

Die zweifelhafte Validität forensischer Schriftvergleichung bei Unterschriftenfälschungen

von

Jürgen Wawersik

Die Validität schriftvergleichender Untersuchungen von Handschriften ist im juristischen Schrifttum mehrfach in Frage gestellt worden^{1, 2, 3, 4} und wurde auch im Schrifttum des Fachgebietes kritisch erörtert.^{5, 6, 7} Das Beispiel eines schriftvergleichenden Gutachtens im Zusammenhang mit einem Prozess vor dem Landgericht Leipzig (AZ 15 O 1187/06) ist geeignet, die von anderer Seite geäußerten Zweifel zu konkretisieren.

Der Kläger, der verpflichtet war, an den Beklagten 72.000,00 EUR zu zahlen, versuchte, die Vollstreckung mit der Begründung abzuwehren, die Forderung sei an einen Dritten abgetreten. Zum Beweis legte er ein angeblich vom Sohn des Beklagten unterzeichnetes Schriftstück vor. Der Beklagte bestritt die Echtheit der Unterschrift. Das Gericht bestellte ein graphologisches Gutachten bei Frau Dr. G. Bromm vom M.S.U. Mannheimer Schrift- und Urkundenlabor Dr. Bromm und Müller-Partnerschaft. Der Gutachterin wurden zwei Schriftstücke mit den strittigen Unterschriften sowie 32 echte Schriftproben vorgelegt.

Das Gutachten kam zu der Schlussfolgerung, die strittigen Unterschriften »stammten mit hoher Wahrscheinlichkeit« vom Namensträger. Nach Durchführung der Beweisaufnahme gelangte das Gericht zu der Überzeugung, dass die strittigen Unterschriften nicht vom Sohn des Beklagten stammten und wies die Klage ab. Die Berufung des Klägers beim Oberlandesgericht in Dresden (AZ 14 U 0730/08) war ebenfalls erfolglos.

Die Gutachterin begründet das Ergebnis mit einer »eingehenden schriftvergleichenden Analyse«, in der »neben dem Schriftsystem und ganzheitlichen Eindruckscharakteren« die »Teilaspekte der individuellen Schreibdynamik, Formung und Gliederung des Schriftbildes sowie die Dominanzverhältnisse zwischen diesen graphischen Aspekten« (s. S. 15 d. Gutachtens) gegenübergestellt wurden. Diese Formulierungen beziehen sich auf Schriftzüge, die lediglich aus drei einfachen graphischen Gebilden bestehen (s. Abb. 1 u. 2), nämlich einer Eingangsschleife und einer Ausgangsschleife, die durch einen sägezahn-, girlanden- oder arkadigen Linienzug (s. Abb. 2) miteinander verbunden sind. Ersichtlich lässt die Gestalt dieser Elemente ohne Kenntnis des Namensträgers keinen Rückschluss auf die Buchstaben zu, für die sie stehen sollen.

1. Die Methodik des Gutachtens

Die »eingehende schriftvergleichende Analyse« soll, dem Wortlaut des Gutachtens folgend, eine »systematische Merkmalsvergleiche« (s. S. 15 d. Gutachtens) zwischen den strittigen Signaturen und den Unterschriften des Namensträgers beinhalten. Wie sich herausstellt, besteht dieser Vergleich im Prinzip darin, nach Ähnlichkeiten der Strichführung zwischen den authentischen und den strittigen Signaturen zu suchen. Dabei bedient sich die Handschriftenvergleiche offenbar weitgehend unverändert einer Systematik, die von Lothar Michel⁸ vor 30 Jahren in die Handschriftenexpertise eingeführt wurde.

Anhand dieses sogenannten »Grundkomponentenmodells«⁹ werden bestimmte Eigenschaften der Schriftzüge, zum Beispiel Strichbeschaffenheit, Bewegungsführung, räumliche Ausdehnung, Flächengliederung usw. überprüft. Das geschieht jedoch ausschliesslich anhand von deskriptiven Merkmalen wie »Strichverbiegungen«, »kantig gestaltete Schleifenbasis«, »dynamische Strichsequenzen«, »Strichspannung« usw., um die Terminologie des Gutachtens zu benutzen, deren Ausprägung ohne den geringsten Ansatz einer quantitativen Objektivierung rein visuell, d. h. durch blossen Augenschein bestimmt wird und somit ausschliesslich auf der subjektiven Einschätzung des oder der Sachver-

ständigen beruht. Dabei bezieht sich die Auswertung jeweils auf einzelne begrenzte Abschnitte der Signaturen. Sofern in irgendeiner der Vergleichssignaturen ein Abschnitt Ähnlichkeit mit einem Abschnitt in den strittigen Signaturen aufweist, wird dieser Umstand als Indikator für die „Echtheit“ gewertet. Das soll im vorliegenden Fall auf nicht weniger als 42 Merkmalsausprägungen zutreffen (s. S. 17 – 24 d. Gutachtens).

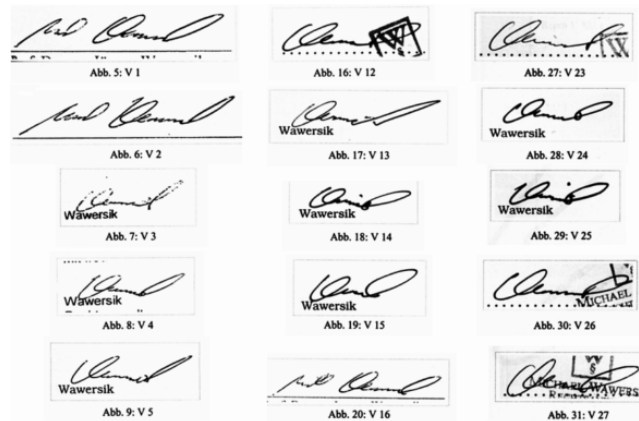


Abb. 1 Ausgewählte Unterschriften des Namensträgers (Reproduktion der Abb. 5 ff. mittels Scanner aus dem Dokumentationsanhang S. 3 – 5 des Gutachtens)

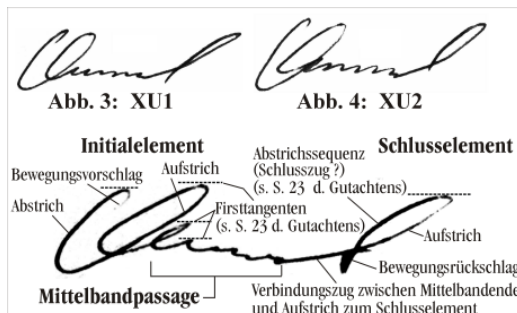


Abb. 2 Die strittigen Unterschriften XU1 und XU2 (Reproduktion der Abb. 3 und Abb. 4 mittels Scanner aus dem Dokumentationsanhang S. 3 des Gutachtens) sowie die Terminologie des Gutachtens zur »Formung und Gliederung des Schriftbildes« am Beispiel der Unterschrift V27 (Abb. 27 des Dokumentationsanhangs)

Das Heraustrennen einzelner Abschnitte mag bei längeren handschriftlichen Texten angemessen sein, um anhand von Buchstaben oder Zeichenfolgen, die mehrfach im Text vorkommen, die Eigenart und Variabilität einer Handschrift im Sinne des Grundkomponentenmodells zu beschreiben. Diese Vorgehensweise ist jedoch nicht oder allenfalls unter Vorbehalt auf Signaturen und zumal solche ohne erkennbaren Buchstabenbezug übertragbar.

2. Kritik der Methodik

Vorab sei darauf hingewiesen, dass die Autorin des Gutachtens neben ihrer Tätigkeit im »Mannheimer Schrift- und Urkundenlabor« Mitglied des »Instituts für Schrift- und Urkundenuntersuchung (ISU) e.V.« ist, das, dem Wortlaut der Internetseite »<http://www.isu-mannheim.de/>« des Instituts folgend, den „Rahmen für die Weiterentwicklung des Fachgebiets "Schrift- und Urkundenuntersuchung" in Forschung, Lehre und Praxis liefert“. Man darf deshalb davon ausgehen, dass die methodische Vorgehensweise, die im Folgenden kritisch erörtert werden soll, der Lehrmeinung des Fachgebietes der forensischen Schriftvergleichung entspricht.^{10, 11, 12, 13, 14}

Im Rahmen der gerichtlichen Auseinandersetzung wurden gegen das Gutachten folgende Einwände vorgetragen:

1. Die 32 Schriftproben des Namensträgers bestehen ausnahmslos aus drei graphischen Komponenten, die keinerlei Buchstabenbezug haben (s. Abb. 1). Das Fachschrift-

tum beschreibt Unterschriften dieser Art als stark »bewegungsorientiert bei mangelndem Buchstabenbezug«, die deshalb »eher Paraphencharakter« haben.¹⁵ Das Missverhältnis zwischen den kurzen Schriftproben und der grossen Anzahl der angeblich charakteristischen Merkmale spricht für eine »*Hypertrophierung des Zuverlässigkeitsgrades des Ergebnisses*«, einer Gefahr, der, dem Schrifttum folgend, selbst die »*besten Schriftsachverständigen gelegentlich unterliegen*«. ¹⁶

2. Die Authentizität der strittigen Signaturen wird wesentlich damit begründet, dass die graphischen Elemente, aus denen sie sich zusammensetzen, innerhalb der Variationsbreite der Signaturen des Namensträgers liegen. Ersichtlich ist die Variationsbreite der authentischen Vergleichssignaturen ausserordentlich gross (s. Abb. 1). Dabei dürfte unbestreitbar sein, dass eine Reihe von Extremwerten in der Ausprägung der Initial- und Schlusselemente sowie der Mittelbandpassage (s. Abb. 1) dem Umstand zuzuschreiben sind, dass sie unter atypischen Randbedingungen (z.B. im Stehen, in Eile, usw.) entstanden sind. Unter diesen Voraussetzungen dürfte es dem Fälscher im vorliegenden Fall nicht besonders schwer gefallen sein, eine Nachahmung herzustellen, die innerhalb dieser Grenzen liegt. Auch das Fachschrifttum räumt ein,^{17, 18} dass es »*einem geschickten Fälscher, der sich die Mühe macht, gründlich zu üben, gelingen kann, ein Nachahmungsprodukt zu erzeugen, dessen Fälschung nicht erkannt, geschweige denn nachgewiesen werden kann. Das ist besonders dann möglich, wenn die echte Unterschrift ihrer Beschaffenheit nach leicht zu fälschen ist.*« Deshalb darf die Feststellung der Gutachterin bezweifelt werden, dass der Fälscher im vorliegenden Fall besonders »*beobachtungsstark*« und »*umsetzungsicher*« gewesen sein musste (s. S. 28 d. Gutachtens), um bei der Reproduktion der drei primitiven graphischen Gebilde, aus denen sich die Signaturen zusammensetzen, die extrem grosse Variationsbreite der Vergleichsunterschriften nicht zu überschreiten.
3. Die strittigen Schriftzüge wurden zur Befunderhebung in einzelne Abschnitte aufgetrennt, um sie mit korrespondierenden Abschnitten der Signaturen des Namenssträgers zu vergleichen. Völlig unberücksichtigt blieb dabei die Ausprägung der benachbarten Abschnitte des jeweiligen Schriftzuges. Man darf unterstellen, dass Unterschriftsleistungen der vorliegenden Art in einer einzigen, zusammenhängenden und zügigen Bewegung entstehen. Wenn die Handschriftenvergleichung diesen Vorgang als »*graphomotorische Automatisierung*« bezeichnet, so erstreckt sich der Bewegungsfluss dann aber doch wohl stets auf das Gesamtgebilde. Deshalb erscheint es unzulässig, einzelne Segmente des Linienzuges aus dem Zusammenhang herauszulösen und mit einem ebenso isolierten Segment aus einer anderen Signatur zu vergleichen. Genau das geschieht aber im vorliegenden Gutachten. Die Befunde bestehen dabei zum Teil aus Spitzfindigkeiten wie »*leichte Strichverbiegungen im Bereich der bogenförmigen Initiale*« oder »*etwas kantig gestaltete Schleifenbasis der zweiten Oberlänge*« (s. S. 18 des Gutachtens) oder sie werden mit Attributen wie »*leicht*«, »*etwas*«, »*schwach*«, »*relativ*« usw. versehen. An keiner Stelle wird die Manifestation einzelner Elemente zu der Ausprägung anderer Elemente innerhalb der gleichen Unterschrift im Sinne eines funktionalen Zusammenhanges des automatisierten Bewegungsablaufs in Beziehung gesetzt.
4. Im Schlusselement der zwei strittigen Unterschriften fehlt ein Merkmal, das in allen Signaturen des Namensträgers vorhanden ist und dessen Unterschriften in einer besonderen und individuellen Weise charakterisiert, nämlich ein sogenannter Bewegungsrückschlag am Ende des Schlusszuges (s. Abb. 3).

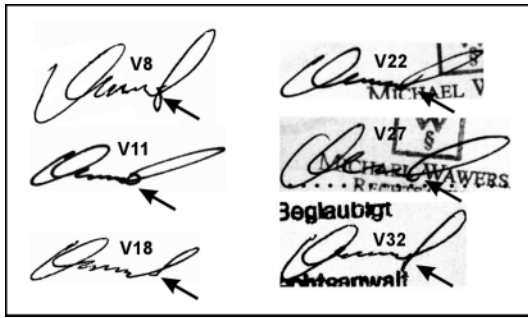


Abb. 3 Der sogenannte »Bewegungsrückschlag« in den Signaturen des Namensträgers

Diesen Einwänden begegnet die Gutachterin mit der Feststellung, »die [...] schriftvergleichenden Analysen [haben] den Eindruck weitgehender Übereinstimmung [...] bestätigt«. Das gelte »insbesondere für diejenigen Merkmalsaspekte, die sich durch einen höheren Grad graphomotorischer Automatisierung, Schreibdynamik und Eigenprägung auszeichnen« (S. 27 d. Gutachtens).

Die Schlussfolgerung, dass die strittigen Unterschriften mit »hoher Wahrscheinlichkeit vom Namenseigner« stammen, beruhe unter anderem darauf, dass »der Sachverständige aufgrund seines **Wissens** und seiner **Erfahrung** die Konfiguration der Befunde in eine summarische Schlussfolgerung« zusammenführt«. »Selbstverständlich« sei berücksichtigt worden, »bei welchen Merkmalen es sich um Übereinstimmungen der **zentralen Tendenz** der Vergleichsschriften oder um Randvarianten« handle (S. 13 der ergänzenden Stellungnahme¹⁹).

Besonders bemerkenswert ist die Begründung, warum der Bewegungsrückschlag im Schlusselement der Vergleichsunterschriften keinen Eingang in die Bewertung gefunden hat: In schriftvergleichenden Untersuchungen würde geprüft, »ob eine strittige Schreibleistung innerhalb der Variationsbreite der Vergleichsschriften liegt oder nicht«. Damit sei »die Richtung des Vergleichs eindeutig vorgegeben: Im vorliegenden Fall waren XU1 und XU2 mit V1 – V32 zu vergleichen und nicht etwa die Vergleichsschriften mit den strittigen Signaturen« (s. S.13 u. S. 25 der ergänzenden Stellungnahme). Ein Fälscher kann somit die Aussicht, mit seiner Fälschung erfolgreich zu sein, allein dadurch erhöhen, dass er sich bei der Nachahmung auf die primitiven Elemente einer Unterschrift beschränkt, dabei aber die besonders charakteristischen, jedoch schwieriger nachzuahmenden Elemente weglässt.

Im Übrigen führt die Gutachterin in ihrer Erwiderung die Begriffe »Erfahrung des Sachverständigen« und »zentrale Tendenz der Vergleichsschriften« ein, die eingehender erörtert seien.

2.1. Die Erfahrung des Sachverständigen

Die Untersuchung von zwei Schriftproben daraufhin, ob sie vom gleichen Urheber oder von zwei verschiedenen Urhebern stammen, entspricht im Prinzip dem Sachverhalt anderer Fachbereiche, zum Beispiel der Medizin, mit Hilfe von geeigneten Methoden eine vermutete Diagnose nachzuweisen oder auszuschließen.

Tab. 1 Die Analogie zwischen der Spezifität (richtig negative Rate) und Sensitivität (richtig positive Rate) von diagnostischen Testverfahren und Methoden zur Identifikation von freihändigen Schriftnachahmungen

Schriftprobe	Test positiv	Test negativ
gefälscht	richtig positiv	falsch negativ
echt	falsch positiv	richtig negativ

Die Validität der Ergebnisse setzt allerdings voraus, dass zuvor in gesonderten Untersuchungen überprüft wurde (s. Tab. 1), wie oft die betreffende Methode unter der Bedingung positiv ausfällt, dass die Diagnose tatsächlich vorliegt (sog. Sensitivität) und wie oft sie unter der Bedingung negativ ausfällt, dass die Diagnose *nicht* vorliegt (sog. Spezifität).

In abgewandelter Form benutzt die Informatik diesen Ansatz, um Methoden zur computergestützten biometrischen Authentifizierung oder Verifizierung von Signaturen zu validieren.^{20, 21} Ermittelt wird dabei (s. Abb. 7) die sogenannte FMR (*»false match rate«*) oder FAR (*»false acceptance rate«*), die den *»falsch negativen Ergebnissen«* des Ansatzes in der Diagnostik (s. Tab. 1) entspricht und die FNMR (*»false non-match rate«*) oder FRR (false rejection rate), die den *»falsch positiven Ergebnissen«* der Diagnostik entspricht.

Erst aufgrund derartiger Daten liesse sich abschätzen,

- wie gross die Wahrscheinlichkeit ist, eine Schriftprobe für echt, d.h. vom Namensträger angefertigt zu erklären, obwohl es sich um eine freihändige Nachahmung handelt und
- wie gross die Wahrscheinlichkeit ist, eine Schriftprobe zu einer freihändigen Nachahmung zu erklären, obwohl sie vom Namensträger stammt.

Im Schrifttum der forensischen Schriftvergleichung sind nur wenige Mitteilungen in dieser Richtung auffindbar. Risinger et al. (1989) zitieren diesbezügliche Studien der *Forensic Sciences Foundation (FSH)* in den USA aus den Jahren 1984 – 1987. Die Ergebnisse wurden von der FSH nicht publiziert, sondern erst von Risinger et al. kritisch gewürdigt. Folgt man deren Schlussfolgerungen, dann gab es bei allen Erhebungen der vier Jahrgänge nur 18% Antworten, die in jeder Hinsicht fehlerfrei waren.

Im deutschsprachigen Schrifttum konnte lediglich eine Dissertationsschrift²² ausfindig gemacht werden, die über ähnliche Erhebungen berichtet. Offenkundig war das Ergebnis, gemessen an den Untersuchungen der FSF, nicht günstiger. Der Autor kam unter anderem zu der Schlussfolgerung:

»[...] die vorhergesagte Überlegenheit von Sachverständigen in der zutreffenden und sicheren Erkennung von Unterschriftenfälschungen [...] hat sich [...] nicht bestätigt«.²³ »[...] der Zusammenhang zwischen Richtigkeit von Vergleichsurteilen und dem Sicherheitsniveau der Befunde [...] überschreite [...] bei den Experten ebenso wenig den Zufallsbereich wie bei Laien. Der Sicherheitsgrad [...] ist demnach ähnlich wie in anderen Bereichen diagnostischer Urteilsbildung nahezu unbrauchbar«.²⁴

Trotz des begrenzten Umfanges dieser Studien dürfte sich demnach kaum bestreiten lassen, dass Schriftvergleichen einer erheblichen Irrtumswahrscheinlichkeit unterliegen. Dennoch wäre es unangemessen, die schriftvergleichende Expertise pauschal zu diskreditieren. In Anbetracht der Vielfalt und Komplexität von Handschriften²⁵ wird der Spezialist bei der Suche nach Übereinstimmungen oder Unterschieden zur Verifizierung von Fälschungen Feinheiten erkennen, die dem Ungeübten entgehen. Es erscheint dagegen höchst fragwürdig, allein deshalb, weil eine strittige Handschrift keine groben Abweichungen aufweist, auf deren Authentizität zu schliessen, denn die graphische Begabung des Fälschers bleibt im konkreten Fall stets unbekannt. Das Fachgebiet der forensischen Schriftvergleichung hat sich vielfach mit den Ausprägungen und Eigenarten von Handschriften sowie den Artefakten bei freihändigen Nachahmungen auseinandergesetzt, um Fälschungen nachzuweisen. Im Gegensatz dazu sucht man dagegen vergeblich nach Untersuchungen, die sich mit den Grenzen der Entscheidbarkeit zwischen Fälschung und Authentizität befassen. Deshalb ist es nicht legitim, wenn ein Gutachter die Validität seiner Schluss-

folgerung mit dem Hinweis auf sein persönliches Wissen und seine persönliche Erfahrung begründet, solange er keine Angaben darüber machen kann, wie viele Expertisen er bislang abgegeben hat und wie viele davon »richtig positiv oder negativ« und »falsch positiv oder negativ« waren.

2.2. Zentrale Tendenz

Wie bereits dargelegt, beschränkt sich die Befundbewertung im Gutachten darauf, einzelne Segmente der strittigen Signaturen mit den entsprechenden Segmenten der authentischen Signaturen zu vergleichen. Dabei wurde lediglich die qualitative Manifestation der Ausprägungen als solche bewertet, ohne das Ausmass der quantitativen Abweichung und ohne die Unterschiede in der Gesamtkonfiguration der jeweiligen Vergleichsunterschrift zu berücksichtigen. Als Beispiel dafür seien ein sogenannter »Bewegungsimpuls am Ende der Mittelbandpassage in XU1« und ein »Bewegungsimpuls in Gestalt eines deckzügigen Winkelzuges in XU2« angeführt, die in ähnlicher Gestalt auch in V9, V24 und V25 vorkommen sollen (s. Abb. 4). Selbst wenn man dem zustimmen will, sind die Verschiedenheit der Gesamtkonfiguration der Vergleichsunterschriften, in denen dieses Merkmal auftaucht, z.B. in Bezug auf Grösse, Neigungswinkel, Position des Merkmals innerhalb der Signatur sowie der räumlichen Verhältnisse der drei Grundelemente des Gesamtgebildes zueinander unübersehbar (s. Abb. 4).

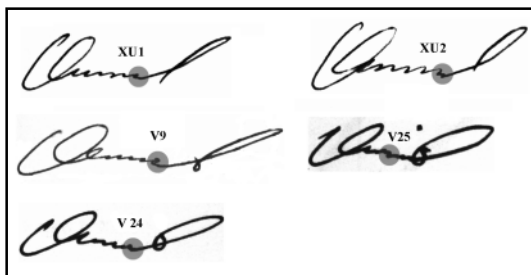


Abb. 4 Ein angeblicher »Bewegungsimpuls« am Ende der Mittelbandpassage in XU1 und ein »deckzügiger Winkelzug« in XU2, die in ähnlicher Gestalt in V9, V24 und V25 nachweisbar sein sollen

Auch andere Merkmale treten in den Vergleichsunterschriften jeweils nur dann auf, wenn die betreffende Signatur zugleich auch an anderer Stelle eine weitere Besonderheit aufweist, ein Umstand, der durchaus plausibel ist, wenn man unterstellt, das sich der Bewegungsablauf im Sinne einer »graphomotorischen Automatisierung« bei unterschiedlichen Randbedingungen in ebenso unterschiedlicher Weise in der Ausprägung der konsekutiven Elemente der Signatur niederschlägt.

Die Fragwürdigkeit dieser Art der Befundbewertung sei am Beispiel der »stark nach rechts geneigten Oberlängenschleife zu Beginn des Namenszuges« (s. S. 23 d. Gutachtens) aufgezeigt. Die Befundbewertung des Gutachtens beruht in diesem Fall auf drei Hilfslinien, die durch das Initial- und das Schlusselement der strittigen Signatur XU2 und einer einzigen Vergleichssignatur V4 gelegt wurden (s. Abb. 5).

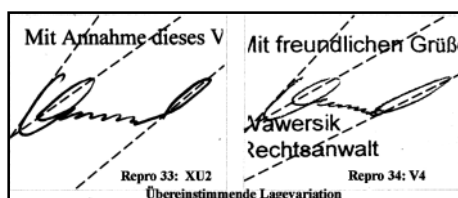


Abb. 5 Repro 33 und Repro 34 des Gutachtens zum Nachweis einer »Lagevariation«

Die Gegenüberstellung der Merkmalsausprägung bei lediglich zwei Signaturen dürfte kaum ausreichen, um Rückschlüsse auf die »Variation« dieses Merkmals zu ziehen. Dennoch leitet die Gutachterin daraus eine »übereinstimmende Lagevariation« ab. Noch weniger nachvollziehbar sind bei dieser Vorgehensweise Rückschlüsse auf eine »zentrale Tendenz«.

Stattdessen hätte man erwartet, dass die verschiedenen Merkmalsausprägungen jeweils an sämtlichen 32 authentischen Signaturen gemessen werden, um daraus statistische Parameter, zum Beispiel den Mittelwert und die Standardabweichung als Mass für eine zentrale Tendenz zu bestimmen.

So ergibt sich zum Beispiel für den Neigungswinkel der »Eingangsschleife« ein Mittelwert von $28,6^\circ$ und eine Standardabweichung von $\pm 3,96^\circ$. Ersichtlich liegt der Neigungswinkel der strittigen Signaturen deutlich ausserhalb dieses Bereiches (s. Abb. 6).

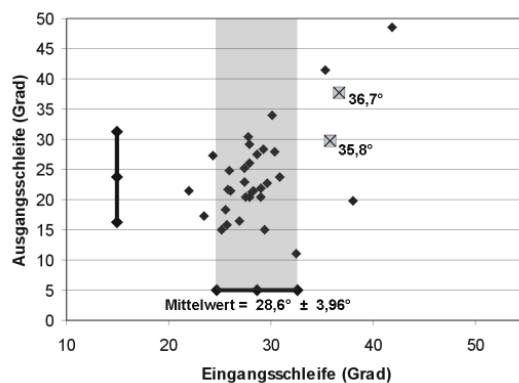


Abb. 6 Der Neigungswinkel des Initialelementes (»Eingangsschleife«) und des Schlusselementes (»Ausgangsschleife«) in den Signaturen des Namensträgers (■) und den strittigen Signaturen (X)

Dieser Befund deckt sich mit Beobachtungen einer Studie über die graphometrische Bestimmung der Merkmalsausprägungen von freihändigen Nachahmungen.²⁶ Dem folgend haben Nachahmungen die Tendenz zu steileren, d. h. grösseren Neigungswinkeln.

In gleicher Weise kann man weitere Ausprägungen (z. B. Breite und Höhe der Signaturen, Breite der Mittelbandpassage, Neigung des Mittelbandes zur Zeilenbasis usw.) metrisch erfassen. Dabei zeigt sich, dass mehrere andere Merkmalsausprägungen bei den strittigen Signaturen gleichfalls ausserhalb des Bereiches von Mittelwert $\pm$ Standardabweichung liegen. Zwar gibt es bei den Signaturen des Namensträgers, ähnlich wie bei dem Neigungswinkel vereinzelt auch Extremwerte bei anderen Merkmalen der authentischen Signaturen, die jedoch im Gegensatz zu den strittigen Signaturen in keinem Fall gemeinsam in einer einzigen Signatur auftreten. Ausserdem fehlt in beiden strittigen Signaturen der Bewegungsrückschlag im Schlusselement, der bei allen Vergleichssignaturen vorhanden ist.

Es mag zutreffen, dass Extremvarianten als Entscheidungskriterium für oder gegen die Authentizität von Unterschriften einen gewissen Stellenwert haben können. Jedoch erscheint schon die Wahrscheinlichkeit für das Auftreten mehrerer Extremwerte in einer einzigen Unterschrift gering. Umso unwahrscheinlicher dürfte es sein, dass eine derartige Merkmalskombination in zwei Signaturen auftritt, die kurz nacheinander unter den gleichen Randbedingungen erstellt wurden.

Die Gutachterin weist diesen Ansatz jedoch mit der Begründung zurück, dass »nach dieser Logik ca. 32% der Vergleichsunterschriften, deren Messwerte ausserhalb der jeweiligen Standardabweichung liegen, zu der Schlussfolgerung führen müssten, dass sie nicht aus der Hand ihres manuellen Urhebers herrühren« (s. S. 11 der ergänzenden Stellungnahme).

3. Fragestellung oder Hypothese

Dem Fachschrifttum folgend, besteht die Befundbewertung schriftvergleichender Untersuchungen im Wesentlichen darin, subjektiv abzuschätzen, inwieweit sich die Merk-

malsausprägungen einer Schreibleistung mit bestimmten Hypothesen in Einklang bringen lassen oder mit anderen Worten: Inwieweit sie die Hypothesen erklären.^{27, 28}

Freilich sind diese Hypothesen nicht mit dem vergleichbar, was man darunter in der Prüfstatistik versteht. Dort werden Hypothesen aufgrund von Beobachtungen oder der Vermutung eines bestimmten Sachverhaltes aufgestellt. Dem folgen Experimente, Untersuchungs- oder Beobachtungsreihen, deren Daten geeignet sind, die Vermutung zu bestätigen oder zu widerlegen. Da die Daten und Prüfwerte keine festen Grössen sind, sondern im Bereich einer bestimmten Wahrscheinlichkeitsfunktion variieren, unterliegen Annahme und Ablehnung der Hypothese einer Irrtumswahrscheinlichkeit, die von dem Signifikanzniveau eines Prüfwertes abhängt. Diese Signifikanz- oder Ablehnungsgrenze muss vor Beginn des Experimentes festgelegt werden, je nachdem, welches Risiko man bei der abschliessenden Testentscheidung in Kauf nehmen will: Die Nullhypothese abzulehnen, obwohl sie zutrifft oder umgekehrt die Alternativhypothese anzunehmen, obwohl sie nicht zutrifft.

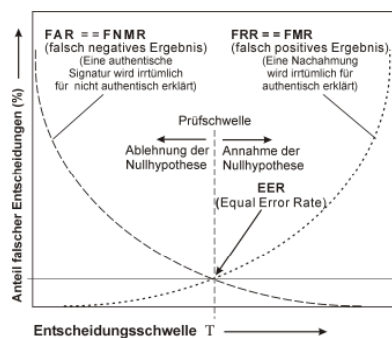


Abb. 7 Die Häufigkeit falscher Entscheidungen in Abhängigkeit vom Prüfwert (modifiziert nach Plamondon u. Lorette)

Nun kommt es bei der Schriftvergleichung nicht darauf an, eine Hypothese durch geplante prospektive Studien zu beweisen oder zu widerlegen, sondern der Untersucher soll eine Diagnose stellen. Es bedarf deshalb eigentlich keiner Hypothesen, sondern lediglich einer Methode, die unabhängig vom subjektiven Eindruck des Untersuchers in der Lage ist, Signaturen, die vom Namensträger stammen, von freihändigen Nachahmungen zu trennen. Auch dabei ist aber das komplementäre Risiko von Fehldiagnosen zu bedenken. Die Validierung von biometrischen Methoden zur computergestützten Authentifizierung oder Verifizierung von Signaturen (s. Abb. 7) zeigt dies deutlich. Eine hohe Prüfschwelle erhöht das Risiko, Fälschungen für authentisch zu erklären und umgekehrt.

Theoretisch sollte die optimale Signifikanzschwelle im Schnittpunkt der beiden Kurven liegen, dem Punkt der »Equal Error Rate« (s. Abb. 7), also einem Prüfwert, bei dem die Risiken eines falsch negativen und falsch positiven Resultats gleich gross sind.

Inwieweit sich diese Ansätze bislang in der praktischen Anwendung bewährt haben, muss an dieser Stelle nicht erörtert werden. Es ist weder die Aufgabe noch das Anliegen dieses Beitrages, Lösungen anzubieten. Der kursorische Exkurs in die Problematik statistischer Gütekriterien dient lediglich dem Zweck, auch den mathematisch-naturwissenschaftlich Unkundigen von der fragwürdigen Vorgehensweise bei Schriftvergleichungen zu überzeugen, nämlich unter Verzicht auf jede Form einer quantitativen Datenerhebung Entscheidungen für oder gegen die Fälschung einer Signatur zu treffen.

Es sei hinzugefügt, dass es inzwischen, folgt man dem Schrifttum, neben den Verfahren der statistischen Datenanalyse weitere Ansatzpunkt zur Verifizierung von Signaturen gibt, die auf der Anwendung neuronaler Netzwerke beruhen.²⁹ Aber auch deren Grundlage sind natürlich quantitative Daten.

Entsprechenden Einwänden begegnet die Gutachterin mit der Feststellung, »der Zweck der Hypothesenbildung in wissenschaftlichen Arbeiten besteht nicht im Gebrauch der Mathematik, sondern in wissenschaftlich fundierten Erklärungen von Sachverhalten in der Realität« (s. S. 4 der ergänzenden Stellungnahme).

Angeblich müssten die Hypothesen »alle relevanten Erklärungsmöglichkeiten über die Entstehung einer begutachteten Schreibleistung abdecken« und zwar vor dem »Hintergrund der Kenntnisse und Erfahrungen« des Sachverständigen, um das »Risiko zu min-

dern, wesentliche Erklärungsansätze ausser acht zu lassen« (s. S. 3-4 der ergänzenden Stellungnahme). Es fragt sich, welche Risiken im vorliegenden Fall bei der Hypothesenformulierung bestehen, die nichts anderes als eine spezielle Formulierung der Fragestellung ist, nämlich:

H_0 : »Die strittigen Unterschriften sind authentisch« und H_1 : »Die strittigen Unterschriften sind gefälscht«.

Dabei wird man vom Sachverständigen erwarten, dass er aufgrund seiner »Kenntnisse und Erfahrungen« geeignete Daten und Kriterien aufzeigt, um diese Fragestellung objektiv und nachvollziehbar zu beantworten oder sie im Zweifelsfall für unentscheidbar zu erklären, nicht aber, um sie durch »bewusste Hypothesenreflektion« aufzuklären.

4. Die Subjektivität schriftvergleichender Befundbewertung

Die Reduktion der Befunderhebung auf die blosse visuelle Einschätzung von Merkmalsausprägungen und deren Vereinbarkeit mit der Nullhypothese erweitert den Deutungsspielraum von Schriftvergleichen fast bis zur Beliebigkeit. Vollständig ausgeblendet wird dabei die Wahrscheinlichkeit für das Auftreten von Extremvarianten. Im vorliegenden Fall lässt sich quantitativ nachweisen, dass in den strittigen Signaturen mindestens sechs Merkmale ausserhalb der zentralen Tendenz der authentischen Signaturen liegen. Selbst dann, wenn man dem Gutachten beipflichten mag, dass »jede Unterschrift aus der Stichprobe V1 - V32 im Hinblick auf irgendeine Merkmalsausprägung eine Extremvariante aufweist« (s. S. 25 d. Erwiderung), so gibt es doch keine Signatur des Namensträgers, in der alle sechs »Extremvarianten« gleichzeitig vorkommen. Die Unwahrscheinlichkeit der Realisation einer derartigen Merkmalskombination in einer einzigen Signatur durch den Namensträger wird nicht in Erwägung gezogen. Für entsprechende Einwände hält die Gutachterin das Argument bereit, wonach »sich die Befundlücke auf Mängel des Vergleichsmaterials zurückführen lässt, die an weiteren, spontan und zeitnah entstandenen Unterschriftenleistungen möglicherweise zu schliessen ist« (s. S. 9 d. Gutachtens).

Geradezu willkürlich ist der Umgang mit der Variabilität der Signaturen. Diesbezüglich behauptet die Gutachterin, die Variationsbreite der Vergleichsunterschriften sei »nicht ungewöhnlich gross« (s. S. 10 d. ergänzenden Stellungnahme), es wirke sich aber »positiv aus, dass die Variabilität der Schriftelemente innerhalb von zwei Unterschriften analysiert werden kann« (s. S. 8 d. Gutachtens). Diese Feststellung ist in Anbetracht der grossen quantitativen Variationsbreite der authentischen Signaturen (s. Abb. 1 u. 2) kaum nachvollziehbar, während die beiden strittigen Signaturen nahezu deckungsgleich sind (s. Abb. 2 u. 8).

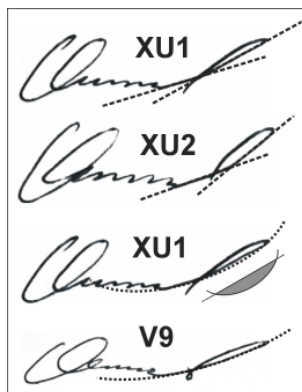


Abb. 8 Die Konfiguration des aufsteigenden Linienzuges im Schlusselement der strittigen Signaturen im Gegensatz zu der Signatur V9 des Namensträgers

Vollends spekulativ ist die Befundbewertung des Schlusselementes (s. Abb. 8). Im Gegensatz zu allen Vergleichsunterschriften besteht der aufsteigende Linienzug, der das letzte Element der Mittelbandpassage mit dem Scheitelpunkt des Schlusselementes verbindet, bei den strittigen Signaturen aus zwei Geraden, die einen stumpfen Winkel bilden (s. Abb. 8). Die Gutachterin beschreibt diesen Befund in der Rubrik »Bewegungsführung und Formgebung« als »kantige Formgebung des Schlusszuges«, die dem Schlusszug in den Vergleichssignaturen V7, V9 und V19 (siehe Abb. 14) entsprechen soll. Tatsächlich folgt dieses Element in den authentischen Signaturen jedoch einem Bogen (s. Abb. 8) oder ist eine durchgehende Gerade ohne Kante.

Das Gutachten räumt zwar ein, dass es sich hierbei um eine Merkmalsbesonderheit handelt, knüpft daran aber die Feststellung, dieser Unterschied sei unter der Hypothese der Fälschung kaum schlüssiger zu erklären als unter der Annahme, dass die Unterschrift vom Namensträger stammt (s. S. 28 d. Gutachtens). Die Begründung dafür ist besonders bemerkenswert:

»Der abweichende Verlauf einer Schleifenbasis stellt eine recht offensichtliche Diskrepanz dar, die ein Fälscher, dem es um eine möglichst formgetreue Reproduktion geht, kaum anstrebt. Es dürfte einem Fälscher darüberhinaus wenig Nachahmungsprobleme bereitet haben, eine rundere Schlinge herzustellen«.

Diese Deutung ist zugleich auch für die Einseitigkeit des Gutachtens bezeichnend, die wohl damit zusammenhängt, dass »bei der Hypothesenbildung die Fragestellung mit den Ergebnissen einer ersten Materialprüfung zusammengeführt« wird. Es liegt auf der Hand, dass die »Einbeziehung einer ersten Materialprüfung« schon vor der Befunderhebung eine »Wunschhypothese« begründet, die dann in der abschliessenden Befundbewertung zu der Feststellung führt, dass der »aus den Vorprüfungen entstandene Eindruck weitgehender Übereinstimmung bestätigt und abgestützt« wurde. Diese Art der Argumentation erfüllt in keiner Weise die Erwartung einer objektiven und vorurteilsfreien Wertung der Befunde in beide Richtungen.

Auch das Fachgebiet räumt die Fragwürdigkeit dieser Vorgehensweise ein. Eine Kommission aus den Leitern der Kriminaltechniken in Bund und Ländern³⁰ hat sich unter anderem zu dieser Problematik wie folgt geäussert:

»Bei der Einschätzung der Befundwahrscheinlichkeit spielt in der Schriftvergleichung der Begriff der »Erklärbarkeit« eine zentrale Rolle. Wenn sich ein Befund unter einer Hypothese gut erklären lässt, so wird der Sachverständige ihn auch irgendwie für wahrscheinlich halten. Mit diesem Ansatz „schloss man **früher** oft auf eine klare Urheberidentität...«. **Dies war einseitig.** Die Erklärbarkeit des Befundes muss mindestens **zweiseitig** geprüft werden«.

Das Gutachten hätte demnach der abschliessenden Schlussfolgerung, wonach sich »die entstandene Konfiguration der schriftvergleichenden Untersuchungen widerspruchsfrei durch die Annahme der Urheberidentität erklären« lässt, mit der gleichen Plausibilität die Schlussfolgerung gegenüberstellen können:

In Anbetracht der Tatsache, dass die Vergleichssignaturen nur aus „begrenzten Schriftkomponenten“ bestehen (s. S. 29 d. Gutachtens), deren äussere Gestalt eine grosse Variationsbreite aufweist, und die wegen ihrer „vereinfachten Schreibweise“ keinen „aussergewöhnlichen Informationsgehalt“ (s. S. 29 d. Gutachtens) besitzen, dürfte es einer „bislang unbekannten Person“ nicht schwer fallen, im Freihandverfahren eine möglichst „formgetreue Nachahmung“ zu produzieren.

5. Die Validität schriftvergleichender Schlussfolgerungen

Die aufgeführten Beispiele zeigen, wie gross der Spielraum ist, um schriftvergleichende Befunde zugunsten der einen oder anderen Richtung zu interpretieren. Der eigene Fall erinnert an die Vorgänge um die Hitler-Tagebücher.³¹ Auch dort handelte es sich, soweit

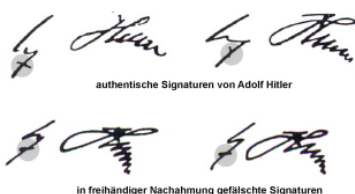


Abb. 9 Authentische und gefälschte Unterschriften von Adolf Hitler (Aus: Brandner W: Die Hitler-Tagebücher, graphologie-news.net, Ausgabe 01_09)

ersichtlich, um namhafte Vertreter des Fachgebietes, die den Dokumenten die Echtheit attestierten. Im Nachhinein meldeten sich zwar Autoren, die dem Gutachten handwerkliche Fehler vorhielten.³² Unter anderem wies Brandner³³ darauf hin, dass sich »bei dem abgekürzten Vornamen Adolf im Abstrich des »f« ein Querstrich befindet« (s. Abb. 9). »Dieser Strich ist in den Originalunterschriften immer schräg nach unten geschrieben, bis fast senkrecht, nie jedoch waagrecht, wie bei den Fälschungen«. Dies ist eine ex post getroffene Befundbewertung des Autors, die nicht ausschliesst, dass ein anderes Gutachten zu der

Schlussfolgerung hätte gelangen können, es würde einem Fälscher »wenig Nachahmungsprobleme bereitet haben«, stattdessen einen senkrechten oder abwärts gerichteten Querstrich zu produzieren.

Wie oft Expertisen, die sich auf eine blosse Schriftvergleichung stützen, den Ausgang eines Streitfalles vor Gericht massgeblich beeinflussen, ist unbekannt. Köller et al. (2004) räumen ein, dass Gerichte sich die Schlussfolgerung von Gutachten nicht unbesehen zu eigen machen müssen, sondern »unter dem Eindruck anderer Indizien« durchaus auch zum gegenteiligen Schluss kommen können. Dabei würden sich »Richter intuitiv zumeist richtig verhalten«.

Dies sollte freilich nicht »zumeist« und »intuitiv«, sondern explizit und in jedem Fall geschehen. Zweifel am Ergebnis sind umso dringender geboten, je kürzer die Schriftproben sind. Im vorliegenden Fall wurde die freihändige Nachahmung einer Unterschrift irrtümlich für echt erklärt. Wie gross das umgekehrte Risiko ist, dass Signaturen als Fälschung deklariert werden, obwohl sie vom Namenseigner stammen, ist schwer zu beantworten. Ein allgemeingültiges Rezept, um Gutachten dieser Art anzufechten, kann es in Anbetracht der Vielfalt handschriftlicher Ausprägungen nicht geben. Deshalb kann dieser Beitrag lediglich Anhaltspunkte aufzeigen, anhand welcher Kriterien sich die Plausibilität der Befunde und die Objektivität der Befundbewertung von Schriftexpertisen abschätzen und gegebenenfalls in Zweifel ziehen lässt. Zumindest bei Unterschriften und allemal dann, wenn sie nur einen schwachen Buchstabenbezug haben, gibt es guten Grund zu besonderer Zurückhaltung gegenüber der Validität von Schriftvergleichen. Als Beweismittel sollten sie deshalb allenfalls einen nachrangigen Stellenwert erhalten.

Anmerkungen:

-
- ¹ Deitigsmann O (1957)
 - ² Falck E (1956)
 - ³ Schröter G (1987)
 - ⁴ Risinger et al. (1989)
 - ⁵ Rieß M (1989)
 - ⁶ Conrad W (1970)
 - ⁷ Köller et al. (2004)
 - ⁸ Michel Lothar (1982)
 - ⁹ Seibt A (2004)
 - ¹⁰ Michel L (s. Anm. 8)
 - ¹¹ Seibt (siehe Anm. 9)
 - ¹² Seibt A (2006)
 - ¹³ Seibt A (1999).
 - ¹⁴ Seibt A (2008)
 - ¹⁵ Hecker M R (1993)
 - ¹⁶ Pfanne H (1971) S. 29
 - ¹⁷ Pfanne S. 109 (s. Anm. 17)
 - ¹⁸ Hecker S. 164 (s. Anm. 16)
 - ¹⁹ »Ergänzende Stellungnahme« vom 05. 12. 2007 des MSU Mannheimer Schrift- und Urkundenlabor
 - ²⁰ Vielhauer C. (2006)
 - ²¹ Plamondon R, Lorette G (1989)
 - ²² Conrad W (1970)
 - ²³ Conrad S. 115 (s. Anm. 23)
 - ²⁴ Conrad S. 140 (s. Anm. 23)
 - ²⁵ Bernard M (1985)
 - ²⁶ Halder-Sinn P. (1989)
 - ²⁷ Seibt (s. Anm. 13)
 - ²⁸ Seibt (s. Anm. 9)
 - ²⁹ Leclerc F, Plamondon F (1994)
 - ³⁰ Köller et al. S. 21 (siehe Anm. 7)
 - ³¹ Picker G (1992)

³² Michel L (1985)

³³ Brandner W. (o. J.) Die Hitler-Tagebücher. graphologie-news.net, Ausgabe 01_09 (Internetpublikation: <http://www.graphologie-news.net/gn/Archiv.html>)

Schrifttum

Bernard M (1985) The art of graphology, Troy, NY, Whitston Publ. Co., VIII, 416 S.

Conrad W (1970) Authentische Unterschriften und Unterschriftsfälschungen. Phil. Diss., Mannheim, XII, 228 S.

Deitigsmann O (1957) Neue juristische Wochenschrift, S. 1867-1869

Falck E (1956) Juristische Rundschau, S. 255 – 257

Halder-Sinn P. (1989) Determinanten der Schriftnachahmungsleistung. In: Conrad W, Stier B (Hrsg.) Grundlagen, Methoden und Ergebnisse der forensischen Schriftuntersuchung, Lübeck, Schmidt-Römhild, S. 273 – 296

Hecker M R (1993) Forensische Handschriftenuntersuchung, Heidelberg, Kriminalistik Verlag, IX, 399 S.

Köller N, Nissen K, Rieß M, Sadorf E (2004) Probabilistische Schlussfolgerungen in Schriftgutachten. Bundeskriminalamt (Hrsg.) Polizei + Forschung Bd. 26, München, Luchterhand, München, Wolters Kluwer Deutschland GmbH, VIII, 164 S.

Leclerc F, Plamondon F (1994) Automatic Signature Verification and Writer Identification - the State of the Art – 1989 - 1993. Internat. Journal of Pattern Recognition and Artificial Intelligence 8, No. 3, S. 643 – 660

Michel Lothar (1982) Gerichtliche Schriftvergleichung – Eine Einführung in Grundlagen, Methoden und Praxis, Berlin, Walter de Gruyter, VI, 299 S.

Michel L (1985) Urkundenexperten und die Aufklärung der Hitler-Tagebuch-Fälschung – eine kritische Nachlese, Mannheimer Hefte für Schriftvergleichung 11, 74 – 82

Pfanne H (1971) Handschriftenvergleichung für Juristen und Kriminalisten, Lübeck, Verlag Max Schmidt-Römhild, 132 S., 1 Falttafel

Picker G (1992) Der Fall Kujau: Chronik eines Fälschungsskandals, Berlin, Ullstein (Ullstein-Buch Nr. 34993; Ullstein-Sachbuch), 140 S.

Plamondon R, Lorette G (1989) Automatic signature verification and writer identification – the state of the art. Pattern Recognition 22, 107 - 131

Rieß M (1989) Die Bedeutung des schriftvergleichenden Gutachtens in Strafverfahren. In: Conrad W, Stier B (Hrsg.) Grundlagen, Methoden und Ergebnisse der forensischen Schriftuntersuchung, Lübeck, Schmidt-Römhild, S. 47 – 69

Risinger D M, Denbeaux M P, Saks M J (1989) Exorcism of ignorance as a proxy for rational knowledge: The lessons of handwriting identification “Expertise”. University of Pennsylvania law review 137, 731 – 792

Schröter G (1987) Juristische Rundschau, S. 258 – 259

Seibt A (1999) Forensische Schriftgutachten – Einführung in Methode und Praxis, München, C.H.Beck, X, 163 S.

Seibt A (2004) Qualitätsmerkmale forensischer Schriftgutachten, Zeitschrift für Schriftpsychologie und Schriftvergleichung 68, 44–62

Seibt A (2006) Forensische Handschriftenuntersuchung als Wissenschaft. Kriminalistik 10, 599 – 608

Seibt A (2008) Unterschriften und Testamente. Praxis der forensischen Schriftuntersuchung, München, Verlag C.H.Beck, XVII, 220 S.

Vielhauer C. (2006) Biometric User Authentication for IT Security. Advances in Information Security Vol. 8, Springer Science + Media Inc., XII, 284 S.

Anschrift des Verfassers:

Prof. em. Dr. med. J. Wawersik

Am Alpengarten 3/4

88131 Lindau Bodensee

e-mail: wawersik@anaesthesie.uni-kiel.de