

Über die Grenzen substantieller Faktorladungen

Thomas Weber

Die analytische Betrachtung von Steinartefakten führt zu einer großen Anzahl von Merkmalen, die an jedem Stück gemessen werden können (WEBER 1986). Diese Daten werden zur Grundlage von Aussagen besonders hinsichtlich der Technologiegeschichte alt- und mittelpaläolithischer Technokomplexe.

Da die Einzelaussagen mehr oder minder unvermittelt nebeneinanderstehen, einander z.T. sogar widersprechen, erscheint der Einsatz multivariater Verfahren sinnvoll, mit deren Hilfe eine Datenreduktion im Interesse größerer Übersichtlichkeit zu erreichen ist. Zu diesen Verfahren gehört die Faktor(en)analyse, die aus einer größeren Anzahl von Variablen über die Untersuchung ihrer Korrelationen sogenannte Faktoren abzuleiten versucht, die einen möglichst hohen Varianzanteil der gemessenen Merkmale erklären sollen. Die Faktoren stellen dabei voneinander unabhängige Größen dar, die man zunächst kausal auffassen wollte - beispielsweise in der Psychologie, für die die Methode entwickelt wurde, als «Intelligenz», «Auffassungsgabe» o.ä.. Heute spricht man neutraler davon, daß Faktoren Zusammenhänge zwischen Merkmalen wiedergeben.

Ausdruck finden diese Zusammenhänge im Ladungsmuster der Faktoren, einer Matrix der Korrelationen zwischen beobachteten Merkmalen und hypothetischen Faktoren. Die Werte können wie ein Korrelationskoeffizient zwischen +1 (höchstmöglicher positiver Zusammenhang) und -1 (höchstmöglicher negativer Zusammenhang) variieren, wobei der Wert 0 Unabhängigkeit des Faktors vom beobachteten Merkmal bezeichnet (vgl. E. WEBER 1974).

Die Berechnung erfolgt über verschiedene Techniken der Faktorenextraktion, von denen die (hier verwendete) Hauptkomponentenanalyse die bekannteste ist. Anschließend kann «rotiert» werden, d.h. das Koordinatensystem der Faktorebenen wird in bezug auf die Merkmalkonfiguration so lange gedreht, bis eine im Sinne des Nutzers optimale Wiedergabe der Merkmalzusammenhänge durch das Ladungsmuster erreicht ist. Besonders verbreitet ist die Varimaxrotation, bei der die Varianz faktoren- (spalten-)weise maximiert werden soll.

Wollen wir mit der Faktorenextraktion eine Datenreduktion erreichen, d.h. gegenüber der Matrix der Korrelationskoeffizienten zwischen allen Variablen zu einer besseren Übersicht über die Zusammenhänge zwischen den Merkmalen kommen, so sollte die Anzahl der Faktorladungen unter der der Korrelationskoeffizienten zwischen den Merkmalen

$$m*(m-1)/2$$

(bei m Merkmalen; "/2" wegen der Symmetrie der Matrix) liegen.

Nicht allein die Anzahl der Faktorladungen entscheidet über die Qualität eines gefundenen Modells, sondern auch ihre Werte spielen eine Rolle dafür. Für deren Einschätzung hat THURSTONE 1947 die fünf Kriterien der Einfachstruktur des Faktorenmusters (zitiert nach ÜBERLA 1968,184) erarbeitet:

Arbeitsgemeinschaften

1. Jede Zeile des Faktorenmusters sollte wenigstens eine Nullladung haben (d.h. jede Variable sollte in wenigstens einer Koordinatenhyperebene liegen).
2. Jede Spalte sollte wenigstens w Nullladungen haben (d.h. jeder Faktor sollte durch w Variablen in seiner Koordinatenhyperebene definiert und lokalisiert sein).
3. Für jedes Spaltenpaar des Faktorenmusters sollte es mehrere Variablen geben, die auf einem Faktor keine, aber auf den anderen hohe Ladungen haben.
4. Bei mehr als vier Faktoren sollten in jedem Spaltenpaar möglichst viele Variablen in beiden Spalten Nullladungen haben.
5. Für jedes Faktoren- und Spaltenpaar des Faktorenmusters sollten nur wenige Variablen in beiden Spalten hohe Ladungen haben.

Diese Kriterien sollten, WITTMANN (1977,191 - zitiert nach SCHILLER 1988,76) zufolge, dafür herangezogen werden "...das Inflationäre Angebot an Rotationsverfahren" zu bewerten, um eine «beste» Ladungskonfiguration im Faktorenraum zu finden.

Für die Bewertung ergibt sich jedoch die Schwierigkeit einer Abgrenzung von «Nullladung» (substantiell niedriger) und «interpretierbarer» (substantiell hoher) Ladung, um eine gefundene Lösung im Hinblick auf THURSTONEs Kriterien auszählen zu können. BARGMANN (1955, zitiert nach SCHILLER 1978,303) hat als Obergrenze ersterer einen Anteil von 10% aus der Quadratwurzel der Kommunalität des jeweiligen Merkmals (also der Summe der Faktorladungsquadrate pro Zeile) angegeben. Abgesehen davon, daß der Bezug zur Kommunalität problematisch erscheint (bei niedriger Kommunalität wird der Maximalwert der Nullladung sehr herunterdrückt), ist diese Grenze insofern zu starr, als sie die Anzahl der einer Faktoranalyse zugrundeliegenden Beobachtungen nicht berücksichtigt. PAWLIK hat deshalb (1968; 1971,182-183) vorgeschlagen, "...in nicht zu kleinen Personenstichproben" Faktorenladungen zwischen $-0,1$ und $+0,1$ als Nullladungen zu klassifizieren und damit die «Hyperebenenstärke», ("hyperplane width") auf $0,2$ festzulegen. Dieser Wert solle bei Stichprobenumfängen von $50...60$ oder weniger auf $0,25...0,30$ steigen und bei (mehr als) 500 «Versuchspersonen» auf $0,1$ abnehmen. Gesicherte interpretierbare Ladungen will PAWLIK ab Beträgen von $0,3$, bei kleinen Stichproben von $0,4$ gelten lassen.

Auch diese Angaben erscheinen subjektiv, ermangeln sie doch einer überzeugenden Begründung. Plausibel wäre m.E. dagegen die Festlegung einer Obergrenze für die Nullladung auf den Zufallshöchstwert des Korrelationskoeffizienten, bezeichnet doch die Ladung eben die Korrelation zwischen Variabler und Faktor. Unter Zugrundelegung einer gleichbleibenden Irrtumswahrscheinlichkeit ergäben sich hohe Werte für kleine Stichprobenumfänge, die sich mit wachsender Anzahl der Beobachtungen - zunehmend langsamer - verringerten. Ähnlich haben schon BUSE & PAWLIK (1978,10) argumentiert.

Schwieriger zu beantworten ist die Frage nach den Untergrenzen substantiell hoher Ladungen. Faktoren als hypothetische Größen bedürfen einer engeren Verknüpfung mit den unmittelbar beobachteten Daten als es der Nachweis einer «nicht mehr zufälligen» Korrelation wäre. Letzten Endes kann die Frage nur (sachlich begründet) subjektiv entschieden werden. Grundsätzlich wären

für höhere Stichprobenumfänge niedrigere Untergrenzen anzusetzen. Die Untergrenzen könnten durch Konfidenzintervalle (zu den angesetzten Werten zu addieren) abgesichert werden, damit man «auch wirklich» von den betreffenden Beträgen der Korrelationskoeffizienten ausgehen kann.

Welche praktischen Schlußfolgerungen ergeben sich bei einem solchen Vorgehen für den Umgang mit Korrelationsmatrizen, die man faktorisieren will? Aussagemöglichkeiten und -grenzen seien an einem Beispiel erläutert.

An 3178 Feuersteinabschlägen der altpaläolithischen Homo-erectus-Fundstelle Bilzingsleben (weniger als 10% des Gesamtbestandes bisher geborgener Flintartefakte: WEBER 1986) wurden zahlreiche metrische Merkmale erfaßt, von denen zwölf in eine Faktorenanalyse eingegangen sind Länge, Breite und Dicke des Stückes, Weite und Tiefe des Schlagflächenrestes, Schlagwinkel, Anzahl dorsaler Negative, Anteil dorsaler Bearbeitung, Anzahl dorsaler Schlagrichtungen, Längen-Breiten-Verhältnis, relative Dicke (Quotient aus Dicke und Mittel von Länge und Breite *100) und Weiten-Tiefen-Verhältnis des Schlagflächenrestes. Die Abbildung zeigt einen «Ideal-Abschlag» zur Erklärung der Meßverfahren.

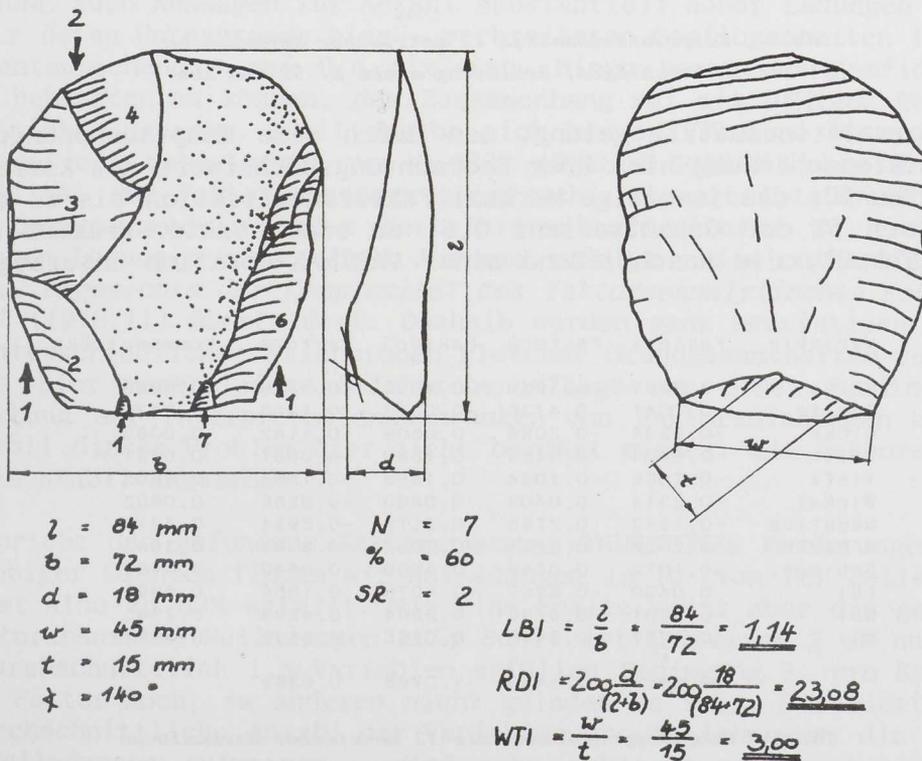


Abb. Feuersteinabschlag mit schematischer Darstellung der in die Faktorenanalyse einbezogenen Merkmale Länge (l), Breite (b), Dicke (d) des Stückes, Weite (w) und Tiefe (t) des Schlagflächenrestes, Schlagwinkel (x), Anzahl dorsaler Negative (N) Anteil dorsaler Bearbeitung (%), Anzahl dorsaler Schlagrichtungen (SR), Längen-Breiten- (LBI), relativer Dicken- (RDI) und Weiten-Tiefen-Index (WTI).

Arbeitsgemeinschaften

Die arithmetischen Mittel der Einzelmerkmale ließen erkennen, daß es sich bei dem Inventar um einen Fund aus dem «mikrolithischen Altpaläolithikum» handelt, den in Vergleich zu anderen Inventaren zusetzen durch die Anzahl der beobachteten Variablen erschwert wird. Faktorenanalytisch sollten die Zusammenhänge zwischen den Merkmalen studiert werden, zumal die Korrelationsmatrix (Tab.1.) bekanntlich auch multiple Zusammenhänge widerspiegelt, die z.B. durch die Analyse der partiellen Korrelationen eliminiert werden könnten. (Es wurde der PEARSON-BRAVALSsche Maßkorrelationskoeffizient verwendet, auch wenn nicht in allen Fällen die Voraussetzung normalverteilter Beobachtungswerte erfüllt war.

	2	3	4	5	6	7	8	9	10	11	12	13	14
2 L	1,0000												
3 b	0,8215*	1,0000											
4 d	0,6500*	0,6289*	1,0000										
5 w	0,3065*	0,6180*	0,4568*	1,0000									
6 t	0,3857*	0,4199*	0,6147*	0,6310*	1,0000								
7 k	0,1847*	0,2057*	0,1998*	0,1327*	0,1450*	1,0000							
8 N	0,0913*	0,1698*	0,0884*	0,1343*	0,0836*	0,1002*	1,0000						
9 %	-0,1617*	-0,1341*	-0,2362*	-0,0702*	-0,1313*	-0,0501	0,3533*	1,0000					
10 SR	0,1170*	0,0890*	0,1175*	-0,0077*	0,0274	0,1150*	0,3378*	0,2946*	1,0000				
12 LBI	0,4081*	0,3840*	0,0283	-0,2054*	-0,0291	-0,0268	-0,0774*	-0,029*	0,0326	1,0000			
13 RDI	0,0731*	0,0196*	0,6054*	0,1434*	0,4463*	0,0583*	-0,0151	-0,1848*	0,0529*	0,0767*	1,0000		
14 WTI	-0,1162*	0,1312*	-0,1957*	0,3156*	-0,3597*	-0,0456*	0,0820*	0,0843*	0,0404*	-0,2745*	-0,3163*	1,0000	

Zufallshöchstwert für $r = 0,034$

Tab.1. Korrelationsmatrix 12 metrischer Merkmale an 3178 Flintabschlägen. Erklärung siehe S. 99 und Abb.

Von der Korrelationsmatrix gelangt man durch eine Hauptkomponentenanalyse (Kommunalitätenschätzung hier über Beobachtungshöchstwert des Korrelationskoeffizienten für das jeweilige Merkmal Faktorenextraktion bis zu einem Minimalwert von 5% der Gesamtvarianz 0,6) zu einer Faktorstruktur mit vier Faktoren (Tab.2), die anschließend einer VARIMAX-Rotation unterzogen wurde (Tab.3).

Variable	Faktor1	Faktor2	Faktor3	Faktor4	Kommunalität
Länge	-0,7281	-0,1551	-0,4995	0,4429	0,9999
Breite	-0,7641	0,4736	-0,0373	0,1905	0,8470
Dicke	-0,9204	-0,2089	0,0508	-0,1142	0,9065
Weite	-0,6404	0,4453	0,1860	-0,0037	0,6431
Tiefe	-0,7088	-0,1064	0,1489	-0,1568	0,5605
Winkel	-0,2314	0,0403	-0,0860	-0,0262	0,0633
Negative	-0,1343	0,2765	-0,4172	-0,3914	0,4217
% Bearb.	0,2128	0,2291	-0,3907	-0,3793	0,3943
Schlagr.	-0,1072	0,0743	-0,4200	-0,3540	0,3187
LBI	0,0460	-0,6187	-0,4075	0,1956	0,5893
RDI	-0,4801	-0,4798	0,2984	-0,4298	0,7345
WTI	0,1297	0,5440	0,0265	0,1413	0,3334
	3,2527	1,5466	1,0765	0,9362	

Tab.2. Hauptkomponentenmatrix 12 metrischer Merkmale an 3178 Flintabschlägen. Erklärung siehe S. 99 und Abb.

Wollen wir die Qualität des gefundenen Musters bezüglich der THURSTONEschen Kriterien beurteilen, so müssen wir für die Festlegung des Nulladungshöchstwertes oben empfohlenem Vorgehen entsprechend auf den Stichprobenumfang zurückgreifen: Da diese Zahl sehr groß ist, reicht bereits ein ganz niedriger absoluter Betrag des Korrelationskoeffizienten für die Annahme eines nicht mehr zufälligen Zusammenhanges zwischen den beobachteten Variablen. Der Zufallshöchstwert liegt (bei einseitiger Fragestellung) unter Annahme einer Irrtumswahrscheinlichkeit von 596 bei 0,0348 (SACHS

Variable	Faktor1	Faktor2	Faktor3	Faktor4	Kommunalität
Länge	-0,8460	0,0173	0,0025	0,5328	0,9999
Breite	-0,8730	0,0476	-0,0192	-0,2868	0,8470
Dicke	-0,7184	-0,6221	0,0166	0,0554	0,9065
Weite	-0,6490	-0,1050	0,0090	-0,4590	0,6431
Tiefe	-0,5240	-0,5272	0,0268	-0,0853	0,5605
Winkel	-0,2264	-0,0669	-0,0865	0,0065	0,0633
Negative	-0,1588	0,0160	-0,6242	-0,0810	0,4217
% Bearb.	0,1645	0,1442	-0,5867	-0,0473	0,3943
Schlagr.	-0,0930	-0,0642	-0,5471	0,0812	0,3187
LBI	0,0614	-0,0730	0,0179	0,7615	0,5893
RDI	-0,1065	-0,8491	0,0423	0,0169	0,7345
WTI	-0,0670	0,4425	0,0312	-0,3634	0,3334
	2,8215	1,6303	1,0452	1,3149	

Tab.3. VARIMAX-rotierte Faktorenmatrix 12 metrischer Merkmale an 3178 Flintabschlägen. Erklärung siehe S. 99 und Abb.

1984,330). Nullladungen im Faktorenmuster müssen also diesen Grenzwert «unterbieten».

Da THURSTONE auch Aussagen zur Anzahl substantiell hoher Ladungen erwartet, wollen wir deren Untergrenze hier - verbreiteten Gepflogenheiten in der Literatur entsprechend - auf 0,6 fixieren. Hinzu kommt das Konfidenzintervall, um behaupten zu können, der Zusammenhang sei mit 95%iger Wahrscheinlichkeit nicht lockerer. Damit ergibt sich bei $n=3178$ eine Untergrenze für die substantiell hohe Ladung von 0,622. (BUSE & PAWLIK haben das Problem der Stabilität von Faktorenmustern untersucht; es ist letzten Endes nicht von dem der zugrundeliegenden Korrelationskoeffizienten zu trennen. Mit analytischen Lösungen wäre -ihrer Aussage zufolge - die mathematische Statistik "...angesichts der Komplexität des faktorenanalytischen Rechenganges zur Zeit" (1978,11) überfordert. Deshalb wurden ganz heuristisch Faktorenmuster unterschiedlicher Stichproben gleicher Grundgesamtheiten verglichen. Auch wenn hier verschiedene Stichprobenumfänge aus einem Fundinventar in Ihrer Wirkung auf Interpretationsschranken von Faktorenladungen betrachtet werden, soll dieses Problem hier nicht berührt werden; die Faktorstrukturen werden als stabil angesehen.

Wie entspricht das gefundene Faktorenmuster THURSTONEs Forderungen ? Unter Annahme obiger Grenzen finden wir Nullladungen in 10 (von 12) Zellen, Kriterium 1 ist also zu 83% erfüllt. Nur eine Spalte weist aber die geforderten 4 (= Faktorenanzahl) Nullladungen auf. Somit ist Kriterium 2 zu nur 25% erfüllt. Durchschnittlich 1,5 Variablen erfüllen Bedingung 3, pro Spaltenpaar in einem Faktor hoch, im anderen nicht geladen zu sein. Der vierten Bedingung- durchschnittliche Anzahl der Variablen pro Spaltenpaar, die in beiden Fällen Nullladungen aufweisen - wird sehr schlecht entsprochen; der Wert liegt bei 0,17. Dagegen tritt der «ungelebte» Fall 5 - hohe Ladungen derselben Variablen in einem Spaltenpaar - nur in einer einzigen Gegenüberstellung (von 6) auf und dort nur bei einem einzigen Merkmal.

Insgesamt können wir mit dem Resultat nicht zufrieden sein. THURSTONEs Kriterien der Einfachstruktur werden durch das vorhandene Muster nur zum Teil erfüllt. Lassen sich also über die Zusammenhänge zwischen den Beobachtungsdaten an den Bilzingslebener Abschlägen keine faktorenanalytisch begründeten Aussagen treffen ?

Arbeitsgemeinschaften

Das Material bildet, wie oben geschildert, nur eine Stichprobe aus dem auf der Grabung seit 1969 geborgenen Fundgut. Natürlich spiegelt auch das gesamte Inventar einer Ausgrabung nur einen begrenzten Ausschnitt der einstmals lebenden Kultur. Welche Aussagen würden- im Hinblick auf unsere Fragestellung - getroffen werden, hätten wir eine kleinere Stichprobe geborgen? (Wie oben erwähnt, sollen uns hier die von BUSE und PAWLIK untersuchten Veränderungen dann beobachteter Faktorenmuster nicht beschäftigen.)

Würde das Abschlaginventar nur jedes zehnte untersuchte Artefakt umfassen (also $n = 318$), läge der Nullladungshöchstwert bei 0,11, das Minimum substantieller Ladung bei 0,667 (unter Anwendung des oben geschilderten Algorithmus). In diesem Falle würde eine Faktorstruktur wie die hier vorgestellte die Bedingungen 1 und 2 100%ig erfüllen, durchschnittlich zwei Merkmale wären pro Spaltenpaar auf einem Faktor hoch, auf dem anderen nicht geladen. Der Wert für Kriterium 4 läge bei 3,17 und Kombinationen hoher Ladungen derselben Variablen gäbe es in keinem Faktorenpaar.

Wenn wir die Stichprobe noch einmal auf 10% ausdünnen und damit die Untergrenze sinnvoll behandelbarer Beobachtungsreihen erreichen ($n = 32$), so reichte die Zone der Nullladungen bis 0,3382, substantielle Ladungen sollten einen Betrag von mindestens 0,792 erreichen. Dann wären Kriterien 1 und 2 wie beim vorangegangenen Versuch erfüllt, aber (wegen des hohen Schwellenwertes für eine hohe Ladung) 3 mit nur 1,33. Durchschnittlich 4,83 Variable wären pro Spaltenpaar in beiden Fällen null-geladen und ein gemeinsames Auftreten hoher Ladungen derselben Variable gäbe es nicht.

Wie auf Grund der Prämissen unschwer vorstellbar, verbessert eine Verringerung des Stichprobenumfangs die Aussagemöglichkeiten, weil die Forderungen, die an eine Stichprobe gestellt werden, weniger scharf formuliert sind. (Andererseits ist mit einer wachsenden Instabilität des Faktorenmusters beim Vergleich mehrerer solcher kleineren Stichproben aus derselben Grundgesamtheit zu rechnen, ein Umstand, der eine analytische Lösung des Problems wünschenswert erscheinen läßt.) Sollte dies nun ein Motiv für den Archäologen sein, bei Aussagen über seine Funde auf kleinere Stichproben zurückzugreifen?

Ein Patentrezept kann nicht gegeben werden. Selbstverständlich erhöhen umfangreichere Beobachtungsreihen die Aussagegenauigkeit, die Stichprobenwerte werden von den «wahren» Werten der Grundgesamtheit immer weniger differieren. Immer feinere Unterschiede zwischen Stichproben können damit aber auch als signifikant herausgearbeitet werden, was zweifellos in den objektiven Abweichungen der Grundgesamtheiten voneinander wurzelt. Bei einer Perspektive von der Gegenwart auf das ältere Paläolithikum allerdings, das aus riesigen geographischen Räumen für Hunderttausende von Jahren durch vergleichsweise wenige Inventare belegt ist, sollten Unterschiede in den «Handschriften» einzelner Steinartefaktproduzenten zurücktreten hinter gravierenden Abweichungen zwischen den großen Technokomplexen. Univariate Vergleiche zwischen den Beobachtungsreihen an Funden dieser Periode legen es genauso nahe wie die Resultate der Faktorenmuster-Bewertungen, daß für eine solche Perspektive Stichprobenumfänge von einigen hundert Objekten ein ausreichend scharfes Bild der technologischen Entwicklung liefern.

L i t e r a t u r

- BARGMANN, R. (1955) Signifikanzuntersuchungen der einfachen Struktur in der Faktorenanalyse. Mitteilungsblatt Mathematische Statistik 7, 1955, 1-24.
- BUSE, L. & K. PAWLIK (1978) Stichprobenfehler und Rotationsambiguität von Faktorladungen. Zeitschrift für experimentelle und angewandte Psychologie 25/1, 1978, 10-34.
- PAWLIK, K. (1968) Dimensionen des Verhaltens. Bern/Stuttgart 1968. (2. Aufl. 1971).
- SACHS, L. (1984) Angewandte Statistik. 6. Aufl. Berlin/Heidelberg/New York/Tokio 1984.
- SCHILLER, W. (1978) Die Signimaxrotation - ein weiterer Versuch, Faktorenstrukturen interpretationsgünstig zu rotieren. Biometrical Journal 20, 1978, 299-316.
- SCHILLER, W. (1988) Vom sinnvollen Aufwand in der Faktorenanalyse. Archiv für Psychologie 140, 1988, 73-95.
- THURSTONE, L.L. (1947) Multiple factor analysis. Chicago 1947.
- ÜBERLA, K. (1968) Faktorenanalyse. Berlin/Heidelberg/New York 1968.
- WEBER, E. (1974) Einführung in die Faktorenanalyse. Jena 1974.
- WEBER, Th. (1986) Die Steinartefakte des Homo erectus von Bilzingsleben. In: MANIA, D. & Th. WEBER, Bilzingsleben III. Homo erectus - seine Kultur und Umwelt. Berlin 1986, 65-230.
- WITTMANN, W. (1977) Faktorenanalytische Modelle, Methodenstudium und Probleme der Reproduzierbarkeit. Diss. Albert-Ludwigs-Universität, Freiburg i.Br. 1977.

Dr. phil. habil. Thomas Weber
Landesamt für archäologische Denkmalpflege Sachsen-Anhalt
Richard-Wagner-Str. 9/10
O-4020 Halle/Saale