachträgliche Bearbeitungen allzu leicht zu Verfälschungen der Meinungen der Beurteiler führen.

Um subjektive Beurteilungen auf einen Nenner zu bringen, kann man paarweise Vergleiche durchführen, Rangordnungen legen oder Skalierungen vornehmen lassen. Über Vor- und Nachteile sowie kritische Momente gibt die Fachliteratur Bescheid.

#### 1.4.2 Einsammlung von Kriteriendaten

Es wurde bereits angedeutet, dass die Brauchbarkeit eines Kriteriums durch verschiedene Faktoren beeinflusst werden kann. Besonders zu beachten ist die Motivation der

Mitwirkenden. Die angemessensten Resultate dürfte man stets dann erzielen, wenn das Experiment in eine Untersuchung eingepasst ist, die vollen Realitätscharakter hat. Personen, die um Anstellung, Beförderung oder Ausbildung nachsuchen, werden z. B. Testaufgaben auf andere Weise angehen als Wehrpflichtige, die für ein Experiment vom kaum populären Aussendienst abkommandiert werden. Dass Rat- und Hilfesuchende auf bestimmte Typen von Fragebögen anders als Stellenbewerber reagieren, wurde bereits erwähnt. Vorgesetzte, die ihre betriebsinternen Personalbeurteilungen vor ihren Chefs und gleichzeitig gar den Beurteilten gegenüber vertreten müssen, werden ihre Aufgabe viel ernster nehmen, als wenn der aus der Aufgabenstellung resultierende Druck nicht gegeben wäre.

#### 1.5 Auswahl der Mitarbeiter

Bei schriftpsychologischen Untersuchungen ist der psychologische Gutachter oder Beurteiler der Handschriftenvariablen das eigentliche Messinstrument, das es zu überprüfen gilt. Selbst wenn reine Messungen an Handschriften vorgenommen werden, können bei den Messenden für sie typische Fehler vorkommen. Gutachter und Beurteiler sind also nicht beliebig austauschbar. Schlüsse von den Ergebnissen des einen Gutachters auf die Leistungsfähigkeit eines anderen oder gar auf die gesamte Wissenschaft von der Psychologie der Handschrift sind somit unstatthaft, da sie der sachlichen Grundlage entbehren.

Als Mitarbeiter bei Experimenten können u.a. Schriftpsychologen und in schriftpsychologischer Ausbildung befindliche Personen, Anfänger wie Experten, Psychologiestudenten oder sogar Laien eingesetzt werden. Es ist aber in jedem Untersuchungsbericht unbedingt und genauestens anzugeben, welcher Kategorie die einzelnen Beurteiler entstammen, welche Ausbildung sie genossen, welche Berufserfahrung sie haben oder welche anderen Besonderheiten sie auszeichnen. Erst dann können die Ergebnisse richtig gewürdigt werden.

Validierungsexperimente auf Gutachterebene sollten nur mit Beurteilern mit langer Berufserfahrung durchgeführt werden. Arbeitet man mit Laien oder Berufsanfängern, dann sollte wenigstens ein Experte eingesetzt werden. Bei negativen Ergebnissen, gewonnen an unerfahrenen Gutachtern, besteht nämlich trotz ausdrücklicher Hinweise auf die begrenzte Qualifikation der Mitarbeiter die Gefahr der Verallgemeinerung durch voreingenommene Kritiker.

Es muss sichergestellt sein, dass die Mitarbeiter keine anderen Kenntnisse über die Versuchspersonen haben oder erhalten, als nach den Versuchsbedingungen und der Zielsetzung vorgesehen und zulässig sind.

#### 1.6 Auswahl der Versuchspersonen

##### 1.6.1 Stichproben und deren Grösse

Nach Angaben der Bevölkerungsstatistiker reicht für die Bundesrepublik Deutschland eine Stichprobe von zirka 2000 Personen aus, um die Gesamtbevölkerung zuverlässig zu repräsentieren. Auf welche verschiedenen Weisen solche Stichproben zusammenge-

stellt werden, kann im aktuellen Falle aus der einschlägigen Literatur ermittelt werden. Für weniger präventive Untersuchungen reichen etwa 100 bis 150 Fälle aus. Eine Erweiterung des Materials über diese Grenze hinaus erbringt nämlich bei wesentlich grösserer Arbeitsbelastung kaum nennenswerten Gewinn. Sinkt dagegen die Grösse der Stichprobe unter 50 Fälle, so erhöht sich das Risiko rapide, dass Zufälligkeiten das Resultat beeinflussen oder gar verfälschen. Im Übrigen lohnt es sich immer mit einer grösseren Stichprobe als vorgesehen zu beginnen, da während der Untersuchung meist ein Materialschwund einzutreten pflegt.

### 1.6.2 Zusammensetzung der Stichprobe

1.6.2.1 Die Zusammensetzung der Stichprobe kann man zwar dem Zufall überlassen, ratsam ist dies jedoch nicht. Meist bestimmt die Fragestellung schon die notwendige Beschaffenheit der Stichprobe.

Die Zusammensetzung kann nach der erlernten Schulvorlage, nach Alter, Geschlecht, Ausbildungsgang und Ausbildungsstand, Beruf, Gesundheitszustand, Konstitution usw. erfolgen. Eine oder auch mehrere Variablen können konstant gehalten werden (z. B. Frauen im Alter von 30 bis 40 Jahren; männliche italienische Gastarbeiter im Hotelgewerbe).

Es sind hier regelhafte Abhängigkeiten zu beachten. Je genauer die Stichprobe definiert ist umso zulässiger sind verallgemeinernde Schlussätze, umso genauere Diagnosen lassen sich bei positiven Resultaten in Zukunft stellen.

1.6.2.2 Eine besondere Art der Stichprobe liegt bei Aufteilung in *Extremgruppen* vor (sehr gut/sehr schlechte Schüler, extreme Linksräg-/extreme Rechtsschrägschreiber). Solche Aufteilungen haben dann einen Wert, wenn man vorfühlen will, ob sich die nähere Untersuchung einer Hypothese lohnt. Dieses Verfahren ersetzt aber keinesfalls die Untersuchung der Variable an einem Kontinuum. In der Praxis hat man gerade mit Extremfällen selten zu tun. An Extremgruppen gewonnene positive Ergebnisse müssen also an einer Normalgruppe nochmals nachgewiesen werden, ehe man ihnen umfassendere Gültigkeit zusprechen darf.

## 1.7 Auswahl von Versuchsgerät und Material

Schriftpsychologische Untersuchungen können durchaus mit §Alltagsmaterial§ durchgeführt werden. Man muss dabei aber in Kauf nehmen, dass das Schriftenmaterial in vieler Hinsicht ausserordentlich heterogen werden kann. Dies kann wiederum die Arbeit und die Ergebnisse zumindest beeinträchtigen. Daher ist es nur zu empfehlen, die äusseren Umstände der Schriftprobenproduktion so weit wie möglich unter Kontrolle zu halten. Papierqualität und §format, Unterlage und Schreibgerät lassen sich dann weitgehend vereinheitlichen. Als Schreibgeräte sollten Bleistifte oder Kugelschreiber verwendet werden. Tintenschriften werden nämlich kaum noch produziert. (Die Forderung nach Anfertigung einer Tintenschrift dürfte heutzutage schon auf Verständnislosigkeit oder gar Widerstand stossen). Welche Geräte und Materialien auch immer verwendet werden: Es sollte sich stets um standardisierte Mittel handeln, die in gleicher Qualität und Ausfertigung jederzeit neu beschafft werden können und ausserdem von den Schreibern als §normal§ betrachtet werden. Für Kontrolluntersuchungen sind genaue

Angaben (z. B. §Kugelschreiber, Schneider, documenta 50, blau§) nicht nur erwünscht, sondern sogar notwendig.

Ein besonderes Problem stellt die Liniierung des Papiers dar. Die Erfahrung zeigt, dass sich viele Versuchspersonen gegen das Schreiben auf unliniertem Papier sträuben. Liniierungen erzeugen nachweisbar materialbedingte Veränderungen einer Reihe von graphischen Variablen. Die Erstellung eines speziellen Liniensystems für schriftpsychologische Experimente könnte das Problem wenigstens teilweise lösen. Ein Beispiel für ein solches System zeigt Abbildung 2:



Abb. 2. Ausschnitt aus einem Schreibbogen, Ideenskizze

Der Papierbogen wird mit parallelen hellgrauen Balken von 0,5 bis 1 mm Breite im gleichen Abstand § etwa wie sogenannte Aerogramme § bedruckt. Der Grauton der Balken sollte so hell wie nur irgend möglich gehalten werden. Der Vorteil für den Schriftpsychologen läge u. a. in einer Erleichterung der Schätzung einer Anzahl von graphischen Variablen. Die Schreiber hätten immerhin eine gewisse § wenn auch von den Schriftpsychologen unerwünschte § Stütze.

Selbstverständlich müssen der günstigste Balkenabstand und die geeignetste Grauskalendifferenz vor Beginn des Hauptversuchs experimentell ermittelt werden.

Bei Bewährung dieses oder eines gleichartigen Systems sollten sich folgende Untersuchungen des gleichen Mittels bedienen. Die benutzte Papierqualität und alle anderen drucktechnischen Details müssen deshalb im Forschungsrapport über die Entwicklung eines derartig geeichten Liniensystems bis in alle Einzelheiten angegeben werden.

## 1.8 Einsammlung des Schriftmaterials

### 1.8.1 Kenntnis des Verwendungszweckes

Da die Kenntnis des Verwendungszweckes bei vielen Probanden oft zur nachweisbaren, unerwünschten Beeinflussung ihrer Handschrift führt, wird die Vereinnahmung von Schriftproben zu Versuchszwecken meist getarnt. Da dieses Verfahren mit ethischen Argumenten angegriffen werden kann, sollten die Versuchspersonen bei Experimenten zumindest spätestens nach Abschluss des letzten Arbeitsmomentes aufgeklärt und das Einverständnis zur weiteren Bearbeitung eingeholt werden. Volle Anonymität und Schutz vor Einsichtnahme in das Material durch Unbefugte muss zugesichert werden und auch gewährleistet sein.

### 1.8.2 Formen der Materialeinsammlung

Die kontrollierte Materialeinsammlung kann durch Diktat, Textabschrift, durch schriftliche Beantwortung von Fragen erfolgen.



Diktate und Abschriften sollen garantieren, dass schriftpsychologische Urteile einzig auf Grund der Schriftvariablen gefällt werden. Diese Steuerung der Materialeinsammlung dürfte jedoch oft einen Verlust oder zumindest eine Einschränkung der Schreibspontaneität zur Folge haben. Zur oft behaupteten Beeinflussung des Gutachters, der zur Beurteilung von Schriftproben hinzugezogen wurde, die in einer ihm völlig unbekannt Sprache abgefasst waren, hoch signifikant mit den Bewertungen seiner sprachkundigen Kollegen korrelierte (Wallner, 1965, 1969[a]). Dabei hatte er nicht einmal die Möglichkeit in Zweifelsfällen exakt zu identifizieren, welche Buchstaben mit gewissen Bewegungsspuren gemeint waren. Das Beeinflussungsargument hat somit wenigstens einen Teil seines Gewichts eingebüsst. Weitere Vor- und Nachteile der Einsammlungsformen sind schon so oft diskutiert worden, dass hier nicht näher auf sie eingegangen zu werden braucht.

Die Diskussion um die Erstellungsformen der Schriftproben verliert in dem Masse an Bedeutung, indem sich Validierungsexperimente mehr und mehr von der Gutachterebene auf die Variablebene bei gleichzeitiger Korrelation mit Aussenkriterien verlagern.

Nur am Rande sei noch das Schreibenlassen von ausgewählten Experimentwörtern erwähnt, da es meist an ganz bestimmte Versuchsbedingungen gebunden ist. Für normale schriftpsychologische Untersuchungen dürfte es kaum geeignet sein.

### 1.8.3 Erwünschter Materialumfang

Diktate und Abschriften geben ein einheitliches Materialvolumen pro Versuchsperson und garantieren ausserdem das Vorkommen bestimmter (gewünschter) Buchstabenkombinationen. Um bei freieren Produktionsformen einen Mindestumfang an Material zu erhalten, wird meist ein Aufsatz usw. von mindestens einer A4-Seite<sup>o</sup>, doch ohne Angabe einer Mindestzahl von Zeilen erbeten. Im Normalfalle dürfte eine derartige Schriftprobe für den Untersuchungszweck ausreichen, obwohl ein viel grösserer Umfang selbstverständlich wünschenswert wäre. Bei vereinzelt Versuchspersonen muss man aber damit rechnen, dass sie absichtlich grossspurig und breitrandig schreiben, um das Pensum mit geringstem Arbeitsaufwand zu erledigen. In diesen Fällen wird das Ergebnis insofern verfälscht, als eine mit Wahrscheinliche bereits vorhandene Tendenz eine unerwünschte, übertriebene Verstärkung erfährt. Dass die Themenwahl Umfang und Ausführung beeinflussen kann, sei als Selbstverständlichkeit nur erwähnt.

### 1.9 Vorbereitung der Wiederholung

Jede Untersuchung sollte im Rahmen des Möglichen von vornherein so angelegt werden, dass sie wiederholt werden kann. Der Wert solcher Kontrolluntersuchungen liegt u. a. in der Möglichkeit, den Nachweis für Stabilität einmal erzielter Ergebnisse erbringen zu können. Verfügt ein Versuchsleiter über ein grosses Material oder kann er es sich ohne besondere Mühe beschaffen, so braucht er das Experiment für diesen Nachweis nicht einmal zu wiederholen. Er kann nämlich sein Material ohne weiteres in zwei Paralleluntersuchungen aufteilen. Dies geschieht am besten schon im Planungsstadium. Als Aufteilungskriterien können regionale, zeitliche und andere Faktoren dienen.

Bei Wiederholungsexperimenten sollte das Zeitintervall bei allen Arbeitsmomenten konstant gehalten werden. Auch die übrigen Umstände der Untersuchung sollten sorg-

fältig berücksichtigt und kongruent behandelt werden. Bemüht man die gleichen Gutachter, so ist dies in der späteren Berichterstattung ausdrücklich anzugeben.

Aber auch andere Interessenten sollten Gelegenheit haben, das Experiment zu überprüfen. Will der Experimentator vermeiden, dass seine Untersuchung auf mangelhafte Weise oder unter falschen Prämissen wiederholt wird, tut er gut daran, seine Arbeitsweise eindeutig und lückenlos darzustellen.

## 2. Durchführung

### 2.1 Allgemeine Gesichtspunkte

Die systematische Aufstellung von Regeln für die praktische Durchführung von schriftpsychologischen Experimenten ist wegen der vielseitigen Möglichkeiten und der Vielzahl der einwirkenden Faktoren nicht möglich. Prinzipiell gelten aber auf jeden Fall alle für Einzel- und Gruppentestungen gültigen Grundregeln, soweit sie nicht durch das Untersuchungsziel aufgehoben oder eingeschränkt werden. Hier sei daher nur auf einige wenige Punkte hingewiesen.

Ist z. B. eine Validierung auf Gutachterebene vorgesehen, dann darf der in Aussicht genommene Gutachter unter keinen Umständen etwa als Instruktor beim Abnehmen der Schriftprobe oder anderer Proben mitwirken. Der derart eingesetzte Gutachter würde in diesem Falle unzulässige Vorhandsinformationen über die Versuchspersonen erhalten. Wirkt er dennoch mit, so ist dies im Untersuchungsbericht anzugeben und zu begründen.

Besondere Beachtung verdienen die Milieufaktoren, unter denen Handschriftenproben, Testresultate usw. zustande kommen. Beleuchtung, Tageszeit und Reihenfolge einer längeren Untersuchung sowie der Schriftprobenabnahme vorangegangene Erfolgs- oder Misserfolgserlebnisse haben Einfluss auf das Untersuchungsergebnis.

Ob es beispielsweise ratsam ist, Untersuchungsergebnisse von Arbeitsgruppen zu mischen, die teils am Vormittag, teils in den späten Abendstunden gearbeitet haben, muss vom Versuchsleiter sorgfältigst erwogen werden. Experimente mit erprobten Testverfahren sprechen jedenfalls dagegen.

Zu den Milieufaktoren bei Gruppenuntersuchungen gehören auch von aussen kommende Störungen und das Verhalten der Versuchspersonen zueinander. Es kommt z. B. immer wieder vor dass sich bei Gruppen eine den Untersuchungsgang störende Haltung herausbildet. Diese braucht nicht einmal negativ gemeint zu sein und braucht auch nicht auf ein Fehlverhalten des Versuchsleiters zurückzugehen. So kann die Abnahme der Schriftprobe etwa schon durch einen Zwischenruf eines Spassvogels von vornherein entwertet werden. Um eine sachgerechte Auswertung zu gewährleisten, sollten daher alle besonderen Vorkommnisse während der gesamten Untersuchung protokolliert werden.

Auf falschen Voraussetzungen beruhende Ergebnisse sollten unbedingt kassiert werden, selbst wenn dies in extrem gearteten Fällen einen völligen Neubeginn erforderlich machen sollte. Eine nähere Begründung dürfte sich erübrigen.

## 2.2 Materialbearbeitung

Auf den Wert von vorgefertigten Formularen zur Erfassung der Untersuchungsergebnisse wurde bereits hingewiesen (1.3.5). Diese Vorschläge gelten selbstverständlich über die Erfassung der graphischen Variablen hinaus für alle ermittelten Daten.

Bei der Durchführung des Experimentes, spätestens aber bei der Durchsicht des gesammelten Materials, kann der Versuchsleiter Daten entdecken, die nach Fragestellung und Zielsetzung gar nicht in die Untersuchung hätten aufgenommen werden dürfen, da sie das Ergebnis verfälschen können (z. B. ein einzelner Ausländer mit mangelnden Sprachkenntnissen; Bleistiftschrift zwischen lauter Kugelschreiberstiften; Vpn, bei denen wesentliche Daten fehlen). In derartigen Situationen möchte der Versuchsleiter das Material durch Herausnahme des Gesamtfalles bereinigen. Obwohl einige Autoren darauf bestehen, dass einmal eingesammeltes Material auch verarbeitet werden muss, möchte ich dem Versuchsleiter das Recht zur angemessenen Korrektur einräumen. Das Herausnehmen von Material wegen zu erwartender Verschlechterung des Ergebnisses ist dagegen ein Fälschungsmanöver und daher verwerflich. Um einem solchen Vorwurf von vornherein vorzubeugen, empfiehlt es sich, eine eventuelle Korrektur von einer neutralen Person nach schriftlicher Anweisung vornehmen zu lassen.

## 3. Berichterstattung

Forschungs- und Entwicklungsarbeiten werden nicht um ihrer selbst willen betrieben. Es liegt vielmehr im allgemeinen Interesse, dass ihre Ergebnisse die bestmögliche und angemessenste Verarbeitung finden. Ob dieses Ziel erreicht wird, hängt jedoch nicht nur von der behandelten Materie ab, sondern auch von der Form, in der sie dargeboten wird.

Ein wissenschaftlicher Rapport wird sein Ziel am ehesten dann erreichen, wenn er kurz und eindeutig und in (für den vorgesehenen Leserkreis) allgemein verständlicher Weise abgefasst ist und alle zum Verständnis und zur selbständigen Urteilsbildung wesentlichen Daten enthält. Um Wiederholungen zu vermeiden sei nur darauf hingewiesen, dass alle in den vorangegangenen Kapiteln behandelten Punkte bei der Darstellung der Ergebnisse berücksichtigt werden müssen ó auch wo dies nicht ausdrücklich gefordert wurde. Zum Forschungsbericht gehört aber nicht nur die ausführliche Beschreibung des Untersuchungsganges, sondern auch die Mitteilung und Erläuterung bewusster Auslassungen. Etwaige Schwächen oder Fehler, die im Laufe der Untersuchung entdeckt wurden oder sich eingeschlichen haben, sollten unbedingt mitgeteilt und diskutiert werden. Gerade solche Mitteilungen können den Weg für weitere Forschungen ebnen und Um- und Irrwege in Zukunft vermeiden helfen. Die sind daher von grossem Informationswert und tun dem Ansehen des Experimentators ganz sicher keinen Abbruch.

Sind diese Bedingungen erfüllt, dann ist es nur ein zusätzliches Plus, wenn der Bericht auch noch in einem flüssigen, leicht lesbaren Stil geschrieben ist.

Spezielle Fachausdrücke und Symbole, deren Kenntnis beim Leser nicht vorausgesetzt werden kann, sollten mindestens einmal erklärt oder definiert werden. Statt umfangreicher Erklärungen kann oft der Hinweis auf Spezialliteratur ausreichend sein.

Illustrationen und Tabellen ergänzen und erleichtern die Lektüre und machen sie vor allem anschaulicher. Oft erlauben sie eine wesentliche Straffung des Textes. Sie haben allerdings nur dann einen Wert, wenn einige Grundbedingungen erfüllt sind: Tabellen sollten nicht nur beziffert werden, sondern mit einem aufschlüsselnden Text versehen sein. Für immer wieder vorkommende statistische Werte und Methoden (Korrelationskoeffizienten, Mittel- und Streuungswerte, Korrelationsmatrizen, Faktorenanalysen) haben sich bestimmte, vereinfachende Darstellungsformen herausgebildet. Dies gilt sowohl für die verwendeten Symbole als auch für die Darstellung (z. B. Fortlassung von Nullen vor dem Dezimalzeichen und neuerdings sogar dieses Zeichens bei Korrelationskoeffizienten). Die Anpassung der eigenen Daten an gebräuchliche Darbietungsarten erleichtert dem Leser die Orientierung wesentlich. Gut bezeichnete und sorgfältig zusammengestellte Tabellen erlauben es geübten Lesern oft sogar sich auch ohne Textlektüre einen schnellen Überblick über Zielsetzung, Umfang und Ergebnisse einer Arbeit zu verschaffen.

Wer dagegen eigene Modelle entwickelt, muss bei Misslingen damit rechnen, dass seine Botschaft den Leser nicht oder nur unvollkommen oder erst nach mühseligem Studium erreicht (soweit sich der Leser dem zu unterziehen bereit ist). Selbst wertvolle Vorschläge und Ergebnisse können davon ó zum Nachteil nicht nur des Verfassers, sondern der ganzen Wissenschaft ó betroffen werden.

### Literaturverzeichnis

- Adolfs, K.*: Faktorenanalytische Untersuchung der gebräuchlichen Handschriftenvariablen. Diss. 1964 (U. Freiburg i. Brg.).
- Fahrenberg, J.*: Graphometrie. Diss. 1961 (U. Freiburg i. Brg.).
- Fischer, G.*: Zur faktoriellen Struktur der Handschrift. Zeitschrift f. experimentelle u. angewandte Psychologie, 1964, 11, 254-280.
- Hofsommer, W., Holdsworth, R., und Seifert, T.*: Reliabilitätsfragen in der Graphologie. Psychologie und Praxis, 1965, 9, 14-24.
- Lockowandt, O.*: Faktorenanalytische Validierung der Handschrift mit besonderer Berücksichtigung projektiver Methoden. Diss. 1966 (U. Freiburg i. Brg.).
- Schneevoigt, I.*: Graphologische Intelligenzdiagnose. Bouvier, Bonn 1968.
- Seifert, T.*: Faktorenanalyse einiger Schriftmerkmale. Zeitschrift f. experimentelle u. angewandte Psychologie, 1964, 11, 645-666.
- Stein Lewinson, Th.*: Entwicklung einer ausdruckskundlichen Faktorenanalyse der Handschrift mit Hilfe des Elektronenrechners. Zeitschrift für Menschenkunde, 1968, 32, 393 ó 413.
- Timm, U.*: Graphometrie als psychologischer Test? Diss. 1965 (U. Freiburg i. Brg.).
- Wallner, T.*: Experimentelle Untersuchungen über die Reliabilität direkt messbarer Handschriftenvariablen. Zeitschrift für Menschenkunde, 1961, 25, 49-78.
- Wallner, T.*: Graphologie als Objekt statistischer Untersuchungen. Psychologische Rundschau, 1965, 16, 282-298.
- Wallner, T.*: Zusammenhänge zwischen Prognosedaten, Handschriftenvariablen und Ausbildungsergebnissen. Psychologische Beiträge, 1966, 9, 449-459.
- Wallner, T.*: Zusammenhänge zwischen graphischen Variablen und Persönlichkeitsbeurteilungen. Zeitschrift für Menschenkunde, 1968, 32, 438-445.

TEUT WALLNER

*Wallner, T.:* Uttagning till kontoristutbildning. Stockholm: Rapport från Personaladministrativa rådet 0019, 1969. ó Deutsch: Der prognostische Wert von Tests und Handschriftenvariablen bei Eignungsuntersuchungen. (In Druck)

*Wallner, T.:* Die Reliabilität schriftpsychologischer Begutachtungen, Zeitschrift für Menschenkunde, 1969, 33, 191-197[a].

*(Dipl.-Psych. Teut Wallner, Rottnevägen 31, S ó 121, 61 Hohanneshov, Schweden)*

Erschienen in Zeitschrift für Menschenkunde, 34, S. 280-300.

Kopiert und in PDF-Format konvertiert von RedaktionGrafologie.ch