Die Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten.

Von

F. Krueger und C. Spearman.

Inhaltsverzeichnis.

Da im Laufe dieser Arbeit öfters auf zwei frühere Arbeiten im *American Journal of Psychology* hingewiesen wird, so soll diese Zeitschrift kurz als *Am. J. Psych.* bezeichnet werden.

I. Vorbedingungen der Untersuchung.

1. Das allgemeine Problem.

In der folgenden experimentellen Untersuchung wird versucht die quantitativen Zusammenhänge zwischen verschiedenen geistigen Leistungsfähigkeiten festzustellen und näher zu betrachten.

Die populäre Meinung geht dahin, dafs es solche Zusammenhänge von sehr allgemeiner Natur gibt. Wenn ein Individuum z. B. als „intelligent" bezeichnet wird, so bedeutet das in der Regel mehr als einen Hinweis blofs auf seine tatsächlich bisher manifestierten Leistungsfähigkeiten. Man drückt in diesem Attribut auch noch die Erwartung aus, dafs das Individuum bei erheblich anderen als den bisher geprüften Leistungen sich ebenfalls mehr oder weniger auszeichnen werde. Die wissenschaftliche Psychologie dagegen steht solchen allgemeinen Zusammenhängen äufserst zurückhaltend und vielfach sogar entschieden ablehnend gegenüber; für nicht wenige Psychologen ist die „Intelligenz" nur ein Name für das zufällige Beieinandersein mehrerer günstiger Dispositionen.

Vielleicht noch fester ist die populäre Überzeugung, dafs Zusammenhänge von speziellem beschränktem Charakter existieren. Von den Kindern z. B., die sich in der Arithmetik auszeichnen, erwartet man, dafs sie auch in bezug auf Algebra nicht zurückstehen werden; von denen, die sich als gute Sänger zeigen, meint man, dafs sie auch das Violinspielen zu erlernen relativ gute Aussicht haben. Solche, wohl nicht ganz aus der Luft gegriffene Ansichten wurden im Anfange des letzten Jahrhunderts auf die Spitze getrieben, indem die Phrenologie ihre bekannten 14 spezifischen intellektuellen „Vermögen" aufstellte. Bald ist dieses Vermögenssystem unter kritischen Einwänden zusammengestürzt.

Aber merkwürdigerweise scheint die wissenschaftliche Psychologie bei diesem blofs negativen Erfolge stehen geblieben zu sein. Es sind zwar in den letzten 20 Jahren ziemlich zahlreiche experimentelle Untersuchungen angestellt worden, um Zusammenhänge zwischen geistigen Leistungsfähigkeiten messend festzustellen. Aber der Erfolg ist bisher äufserst dürftig gewesen; die Ergebnisse sind einander widersprechend; und es hat sich, besonders in den letzten Zeiten, eine starke Neigung entwickelt, die fraglichen Zusammenhänge überhaupt zu bestreiten.[1]

Dieses unbefriedigende Resultat, sowie die aufserordentliche theoretische und praktische Wichtigkeit des Problems, haben uns zu dem Versuche geführt, solche Zusammenhänge auf eine neue, wie wir meinen, gründlichere Weise zu untersuchen. Leider haben sich zu diesem Zwecke eine etwas gedrängte Behandlung und ein zwar prinzipiell einfaches, aber doch zeitraubendes rechnerisches Verfahren nicht umgehen lassen, die die Geduld des Lesers manchmal auf die Probe stellen müssen.

2. Die Notwendigkeit, den Grad einer Abhängigkeit quantitativ zu bestimmen.

Bei dem Versuche, die Abhängigkeitsbeziehungen zwischen verschiedenen Leistungsfähigkeiten von neuem zu untersuchen, stöfst man vor allem auf die Tatsache, dafs diese Abhängigkeiten jedenfalls nicht absolut sind. Es kommt tatsächlich vor, dafs ein als „intelligent" bezeichnetes Individuum in einigen Hinsichten geistig nur Geringes zu leisten vermag; dafs etwa der gute Arithmetiker doch kein besonderes Talent für die Algebra aufweist. Man wird höchstens ermitteln können, dafs die eine Begabung eine gröfsere oder kleinere Tendenz hat, die andere zu begleiten. Bevor man also weiter gehen kann, mufs man erst den Grad eines partiellen Zusammenhanges berechnen können.

Zu diesem Zwecke werden wir die jetzt ziemlich allgemein bekannte Methode benutzen, die namentlich von Bravais,[2] Galton[3] und Pearson[4] ausgebildet worden ist. Man zieht zur Untersuchung eine Reihe von Fällen heran, die zahlreich genug und vor allem sorgfältig genug ausgewählt sind, um als Probe

[1] Die geschichtliche Seite dieser Frage ist in einer früheren Arbeit ausführlich behandelt worden, so dafs wir jetzt nicht mehr darauf einzugehen brauchen. (*Am. J. Psych.*, 1904, 15, S. 206.)

[2] Mémoires par divers savants, 1846, Paris, T. IX, S. 255—332.

[3] Proceedings, Royal Society of London, 1886, Bd. 40 u. 45.

[4] Phil. Trans., R. S. London, 1896, Bd. 137 A, S. 164.

der ganzen Klasse dienen zu können. Bei einigen Fällen werden sich die zwei miteinander verglichenen Merkmale mehr, bei anderen weniger proportional zeigen; durch eine einfache Rechnung erhält man schliefslich einen einzelnen sog. Korrelationskoeffizienten, der den Gesamtgrad der Proportionalität zwischen den zwei Arten von Merkmalen zum Ausdruck bringt.[1] Dieser, gewöhnlich durch das Symbol *r* dargestellte Koeffizient hat die bequeme Form, dafs er bei vollkommener Proportionalität der beiden miteinander verglichenen Wertreihen = *1*, bei vollkommen umgekehrter Proportionalität = —*1*, und bei völliger Unabhängigkeit = *0* ist. Statt reeller Werte kann man ebenso gut zwei Rangordnungen miteinander vergleichen; dann stellt *r* den Grad ihrer Tendenz zur Übereinstimmung dar. Durch diese Methode werden die Abhängigkeitsgrade zwischen den verschiedensten Merkmalen quantitativ miteinander vergleichbar.[2]

Es ist aber offenbar möglich, dafs die Abhängigkeit zwischen zwei Merkmalen nicht die einfache Proportionalität, sondern eine

[1] Ein Beispiel der Rechnungsweise wird auf S. 113—114 gegeben.

[2] Vor kurzem hat Herr Prof. Dürr an den einen von uns die Frage gerichtet, welche Bedeutung diesem Werte *r* eigentlich zukomme. Seinerseits konstatiert er, dafs er unter einer quantitativ mefsbaren Korrelation zwischen zwei Merkmalen „nur dann etwas würde denken können", wenn er darunter die Prozentzahl der Fälle verstehen dürfe, in denen der höhere Grad des einen Merkmales mit dem höheren Grade des anderen zusammentrifft (*diese Zeitschrift* 42, S. 470—472). Nun aber dürfte diese Beschränkung aller Korrelationsmessung auf einen einzigen Gesichtspunkt schwerlich haltbar sein. Vielmehr läfst sich eine grofse Anzahl mehr oder weniger gleichberechtigter Korrelationskoeffizienten aufstellen, und für jeden solchen Koeffizienten wiederum sehr verschiedene Gesichtspunkte wählen. Wollen wir hier einen besonders einfachen Gesichtspunkt für den Koeffizienten *r* kurz erläutern. Betrachten wir zuerst diejenigen Korrelationen, wo es sich um die Übereinstimmung zwischen zwei Rangordnungen handelt (denn dies ist der Fall in der gegenwärtigen sowie in der vom Herrn Prof. kritisierten Abhandlung). Eine solche Übereinstimmung ist — nach dem anerkannten Wortgebrauch — desto höher anzusetzen, je kleiner die Diskrepanzen zwischen den zugeordneten Werten sind. Aber es ist mathematisch evident, dafs je kleiner die Summe der Quadrate dieser Diskrepanzen (im Verhältnis zu der durch blofsen Zufall zu erwartenden Summe) ausfällt, desto gröfser der Wert *r* werden mufs. Und ganz Analoges gilt, wenn statt Rangordnungen reelle Werte verglichen, und dann statt des Grades der Übereinstimmung derjenige der Proportionalität gemessen werden sollen. Der Wert *r* weist also ohne weiteres eine Bedeutung auf, die uns völlig berechtigt, ihn als ein quantitatives Mafs der Korrelation anzusehen.

spezielle kompliziertere Form darbiete. Diese spezielle Form wird dann im Koeffizienten r (allein genommen) gar nicht ausgedrückt.[1] Trotzdem bleibt aber r fast immer ein sehr annäherndes Maſs des Abhängigkeits**grades**: erst wenn die Form in ganz extremer (und dann ohne weiteres auffallender) Weise von der einfachen Proportionalität abweicht, wird die Bestimmung selbst des Grades erheblich fehlerhaft sein.

Glücklicherweise können für die gegenwärtige Abhandlung alle solche Schwierigkeiten unbeachtet bleiben. **Denn sie hat in bezug auf die Bedeutung von r nichts weiter als folgende drei Voraussetzungen nötig: Derjenige Abhängigkeitsgrad ist als höher anzusehen, wo r erheblich gröſser ausfällt. Wenn r nahezu $= 1$, so ist die Abhängigkeit beinahe vollkommen. Wenn dagegen r wenig über 0 liegt, so ist fast keine Abhängigkeit vorhanden.** Diese drei Sätze dürften wohl hinsichtlich der tatsächlich in Betracht kommenden Korrelationen (d. i. zwischen Rangordnungen nach geistigen Leistungsfähigkeiten) kaum von jemand ernstlich bestritten werden.

In allen Fällen muſs neben r auch sein sog. wahrscheinlicher Fehler berechnet werden. Alle Ergebnisse (bzw. Ergebnisunterschiede), die nicht wenigstens zweimal gröſser als der wahrscheinliche Fehler (bzw. wahrscheinliche Unterschied) sind, dürfen nicht wissenschaftlich verwertet werden, da sie fast ebenso leicht vom bloſsen Zufall herrühren könnten. Um den Zufall ganz sicher auszuschlieſsen, sollte das betreffende Ergebnis den wahrscheinlichen Fehler um mindestens das Fünffache übersteigen. Diese Maſsregeln in bezug auf den wahrscheinlichen Fehler sind von besonderer Wichtigkeit bei Versuchen, wo die verglichenen Reihen kurz und die Rolle des bloſsen Zufalls dementsprechend groſs ist; dies ist der Fall bei den Versuchen, worüber hier berichtet wird. (Die Formel für die wahrscheinlichen Fehler befindet sich auf S. 114.)

3. Methode zur Kompensation der zufälligen Fehler. Die „Ergänzungsformel".

Die Feststellung irgend einer Korrelation ist einer besonderen Gefahr dadurch ausgesetzt, daſs sie sich nicht unmittelbar auf

[1] Es ist fraglich, ob eine ganz befriedigende Bestimmung solcher spezieller Formen bis jetzt überhaupt erzielt worden ist.

die zu vergleichenden Tatsachen selbst bezieht, sondern zunächst nur auf die aus diesen gewonnenen Messungswerte. Jede Art Messung, mag sie noch so sorgfältig ausgeführt werden, bleibt immerhin mit den sog. zufälligen Fehlern behaftet; wenn es sich um psychologische Messungen, und gar um die Messung einer geistigen Leistungsfähigkeit handelt, so können solche Fehler eine recht beträchtliche Gröfse erreichen.

Nehmen wir beispielsweise an, dafs man die Länge und die Breite einer gewissen Anzahl von Gegenständen gemessen hat, und ferner, dafs jede Messung ebenso leicht und um gleiche Beträge zu grofs als zu klein ausfallen kann. Unter diesen Voraussetzungen ist es klar, dafs — wie grofs auch die einzelnen Messungsfehler sein mögen — der Durchschnittswert jeder der beiden Messungsreihen ebenso gut gröfser als kleiner wie der Durchschnitt der wirklichen Gröfsen sein kann; und wenn die Gegenstände nur zahlreich genug sind, so wird die Diskrepanz der gemessenen von der wirklichen Durchschnittslänge (bzw. Breite) verschwindend klein sein.

Ganz anders liegt nun die Sache, wenn man die Korrelation zwischen der Länge und der Breite dieser Gegenstände berechnen will. Denn je mehr die Messungswerte von zufälligen Fehlern beherrscht werden, desto mehr wird ein etwa vorhandener Zusammenhang zwischen Länge und Breite verdeckt; die scheinbare Korrelation wird also (natürlich innerhalb der durch den wahrscheinlichen Fehler bestimmten zufälligen Schwankungsbreite) nie zu grofs, sondern immer im genannten Mafse zu klein ausfallen. Und einer solchen Störung kann keine Ausdehnung der Versuchsreihen, oder selbst Wiederholung der ganzen Untersuchung im geringsten abhelfen. Diese illusorische Verkleinerung der Korrelation durch die Messungsfehler scheint auch in der Tat eine Hauptursache des Widerstreits der bisherigen Ergebnisse zu sein; in den Versuchen, wo die Messungsfehler sehr grofs gewesen sind, ist eine, wenn auch wirklich vorhandene Korrelation nicht hervorgetreten; infolgedessen ist sie irrtümlicherweise verneint worden. Die Feststellung einer kleinen Korrelation ist also zunächst zweideutig; sie kann auf einen wirklichen Mangel an Zusammenhang hindeuten, oder aber nur auf grofse zufällige Messungsfehler.

Um nun diese Zweideutigkeit zu überwinden, müssen alle Messungen mindestens zweimal durchgeführt werden. Nehmen wir der Anschaulichkeit halber an, daſs dies in unserem obigen Beispiel (Korrelation zwischen der Länge und der Breite von Gegenständen) geschehen sei; es könnte sich daraus etwa folgende Tabelle ergeben:

Tabelle I.

Gegenstände	Länge			Breite		
	Reihe der wirklichen objektiven Gröſsen von L	Messungsreihe L_1	Messungsreihe L_2	Reihe der wirklichen objektiven Gröſsen von B	Messungsreihe B_1	Messungsreihe B_2
1	x_1	114	110	y_1	55	58
2	x_2	109	107	y_2	56	55
3	x_3	119	125	y_3	60	59
4	x_4	103	106	y_4	52	53
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Man erlangt also die zwei Messungsreihen L_1 und L_2 (bzw. B_1 und B_2) als Maſs der objektiven Reihe x (bzw. y). Der Korrelationskoeffizient zwischen L_1 und L_2 muſs offenbar desto gröſser ausfallen, je kleiner die zufälligen Messungsfehler gewesen sind[1]; er dient also als ein „Zuverlässigkeitskoeffizient" der Methode zur Messung der Länge der Gegenstände. In derselben Weise ergibt die Korrelation zwischen B_1 und B_2 einen „Zuverlässigkeitskoeffizienten" für die Methode der Breitenmessungen.

Diese „Zuverlässigkeitskoeffizienten" lassen sich auf genaue rechnerische Weise verwerten, mit dem Erfolge, daſs irgend eine „rohe", d. h. durch die zufälligen Fehler herabgesetzte Korrelation wieder zu ihrem vollen Werte ergänzt werden kann. Um dies zu erreichen, benutzt man folgende „Ergänzungsformel", wie wir sie bezeichnen wollen:

$$\mathbf{AB} = \frac{M\,(A_1B_1,\ A_1B_2,\ A_2B_1,\ A_2B_2)}{M\,(A_1A_2,\ B_1B_2)} \tag{a}$$

[1] Immer unter Abstraktion von Schwankungen innerhalb der durch den wahrscheinlichen Fehler bestimmten Grenzen.

wo A_1 und A_2 (bzw. B_1 und B_2) = zwei gleich genauen voneinander unabhängigen Messungsreihen für **A** (bzw. **B**); A_1B_1 (bzw. A_1B_2, A_1A_2 usf.) = der rohen Korrelation zwischen A_1 und B_1 (bzw. A_1 und B_2, A_1 und A_2 usf.); wo M den Mittelwert bedeutet[1], und wo **AB** = der gesuchten vollständig ergänzten, also reinen Korrelation zwischen **A** und **B** ist.[2]

Schliefslich sei bemerkt, dafs bei der Ergänzung von Korrelationskoeffizienten der Begriff „Messungsfehler" meistens in einem sehr weiten Sinne aufgefafst werden mufs; besonders dann, wenn es sich nicht (wie in unserem fingierten Beispiele) um konkrete Gegenstände handelt, sondern (wie viel gewöhnlicher) um Gegenstände von einem mehr oder weniger abstrakten Charakter. In unserer Untersuchung z. B. wollten wir die Leistungsfähigkeiten der verschiedenen Versuchspersonen unter genau vergleichbaren Umständen, also unter Abstraktion von ihrer zufälligen augenblicklichen „Disponiertheit"[3] messen; da aber die zufälligen Schwankungen der Disponiertheit sich tatsächlich nicht weg-

---

[1] Für den Zähler sollte das arithmetische Mittel benutzt werden, wenn (wie gewöhnlich) die zwei Messungsreihen A_1 und A_2 (bzw. B_1 und B_2) als gleich genau betrachtet werden können; sonst ist das geometrische Mittel theoretisch erforderlich. Für den Nenner ist nur das geometrische Mittel theoretisch richtig. Aber sowohl für den Nenner wie für den Zähler ist es doch bei sehr kurzen Reihen vorteilhafter, das arithmetische Mittel anzuwenden, da dieses den zufälligen Abweichungen weniger preisgegeben ist.

[2] Diese Ergänzungs- sowie die unten aufgestellte „Korrektionsformel" sind im *Am. J. Psych.* (1904, 15, S. 90) mitgeteilt worden; ihre ausführliche mathematische Begründung wird aber erst jetzt in derselben Zeitschrift veröffentlicht werden. Es sei gleich hier erwähnt, dafs der Beweis der Ergänzungsformel streng und allgemein ist; insbesondere beruht er auf keinerlei speziellen Voraussetzungen in bezug auf die Verteilungsgesetze der beiden verglichenen Merkmale, oder in bezug auf das Verteilungsgesetz oder die Gröfse der Messungsfehler. Erst in der Praxis wird die Formel insofern ungenau, als sie notwendigerweise nicht auf die gesamten Werte des betreffenden Kollektivgegenstandes, sondern nur auf probeweise herausgegriffene zufällige Gruppen davon angewandt wird. Durch Ausdehnung der Versuchsreihen läfst sich diese Unsicherheit beliebig herabsetzen.

(In den Fällen, wo man sich andersgestalteter oder mehrerer Koeffizienten bedienen möchte, sollte man für jeden solchen Koeffizienten die entsprechende Ergänzungsformel bestimmen. Dasselbe gilt auch für die Korrektionsformel.)

[3] d. h. der von Tag zu Tag sich ändernden psychologischen Gesamtverfassnng.

schaffen lassen, so müssen sie — für unsere Zwecke — mit zu den zufälligen „Messungsfehlern" gerechnet werden.

Auf Grund der angeführten Überlegungen ist jede Versuchsperson zunächst von Krueger in den verschiedenen zu vergleichenden Leistungen geprüft worden. Eine Woche später wurde sie zur selben Tagesstunde nach genau denselben Methoden von Spearman untersucht. Bis zum Schlusse der Versuche hat keiner von uns beiden irgend etwas über die Ergebnisse des anderen erfahren.

4. Methode zur Vermeidung und Elimination konstanter störender Faktoren. Die „Korrektionsformel".

Die im vorigen Abschnitte betrachteten Abweichungen der gemessenen von den wahren Werten waren zufälliger Art, d. h. die Abweichung für jeden Gegenstand und für jede Messung war von allen anderen Abweichungen unabhängig. Es kann aber auch vorkommen, daſs die Abweichungen alle Gegenstände in gleicher Weise treffen; es könnte etwa in unserem oben fingierten Beispiel eine ganz konstante Tendenz bestanden haben, die Länge um, sagen wir, ein Zwanzigstel zu überschätzen; in solchem Falle sind die Einzelmessungen sowie ihr Durchschnittswert gestört, aber die Korrelation bleibt unberührt.[1]

Schlieſslich kann es geschehen — und dieser Fall soll im gegenwärtigen Abschnitte behandelt werden —, daſs die Abweichungen der gemessenen von den wahren Werten zwar von einem zum anderen Gegenstand variieren, aber für verschiedene Messungen desselben Gegenstandes konstant bleiben. Dann wird die Korrelation wiederum gestört. Und solche Störung wird nicht zusammen mit allen zufälligen Störungen eliminiert; sie bedeutet vielmehr einen für sich zu betrachtenden störenden Faktor, der nur durch spezielle Maſsregeln unschädlich gemacht werden kann. Ferner kann er nicht nur verkleinernd (wie die zufälligen Fehler), sondern unter Umständen auch vergröſsernd auf die sich ergebende Korrelation einwirken.

---

[1] Solange die Abweichung der gemessenen von den wahren Werten alle Gegenstände während einer und derselben Messungsreihe in gleicher Proportion trifft, ist es einerlei, ob diese Abweichung von einer zur anderen Messungsreihe variiert. Der Korrelationskoeffizient wird allerdings etwas gestört, wenn der konstante Fehler groſs und dabei den gemessenen Werten nicht proportional ausfällt.

Als derartige störende Faktoren sind nun alle beeinflussenden Momente zu betrachten, die dem fraglichen Zusammenhange nicht streng angehören. In unseren Versuchen z. B. wollten wir den Zusammenhang zwischen verschiedenen angeborenen Leistungsfähigkeiten feststellen. Aber nehmen wir etwa den Fall an, dafs die Versuchspersonen verschiedenen Altersstufen angehörten, und ferner, dafs die beiden jeweils miteinander verglichenen Fähigkeiten sich mit zunehmendem Alter verbesserten. Dann müfsten die Werte für die zwei Fähigkeiten durch ihren gemeinsamen Zusammenhang mit der Altersstufe auch einen gewissen Zusammenhang miteinander bekommen. Aber dadurch entsteht eine scheinbare Korrelation zwischen den Fähigkeiten, selbst wenn keine wirkliche (im Sinne unserer Untersuchung) vorhanden ist; oder, wenn auch noch eine wirkliche Korrelation besteht, wird diese in illusorischer Weise vergröfsert.[1]

Nehmen wir jetzt den Fall an, dafs nur die eine der beiden Fähigkeiten von einem und demselben nicht zugehörigen Faktor abhängig sei. Dadurch wird offenbar kein künstlicher Zusammenhang mit der zweiten Fähigkeit hervorgerufen; im Gegenteil, da die Werte für die erste Fähigkeit gezwungen worden sind, sich nach dem fremden Faktor zu richten, so sind sie von einer etwa sonst bestehenden Proportionalität zur zweiten Fähigkeit mehr oder weniger abgelenkt; diesmal wird also die Korrelation in illusorischer Weise verkleinert.[2]

Es ist nun ganz und gar unmöglich, die unzähligen derartigen Störungen faktisch absolut auszuschalten. Wir können nur hoffen, sie so klein zu machen, dafs sie vernachlässigt werden dürfen. Und dazu müssen wir sie messen können. Zu diesem Zwecke läfst sich jede der hier gemeinten verschiedenartigen Störungen durch folgende einzige „Korrektionsformel" ausdrücken:

[1] Gerade eine solche Verfälschung der Ergebnisse durch den unberücksichtigten Einflufs des Alters hat eine ganze Reihe der früheren Arbeiten über den Zusammenhang zwischen geistigen Fähigkeiten wertlos gestaltet.

[2] Hierher gehört der sehr gewöhnliche Fall, dafs beide Fähigkeiten von Übung abhängig sind, und dafs die Übung in der einen Fähigkeit keinen Zusammenhang mit der in der anderen Fähigkeit hat (sollten aber in beiden Fähigkeiten dieselben Personen geübter als die anderen sein, so sind beide Fähigkeiten abhängig von demselben störenden Faktor; es tritt dann der Fall ein, der im vorhergehenden Absatze beschrieben worden ist).

$$\overline{\mathbf{AB}} = \frac{\mathbf{AB} - \mathbf{AC} \cdot \mathbf{BC}}{\sqrt{(1 - \mathbf{AC}^2)\,(1 - \mathbf{BC}^2)}}\text{[1]} \qquad \text{(b)}$$

wo **AB** = der direkt ermittelten, also „scheinbaren“ Korrelation zwischen **A** und **B**; **AC** (bzw. **BC**) = der direkt ermittelten Korrelation zwischen **A** (bzw. **B**) und irgend einem nicht zugehörigen Faktor **C**; und $\overline{\mathbf{AB}}$ = der gesuchten „wahren“ Korrelation zwischen **A** und **B** ist.[2]

Im gewöhnlichsten Falle, wo man nur den Einflufs der Abhängigkeit der einen von den beiden zu vergleichenden Merkmalen mit einem fremden Faktor eliminieren will, reduziert sich obige Formel auf:

$$\overline{\mathbf{AB}} = \frac{\mathbf{AB}}{\sqrt{1 - \mathbf{AC}^2}} \qquad \text{(c)}$$

Man wird bemerkt haben, dafs der Unterschied zwischen den störenden Faktoren dieses Kapitels und den zufälligen Fehlern des vorigen eigentlich mehr methodisch als faktisch ist. In der gegenwärtigen Untersuchung z. B., haben wir die Schwankungen der Disponiertheit dadurch als blofs zufällige Fehler betrachten können, dafs wir die zweite Prüfung jeder Versuchsperson erst eine Woche später vornahmen; die Versuchsperson bekam also Zeit genug, um in eine ganz neue Disponiertheit zu ge-

[1] Die Begründung auch dieser Korrektionsformel ist streng und allgemein (s. Anm. 2, S. 57), mit Ausnahme einer unvermeidlichen Beschränkung. Diese läfst sich am obigen Beispiel der verschiedenen Altersstufen klar machen. Die betreffenden Störungen der Korrelationskoeffizienten rühren nur von der Vermischung der Altersstufen her; sie sind offenbar beseitigt, sobald man ausschliefslich Vpn. vom gleichen Alter heranzieht. Dabei ist es gewöhnlich nicht von Belang, welche besondere Altersstufe gewählt wird; denn die „wahre“ Korrelation bleibt, selbst wenn sie von der scheinbaren ganz verschieden ist, meistens für alle Altersstufen innerhalb weiter Grenzen annähernd konstant. Aber diese Grenzen können immerhin schliefslich überschritten werden, sodafs die wahre Korrelation bei der einen extremen Altersstufe von derjenigen bei der anderen extremen Stufe erheblich abweicht. Wenn man nun solche disparate Stufen in eine und dieselbe Untersuchung zusammenfassen und auch hier die „wahre“ Korrelation verlangen will, so kann diese natürlich nur die Bedeutung eines Mittelwertes haben. Und auf diese Bedeutung eines Mittelwertes beschränkt sich notwendig in solchen extremen Fällen auch unsere Korrektionsformel.

[2] *AB*, *AC* und *BC* müssen schon vor der Ansetzung der Korrektionsformel „ergänzt“ werden.

langen. Hätten wir dagegen die zweite Prüfung sofort nach der ersten vornehmen müssen, so wäre es unumgänglich gewesen, den Einfluſs der Disponiertheit als eines konstanten störenden Faktors gründlich für sich zu untersuchen.

Aus allen diesen Erwägungen über die störenden Faktoren ergeben sich zwei Folgerungen von entscheidender Wichtigkeit. Man muſs nämlich gerade im Gebiete der Korrelationsforschung mit viel gröſserer Schärfe die Frage stellen, als das bisher zu geschehen pflegte; d. h. man muſs von vornherein möglichst eindeutig die Tatbestände bestimmen, zwischen denen der zu untersuchende Zusammenhang bestehen soll. Aus dieser Forderung ergibt sich zweitens, daſs man niemals eine Korrelation festzustellen versuche, bis man durch eine eingehende Voruntersuchung alle nicht zugehörigen Faktoren glaubt ermittelt zu haben, die doch auf die zu vergleichenden Merkmale einen wesentlichen Einfluſs ausüben können. Ein Zusammenhang läſst sich also keinesfalls bloſs durch die mechanische Berechnung eines Korrelationskoeffizienten feststellen. Die mathematischen Hilfsmittel muſs man zwar besitzen, aber auſserdem eine gründliche Kenntnis der betreffenden Tatsachen.

Um nun diese Voruntersuchung uns leichter zu machen, haben wir für unsere Versuche vier geistige Fähigkeiten ausgewählt, die schon früher eingehend untersucht worden sind: erstens den sog. Raumsinn, worüber bekanntlich sehr zahlreiche Arbeiten vorliegen; sodann die Fähigkeiten des Addierens und des Auswendiglernens, worüber besonders die Kräpelinsche Schule die für unsere Zwecke wichtigsten Tatsachen ermittelt hat; und viertens die Unterscheidungsfähigkeit für Tonhöhen, die schon bei einer früheren Untersuchung über Korrelationen von dem einen von uns in der fraglichen Richtung geprüft wurde.[1] Dazu fügten wir schlieſslich fünftens noch die bekannte „Kombinationsmethode" von Ebbinghaus, da diese bereits beachtenswerte Ergebnisse geliefert hatte.[2]

Eine Unbequemlichkeit in der Benutzung von historisch vorliegenden Angaben über beeinflussende Faktoren entsteht

---

[1] *Am. J. Psych.*, 1904, Bd. 15, S. 226 ff.

[2] *Zeitschr. für Psychologie*, 1896, Bd. 13, S. 401; auch daselbst, 1901, Bd. 30, S. 196. Prof. Ebbinghaus war so freundlich, uns einige seiner zu jenen Zwecken hergestellten Texte samt Anweisungen über deren Gebrauch zur Verfügung zu stellen.

daraus, daſs solche Angaben in sehr verschiedenen Formen auftreten und nur selten eine Berechnung der Korrelation zwischen dem beeinflussenden und dem beeinfluſsten Merkmale nach der r-Methode gestatten. Man muſs vielmehr über eine ganze Anzahl von Rechnungsmethoden verfügen, mittels deren man dann von den verschiedensten Angaben doch wenigstens eine rohe Annäherung an r zu gewinnen imstande ist. Auf diese zahlreichen Hilfsrechenmethoden kann hier nicht eingegangen werden[1]; es sei nur bemerkt, daſs es dabei nicht auf den absoluten Einfluſs eines fremden Faktors ankommt, sondern auf das Verhältnis dieses Einflusses zur mittleren Variation der Werte des betreffenden Merkmals. Es sei ferner daran erinnert, daſs Störungen der Ergebnisse erst dann zu vernachlässigen sind, wenn sie beträchtlich kleiner sind, als der wahrscheinliche Fehler. Will man also durch sehr ausgedehnte Versuchsreiben den wahrscheinlichen Fehler verkleinern, so muſs man auch dafür sorgen, daſs die unzugehörigen Faktoren mit proportional zunehmender Strenge ferngehalten werden.

Trotz aller Bekanntheit der von uns zur Untersuchung gewählten Leistungen blieb diese Voruntersuchung immer noch der ausgedehnteste und mühseligste Teil der ganzen Arbeit. Aus der vergleichenden Betrachtung des bisher vorliegenden Materials ergab sich uns schlieſslich, daſs die Fähigkeiten des Addierens und des Auswendiglernens den kleinsten Störungen ausgesetzt seien. Hier konnten viele der am meisten erörterten beeinflussenden Faktoren — bei der von uns geplanten Ausdehnung der Versuche — unmöglich störend wirken. Die Verschiedenheit des Alters und der allgemeinen Bildung war bei unseren Versuchspersonen zu gering, um in Betracht zu kommen. Ebenso hatten wir von der Tatsache nichts zu befürchten, daſs die Versuchspersonen in etwas verschiedenem Grade dem Genusse des Alkohols, des Tabaks und des Koffeins ergeben waren. Es machte für unsere Zwecke nichts aus, daſs die einen Versuchspersonen ihre günstigere Arbeitsdisposition am Morgen, die anderen am Abend hatten. Es war nicht von Belang, ob einige Versuchspersonen zu einer früheren Stunde oder weniger reichlich als die anderen gegessen hatten. Selbst geistige und körperliche

---

[1] Dafür muſs auf *Am. J. Psych.*, **15**, S. 82—88, sowie S. 236—238 verwiesen werden.

Ermüdung, innerhalb der zu erwartenden Grenzen, bedurften keiner besonderen Rücksicht. Etwas bedenklicher waren die individuellen Unterschiede in bezug auf Übung; aber auch in dieser Hinsicht schien es zu genügen, wenn wir keine speziell begünstigten Personen zu unseren Versuchen zuließsen (z. B. solche nicht, die früher an ähnlichen Versuchen teilgenommen hatten).

In bezug auf Töne ergab sich ähnliches. Nur der Faktor der Übung erforderte erheblich mehr Berücksichtigung. Es mußsten nicht nur die Versuchspersonen ausgeschlossen werden, die sich an derartigen Versuchsreihen bereits beteiligt hatten, sondern auch andererseits die, welche musikalisch ausnahmsweise ungeübt waren.

Gröfsere Schwierigkeiten begegneten uns in bezug auf den Raumsinn. Hier schienen, nach den (allerdings bestrittenen) Angaben mehrerer Forscher, nicht nur die Übung, sondern auch die Ermüdung sowie kleine Unpäßslichkeiten einen Einfluß von bedrohlicher Gröfse zu haben. Es mußste also sorgfältig vermieden werden, daß Versuchspersonen in sehr ermüdetem oder sonst angegriffenem Zustande geprüft würden.

Solange die Versuchspersonen an unseren Beobachtungen überhaupt beteiligt waren, durften sie nicht anderweitig in verwandter Weise experimentieren. Auch waren sie gebeten, während dieser Zeit jeden Meinungsaustausch über die Versuche zu unterlassen.

Schließlich wurden alle noch irgendwie bedenklich erscheinenden Umstände während der Versuche selbst notiert, um ihren Einfluß wenigstens nachträglich mittels der Korrektionsformel eliminieren zu können. Notiert wurde jedesmal Alter, Beschäftigung mit Musik sowie mit Mathematik, allgemeiner und auch momentaner Gesundheitszustand, Dauer des letzten Schlafes, körperliche und geistige Ermüdung, Zerstreutheit (aus emotionellen oder anderen Gründen), Zwischenzeit seit der letzten Mahlzeit, Genuß von Tabak und Alkohol, Temperatur im Zimmer und im Freien, Feuchtigkeit der Luft und Höhe des Barometers. Ferner haben wir uns bemüht, die verschiedenen Grade des Eifers zu schätzen, mit welchen die Versuchspersonen sich den Beobachtungen hingaben.

II. Beschreibung unserer Versuche.

1. Allgemeines.

Die uns freundlich zur Verfügung stehenden Versuchspersonen besaſsen insofern eine gewisse Verwandtschaft, als sie sämtlich Psychologie studierten. Sie alle, mit Ausnahme eines Privatdozenten der Universität, der sich seit Jahren theoretisch eindringlich mit Psychologie beschäftigt hatte, waren Teilnehmer des experimentell-psychologischen Einführungskursus, den der eine von uns im Leipziger Institute leitete. Alle besaſsen eine gewisse Übung in psychologischen Beobachtungen. Personen von ungewöhnlicher spezieller Geübtheit wurden, wie gesagt, nicht herangezogen. Andere Vorsichtsmaſsregeln zum Ausschluſs störender Faktoren sind im vorigen Abschnitt erwähnt (S. 63).

Bei der Auswahl der Untersuchungsgebiete und Methoden lieſsen wir uns von folgenden Gesichtspunkten leiten. Zunächst, wie schon erwähnt, beschränkten wir uns auf solche Prüfungen, für die eine einigermaſsen ausgebildete Methodik bereits vorlag, und worüber ausgedehnte Versuchsreihen auch von anderen Autoren angestellt waren. Ferner kamen für unseren Zweck nur solche Versuchsanordnungen und meſsbare psychische Funktionen in Betracht, von denen wir durch ausgiebige Vorversuche (an anderen als den schlieſslich herangezogenen Personen) uns überzeugt hatten, daſs sie in verhältnismäſsig kurzer Zeit zu vergleichbaren Zahlenwerten von einiger Konstanz führen konnten. Die mit einem Beobachter an einem Tage anzustellenden Versuche durften aus naheliegenden Gründen den Zeitraum einer Stunde niemals erheblich überschreiten, und doch galt es, fünf verschiedenartige Leistungen in einer Versuchsstunde nacheinander zu prüfen. Aus diesen Gründen muſsten wir mehrere an sich mögliche und wertvolle Verfahrungsweisen, z. B. kompliziertere Gedächtnismethoden, aus dem endgültigen Versuchsplane ausscheiden.

Während der Vorversuche entschieden wir uns für folgende zeitliche Anordnung der in jeder Stunde anzustellenden Versuche:

a) Tonhöhenunterscheidung. Diese Leistungen schienen uns von der Verschiedenheit der Dispositionen, in welchen die Beobachter wegen verschiedener unmittelbar vorhergehender Einflüsse an die Beobachtungen herantraten, besonders unabhängig zu sein. Die Prüfung war, für eine Tonlage, in 15—20 Minuten mit hinreichender Genauigkeit zu vollenden. Es folgte

b) die EBBINGHAUSsche Kombinationsmethode, deren Durchführung (4 Minuten, im Anschlufs an einen orientierenden Vorversuch) gewöhnlich nicht ermüdend wirkte und eher angenehm empfunden wurde. Dann erst folgte

c) die Prüfung der taktilen Raumschwelle an drei (einschliefslich einiger Vorbereitungsversuche: vier) Hautstellen. Sie nahm durchschnittlich etwa 20 Minuten in Anspruch. Diese Versuche durften nicht zu Beginn der Stunde vorgenommen werden, sondern erst nach Adaptation der Haut an die Zimmertemperatur, die ziemlich konstant blieb (Luftheizung) und durchschnittlich 15° betrug. Hieran schlofs sich

d) das sogleich näher zu beschreibende, zeitlich genau (auf insgesamt 7 Minuten) begrenzte Addieren einstelliger Zahlen; und endlich, in einem anderen Zimmer des Institutes, das am meisten ermüdende

e) Auswendiglernen von sukzessiv dargebotenen Zahlenreihen.

Die Apparate und das Material für die Kombinationsmethode, für das Addieren und das Auswendiglernen waren jeweils vor dem Eintritt des Beobachters ganz gebrauchsfähig bereit gestellt. Für jede Stunde und jede Versuchsperson wurde ein Exemplar des als Anhang I abgedruckten, vervielfältigten Protokolls vom Versuchsleiter benutzt. Die den Versuchen jedesmal vorangehenden Fragen und ihre Beantwortung halfen uns, die Beobachter in eine gleichartige, ruhige Stimmung zu versetzen. Ein Teil dieser Fragen, der in der ersten Versuchsreihe ein für allemal beantwortet war, konnte bei der Wiederholung der Versuche, eine Woche später (durch SPEARMAN), unberücksichtigt bleiben. Nach dem Abschlufs der Versuche wurden jedesmal die unter Nr. 7 und 8 des Protokolls aufgeführten Bemerkungen notiert.

2. Besonderes zu den einzelnen Versuchsmethoden.

Alle die genannten Prüfungen hatten wir probeweise wiederholt gemeinsam durchgeführt, namentlich auch gegenseitig an uns selbst, in der endgültigen Reihenfolge, wobei wir ganz bestimmte speziellere Verfahrungsweisen ausbildeten und uns einübten, die wir dann immer möglichst gleichartig festzuhalten suchten.

a) Die Tonhöhenunterscheidung.

Als akustischer Apparat diente uns ein von dem einen von uns konstruiertes Monochord mit Millimeterskala, dessen beide Stege eine Noniuseinteilung besafsen, so dafs 0,1 mm abgelesen werden konnten. Die Spannung der Saite wurde vor jeder Versuchsreihe durch Vergleich mit einer konstanten Stimmgabel gleichgemacht. Waren die beiden benutzten Saitenstrecken (links von dem einen, rechts vom anderen Stege) durch die Schwebungsmethode auf Tongleichheit eingestellt, so waren die abgelesenen Saitenlängen nur ausnahmsweise um mehr als 0,1 mm verschieden. Da nun in dem verwerteten Tongebiete (350—370 v. d.) 1 mm Saitenlänge durchschnittlich einer Schwingung entsprach, so konnte in dieser Tonlage annähernd auf 0,1 v. d. genau abgestimmt werden. Bei der kurz bemessenen Zeit mufsten wir uns begnügen, die qualitative Schwelle in nur einer Tonlage zu bestimmen. Um jedoch die Gewöhnung an einen ganz bestimmten Ton und gewisse zufällige Fehler auszuschliefsen, wanderten wir innerhalb des angegebenen Tonbereiches für jeden Einzelversuch unregelmäfsig um einige Schwingungen.

Analoges gilt von den Raumschwellenversuchen: hier wurde, zur Vermeidung von Ermüdung u. dgl., das Ästhesiometer innerhalb des gerade untersuchten Hautgebietes jedesmal um wenige Millimeter in der Horizontalen verschoben aufgesetzt. Noch in zwei anderen Hinsichten war das Verfahren bei den Tonversuchen dem beim Raumsinne analog. In beiden Fällen wurde von deutlich überschwelligen Reizen aus absteigend vorgegangen, und zwar zuerst in gröfseren, nahe der Schwelle in annähernd gleichen kleineren Schritten (die hier im Maximum 1 Schwingung bzw. 2 mm betrugen). Zweitens hielten wir auf beiden Untersuchungsgebieten die Regel inne, dafs wir die Schwelle da als erreicht annahmen, wo unter 10 Fällen mit derselben Reizgröfse mehr als 2 falsche Urteile auftraten; bei 3 falschen Urteilen galt uns die Schwelle als bereits unterschritten, und nötigenfalls wurde das arithmetische Mittel aus zwei Reizwerten als „die“ Schwelle festgesetzt.

Das als Anhang I abgedruckte Protokoll ist als Probe für den Verlauf einer Versuchsstunde mit den tatsächlich gewonnenen Aussagen und Urteilen einer Versuchsperson (von etwa mittleren

Leistungsfähigkeiten) ausgefüllt. Als Urteilssymbole benutzten wir der Einfachheit wegen: „/" für richtige, „—" für falsche, „o" für zweifelnde Urteile.

Bei den Tonversuchen — Nr. 2 des Protokolls — wurde in unregelmäßsigem Wechsel bald der objektiv höhere bald der tiefere Ton zuerst geboten. Der Beobachter hatte jedesmal den zuzweit gehörten Ton als „höher" oder „tiefer" gegenüber dem vorigen zu beurteilen. Um Zeit zu sparen, haben wir, wie man sieht, bei den überschwelligen Reizen gewöhnlich weniger als 10 Urteile eingefordert; das geschah nur so lange, als die Urteile mit übergewöhnlicher Sicherheit und Regelmäßsigkeit abgegeben wurden. Zuweilen war es nötig um einen Schritt, aufsteigend, zurückzugehen.

b) Bei der Kombinationsmethode hielten wir uns genau an die von ihrem Urheber Ebbinghaus erprobten und ausführlich mitgeteilten Vorschriften.[1] Herr Professor Ebbinghaus hatte die grofse Freundlichkeit, uns Originale der von ihm hergestellten und vielfältig geprüften Texte zu überlassen. Wir wählten aus der für die Oberklassen geschaffenen Gruppe sechs Abschnitte, drei für die von Krueger, und drei für die von Spearman zu leitenden Versuche. Die Texte wurden auf mechanischem Wege vervielfältigt und in der gleichen Form wie die Breslauer Originale vorgelegt.

Der Versuch begann jedesmal, nach einer kurzen Vorbesprechung, damit, dafs die Versuchsperson an einem eigens hierfür bereitgehaltenen Texte die geforderten Ergänzungen unter den Augen des Versuchsleiters probeweise vollzog und damit solange fortfuhr, bis sie über ihre Aufgabe völlig im klaren war. Verlangt wurde in erster Linie sinnvolle Richtigkeit der Ergänzungen, in zweiter Linie Lückenlosigkeit, d. h. möglichste Geschlossenheit des Textes; erst in dritter Linie möglichst schnelles Arbeiten.[2]

Auf diesen Vorversuch folgte, mit Benutzung eines neuen Textes, der Hauptversuch, der nach genau 4 Minuten vom Versuchsleiter unterbrochen wurde.

Ein weiterer Text wurde jeweils in Reserve gehalten, für

---

[1] *Zeitschr. f. Psychol.*, 1897, 13, S. 402 ff.; besonders S. 414—417 u. 423 ff.
[2] Ähnlich bei Ebbinghaus a. a. O. S. 424.

den Fall unvorhergesehener Störungen des Hauptversuches — die jedoch niemals eintraten.

Bei der Auswertung des so gewonnenen Materials hielten wir uns genau an die von Ebbinghaus begründete Berechnungsweise. Es wurde also 1. die Zahl der überhaupt ausgefüllten Silben gezählt; 2. die Zahl der dabei übersprungenen, 3. der sinnwidrig ausgefüllten Silben. Jede ausgelassene Silbe wurde als ein halber Fehler gerechnet, jede sinnwidrige als ein ganzer; die so berechnete Summe der Fehler wurde von der Gesamtsumme der ausgefüllten Silben (1) abgezogen; und dieser Wert galt dann als das Quantum der Kombinationsleistung. Deren Güte dagegen wurde nach der Gesamtzahl der Fehler in Prozenten der Gesamt-Bruttoleistung (1) bemessen.

Die Kombinationsversuche waren die einzigen, zu denen nicht alle unsere Versuchspersonen zugezogen werden konnten. Die Ausländer nämlich, d. h. die (vier) Nichtdeutschen mufsten davon ausgeschlossen werden; denn wir hatten uns in Vorversuchen überzeugt, dafs die Leistungen von Ausländern, selbst wenn sie die deutsche Sprache in hohem Mafse beherrschten, wesentlich hinter dem Durchschnitt deutscher Teilnehmer zurückblieben (wie bei deutschen Texten nicht anders zu erwarten war).

Bei der Qualität unserer Versuchspersonen und bei der Einrichtung unserer Versuche kamen übrigens gewisse Schwierigkeiten nicht in Betracht, mit dem Ebbinghaus bei seinen Kollektivversuchen an Schulkindern zu rechnen hatte; aufgabewidrige Leistungen aus übermütiger Absicht, gegenseitiges Abschreiben u. dergl.

In Anhang II ist als Beispiel der Kombinationsversuche einer der Texte abgedruckt, wie ihn eine Versuchsperson von (hierfür) mittlerer Leistungsfähigkeit ausgefüllt hat. Die Fehler und Lücken der Ausfüllung sind unterstrichen.

c) Die Raumschwellen bestimmten wir, wie das Protokoll zeigt, nacheinander an 3 Hautstellen von mittlerer Empfindlichkeit. Als Vorübung ging jedesmal eine kurze, zunächst wissentliche Versuchsreihe am linken Handrücken voraus. Auf jeder Reizstufe wechselten wir unregelmäfsig und unwissentlich zwischen Doppelberührungen und solchen mit nur einer Spitze.

Das benutzte Ästhesiometer, eigens für unseren Zweck von dem einen von uns konstruiert, war besonders leicht (25 g Magnalium) handlich und nicht länger als nötig. Die genau gleichförmig abgestumpften Spitzen bestanden aus Hartgummi.[1] Eine wesentliche Neuerung unseres Ästhesiometers bestand darin, daſs es auſser den beiden parallelen Spitzen, noch eine dritte, schräg gerichtete besaſs, die uns gestattete, Einzelberührungen von gleicher Druckrichtung und Qualität als die zweispitzigen auszuüben.

Daſs wir nach dem linken Jochbein am Schlusse dieser Versuchsreihe auch das rechte Jochbein untersuchten, geschah zur Kontrolle der Übung und anderer während der Versuche selbst etwa sich ändernder Faktoren. Durch die Bestimmung dreier Schwellenwerte, aus denen schlieſslich Durchschnittswerte gewonnen werden sollten, hofften wir zufällige Schwankungen auszugleichen, und also die Prüfung des Raumsinnes jeder Person etwas zuverlässiger zu gestalten.

d) Für die Additionsversuche bestand das Material aus maschinengeschriebenen Vertikalreihen von je 70 einstelligen Zahlen. Das Verfahren war den in der Kräpelinschen Schule gebräuchlichen Additionsmethoden ähnlich. Nur daſs jede 10. Zahl von den darüberstehenden durch einen horizontalen Strich getrennt war, und die Versuchsperson immer 10 so abgeteilte Zahlen zusammenzuaddieren und das Ergebnis danebenzuschreiben hatte. Auf diese Weise schien uns der Einfluſs der bloſsen Schreibfertigkeit hinreichend herabgedrückt zu sein, während wir doch zugleich für die Qualität der Leistung (die Genauigkeit der Additionen) ein gewisses Maſs gewannen. Dieses Genauigkeitsmaſs bestand in der Prozentzahl der richtigen Summenwerte.

Auf einen Vorversuch von 1 Minute folgten 2 Hauptversuche, jeder genau 3 Minuten dauernd; wir machten 2 Hauptversuche, um auch hier eine gewisse Kontrolle zu gewinnen, und um die Zuverlässigkeit der Prüfung zu erhöhen. War die Versuchsperson beim Abschluſs eines Hauptversuchs nicht zufällig gerade am Ende einer Zehnergruppe, so hatte sie die augenblicklich gewonnene Summe neben die Zahl zu schreiben, bei der sie eben stand. Für die Berechnung der Additionsgenauigkeit wurden solche abgebrochenen Endgruppen von den anderen nicht unterschieden.

---

[1] Abnehmbare Kappen aus Celluloid sind dauerhafter und praktischer. Vgl. Katalog E (1906) des Universitätsmechanikers F. Köhler (Leipzig), S. 27, wo auch das Monochord kurz beschrieben ist.

e) Das Auswendiglernen. Die als Lernmaterial dienenden einstelligen Zahlen waren auf die Kartonscheibe des Wirthschen Gedächtnisapparates[1] derart aufgedruckt, daſs zwischen je zwei aufeinanderfolgende Zahlen einer Reihe eine Pause fiel, von der gleichen Dauer wie die der einzelnen Darbietungen. Für die Darbietungen (d. h. die Expositionen der einzelnen Zahlen) wählten wir die Zeit von 0,75 Sek., die nach Reuthers umfangreichen Beobachtungen[2] und auch nach unseren Vorversuchen durchschnittlich als die bequemste und günstigste gelten kann. Vor dem Anfang und am Ende jeder Lernreihe erschien auf der rotierenden Scheibe eine Signalfigur.

Bei der Herstellung der Zahlenmaterials benutzten wir die Erfahrungen früherer Beobachter hinsichtlich der Vermeidung assoziativ begünstigter Zahlenfolgen.

Wir arbeiteten jedesmal nacheinander mit 6-, 8-, 10- und 12gliedrigen Zahlenreihen; diese hatten sich in den Vorversuchen als von mittlerer bis ziemlich groſser Lernschwierigkeit erwiesen. Von jeder der genannten Reihenarten wurden nacheinander 3 verschiedene Reihen gelernt. Insgesamt wurden also 12 Zahlenreihen dargeboten[3]; dadurch sollte auch in diesem Gebiete eine genügende Kontrolle irgendwelcher während der Versuche etwa sich ändernder Faktoren, sowie eine genügende Zuverlässigkeit der Prüfung erzielt werden.

Jede Zahlenreihe wurde einmal dargeboten. Nach dem Schluſssignal hatte die Versuchsperson das Behaltene sofort auf ein bereitliegendes Blatt niederzuschreiben. Eine Kontrolle durch Reproduktion des Gelernten wollten wir nicht entbehren. Das Schreiben aber hatte sich in den Vorversuchen als weniger störend herausgestellt, gegenüber dem lauten Hersagen.

Bei der Reproduktion waren die Versuchspersonen angewiesen, jeder niedergeschriebenen Zahl eine bestimmte Stelle unter den vorhandenen 6 bzw. 8, 10 oder 12 Zahlen anzuweisen. Für die Auswertung zogen wir im Prinzip ausschlieſslich diese Stellenordnung in Betracht. Die Stellenordnung war hier notwendig das weitaus wichtigste Moment, da der Inhalt der Zahlen-

---

[1] Vgl. *Philosoph. Studien* 15; auch Katalog 18 des Präzisionsmechanikers Zimmermann, Leipzig.

[2] Beiträge zur Gedächtnisforschung, Diss. Leipzig, S. 43. 1905.

[3] Für den Fall von Störungen (der im ganzen dreimal eintrat) wurden von jeder Länge zwei Reservescheiben bereitgehalten.

reihen (besonders der längeren) immer ungefähr derselbe bleiben mufste; durch ausschliefsliche Berücksichtigung der Stellenordnung wurden also die Ergebnisse nur eindeutiger. Auch wurden auf diese Weise die gewöhnlichen, komplizierten und mehr oder weniger willkürlichen Berechnungsmafsregeln umgangen. Wir brauchten nur die Korrelation zwischen der wirklichen (dargebotenen) Reihenfolge der Zahlen und der von den Versuchspersonen niedergeschriebenen zu berechnen. Dazu benutzten wir eine bereits von dem einen von uns beschriebene Methode, wodurch die Berechnung unverhältnismäfsig leichter und rascher geschieht, als nacb der r-Methode.[1] Für ausgelassene sowie falsch angegebene Zahlen wurden die wahrscheinlichen zufälligen Werte in die Rechnung eingesetzt, so dafs die Versuchsperson durch Urteile aufs Geratewohl nichts zu gewinnen hatte.

III. Unsere Ergebnisse.

1. Höhe der rohen Korrelationen.

Die Rohergebnisse unserer Versuche sind in Tabelle II wiedergegeben.

Es mufsten hier zunächst die verschiedenen Werte für dieselbe Versuchsperson in demselben Leistungsgebiete auf einen einzigen Wert reduziert werden. Wenn solche zu kombinierenden Werte genügend homogen erschienen, haben wir sie einfach summiert; beispielsweise haben wir im Gebiete des Raumsinnes die Schwellen an der Hand und an den beiden Jochbeinen für je eine Versuchsperson zusammen addiert. Waren dagegen die zu kombinierenden Werte völlig heterogen, so haben wir sie dadurch miteinander vergleichbar gestaltet, dafs wir sie in Ordnungszahlen überführten. In bezug auf die Geschwindigkeit des Addierens z. B. bekam die Versuchsperson, welche die gröfste Anzahl von Ziffern addiert hatte, die Ordnungszahl 1; die Versuchsperson mit der nächst gröfsten Anzahl von Ziffern bekam die Ordnungszahl 2; usf. Ganz analog wurde die Additionsgenauigkeit behandelt. Dann nahmen wir für jede Versuchsperson die Summe ihrer zwei Ordnungszahlen.

Die so gewonnenen einzigen Werte für je eine Versuchs-

[1] Zur weiteren Ausführung und Begründung dieser leichten Methode Korrelationen zu berechnen vgl. *British Journ. of Psychol.* 2, S. 56. 1906.

Tabelle II.

Unsere rohen Ergebnisse.

Versuchsperson	Schwellen für Tonunterschiede (Schwing. i. d. Sek.)		Raumschwellen (mm)						Addieren								Auswendig lernen (Koeffizienten der Korrelationen zwischen den reproduzierten und den dargebotenen Zahlenordnungen)								Kombinationsmethode (Ebbinghaus)			
			R Hand		*R* Jochbein		*L* Jochbein		Geschwindigkeit (Anzahl der addierten Ziffern)				Genauigkeit (Prozentz. der richtigen Additionen)				6-Ziff.		8-Ziff.		10-Ziff.		12-Ziff.		Quantum		Güte	
									I		II		I		II													
	K	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*
A	2,5	3	9	11	15	12	17	12	125	139	111	125	100	80	91	100	1,00	0,67	0,45	0,58	0,92	0,63	0,70	0,47				
B	2	2,5	17	14	21	18	19	15	133	135	127	130	100	80	91	93	0,94	0,83	0,96	0,61	0,65	0,38	0,46	0,53	57	59	1,7	1,7
C	27	16	18	14	16	15	20	12	99	110	102	110	91	100	82	92	0,72	0,66	0,46	0,48	0,38	0,38	0,24	0,31				
D	1,5	2	14	14	14	15	11,5	13,5	221	243	292	221	92	96	100	100	1,00	0,94	1,00	0,80	0,85	0,75	0,58	0,76	58	80	3,2	3,2
E	28	22	23	22	14,5	16	16,5	16	123	158	137	157	92	71	93	94	0,89	0,94	0,84	0,69	0,59	0,46	0,53	0,53				
F	7	8	11	14	15,5	15	14,5	11	169	204	186	230	83	96	80	72	0,89	0,72	0,81	0,58	0.45	0,53	0,46	0,30				
G	6,5	6	†	23	†	17	†	15	126	135	124	125	93	67	77	85	1,00	1,00	0,93	0,76	0,79	0,58	0,61	0,62	22,5	59	40,8	14,5
H	2,5	0,3	17	25	14	17	19	17	182	156	176	180	100	100	84	100	1,00	0,83	0,67	0,86	0,62	0,61	0,44	0,25	76	79	3,8	1,2
I	1,5	1,2	13,5	14	15,5	11	12	9	142	165	162	144	92	100	100	100	0,83	0,72	0,67	0,82	0,30	0,41	0,39	0,37	57,5	63,5	4,2	11,8
J	1,5	2,5	13	14	18,5	11	14	16	178	204	180	218	90	100	80	96	0,94	0,94	0,93	0,88	0,78	0,68	0,42	0,36	42	73,5	6,7	8,1
K	11	9	16	20	13,5	19	15	17	135	151	148	144	86	94	94	100	1,00	1,00	0,82	0,89	0,77	0,63	0,53	0,46	50,5	48	4,7	4,4

† Diese Versuche konnten nicht durchgeführt werden, da das Ästhesiometer zufälligerweise beschädigt wurde.

person in je einem Leistungsgebiete wurden schliefslich wiederum in Ordnungszahlen übergeführt.[1] Nur auf diese Weise ist es möglich, die Werte für ein Leistungsgebiet mit denen für ein anderes vergleichbar zu machen. Wir gelangten also — Tabelle III — zu zwei Rangordnungen für jede der fünf geprüften Leistungs-

Tabelle III.

Rangordnungen nach unseren Ergebnissen (Tab. II).

Versuchsp	Tonunterscheidung			Raumsinn		Addieren		Auswendig lernen		Kombin.-methode	
	K	V	*S*	*K*	*S*	*K*	*S*	*K*	*S*	*K*	*S*
A	5½		6	3	2	7	10	7½	5½		
B	4		4½	11	7	4	7½	7½	5½	3	4
C	10		10	8½	4½	10	9	11	11		
D	2		3	1	6	1	1	1	1	1½	1
E	11		11	8½	8	6	7½	5	7		
F	8		8	3	3	9	5	10	9		
G	7		7	(9)†	9	11	11	2½	2	7	7
H	5½		1	7	11	3	2	6	8	1½	2
I	2		2	3	1	2	4	9	10	4	5
J	2		4½	6	4½	5	3	4	4	6	3
K	9		9	5	9	8	6	2½	3	5	6

† Diese Versuchsreihe blieb, wie gesagt, ohne Ergebnis. Die eingeklammerte Zahl soll nur dazu dienen, die Tabelle übersichtlicher zu machen. Bei allen Berechnungen, in welche diese Zahl hätte eingehen müssen, wurden die Leistungen der Versuchsperson G weggelassen.

fähigkeiten; die eine ist aus den Kruegerschen, die andere aus den Spearmanschen Werten gewonnen.

Schliefslich wurde, nach der früher (S. 52—53) entwickelten Methode, zwischen je zwei Rangordnungen der Korrelationskoeffizient berechnet.

Bevor wir näher auf die sich ergebenden Korrelationswerte eingehen, mufs noch einmal betont werden, dafs die Versuchsreihen kurz waren. Dadurch sind die Ergebnisse keineswegs ungültig, wohl aber weniger ausgiebig geworden.

[1] Die Ordnungszahlen ergeben in manchen Hinsichten zuverlässigere Mafsstäbe, als die reellen Zahlenwerte (vgl. *Am. J. Psych.* **13**, S. 81—82. 1904; auch *British Journ. of Psychol.* **2**, S. 57. 1906).

Wir müssen nämlich ihre kleineren Züge — gleichviel ob sie mit unseren Deutungen stimmen oder nicht — als zunächst vermutlich blofs zufällig vernachlässigen. Wir beachten ausschliefslich diejenigen Zahlen und Zahlenunterschiede, die mindestens zweimal gröfser als der wahrscheinliche Fehler (bzw. wahrscheinliche Unterschied) ausgefallen sind. Hierbei ziehen wir möglichst überall Mittelwerte in Betracht, und zwar immer diejenigen, die sämtliche beteiligten Fälle umfassen.

Man wird bemerken, dafs der wahrscheinliche Fehler, nach der Formel, desto kleiner ausfällt, je gröfser die betreffende Korrelation ist. Es sei aber hinzugefügt, dafs diese Formel sehr lange Vergleichreihen von reellen Zahlenwerten voraussetzt; bei Ordnungszahlen und kurzen Reihen, wie die unsrigen, behält die Formel nur für kleine und mittelgrofse Korrelationskoeffizienten eine genügende Genauigkeit; bei sehr grofsen Korrelationskoeffizienten dagegen wird der wahrscheinliche Fehler zwar immer kleiner, je gröfser die Korrelation, aber noch lange nicht in dem Mafse, wie es von der Formel gefordert wird. Leider ist der wahrscheinliche Fehler verhältnismäfsig gröfser in all den Korrelationen, in welche die Ebbinghaussche Kombinationsmethode eingeht, da wir für diese Versuche, wie schon gesagt, über weniger Versuchspersonen verfügten.

Dadurch, dafs jede Fähigkeit zweimal gemessen wurde, gewannen wir mit Bezug auf je zwei Fähigkeiten, sagen wir A und B, vier Korrelationswerte: einen Wert für die Korrelation zwischen Kruegers Ergebnissen für A und seinen Ergebnissen für B; einen analogen Wert zwischen Spearmans A und B; einen dritten Wert zwischen Kruegers A und Spearmans B, und schliefslich einen Wert zwischen Kruegers B und Spearmans A.

Zuerst haben wir jedesmal das Mittel aus diesen vier Werten gezogen. Daraus ergeben sich die mittleren rohen Korrelationen zwischen den verschiedenen Fähigkeiten, wie sie in Tabelle IV, nebst ihrem wahrscheinlichen Fehler, aufgeführt sind.

Sieben dieser Korrelationswerte (die klein gedruckten) sind ersichtlich noch nicht zweimal gröfser als ihr wahrscheinlicher Fehler; in allen diesen Fällen müssen wir also sagen, dafs zunächst keine Korrelation zur Erscheinung gekommen ist (ob eine Korrelation doch möglicherweise könnte vorhanden sein oder nicht, werden wir erst später mit Hilfe der Ergänzungs- und

Korrelationsformel zu entscheiden imstande sein). Die drei anderen Korrelationen dagegen übersteigen ihren wahrscheinlichen Fehler um mehr als das fünffache; sie haben einen Durchschnittswert von 0,70 (0,79, 0,67, 0,59), und können also als ziemlich sichergestellt angesehen werden.

Tabelle IV.

Durchschnittliche rohe Korrelationen nach anderen Ergebnissen.

Verglichene Rangordnungen (Tab. III)	Durchschn. Korrelationskoeffizienten	Wahrsch. Fehler †
Addieren und Kombinationsmethode	+ 0,79	± 0,06
„ „ Tonunterscheidung	+ 0,67	0,08
„ „ Raumsinn	+ 0,19	0,20
„ „ Auswendiglernen	+ 0,14	0,20
Kombinationsmethode und Tonunterscheidung	+ 0,59	0,12
„ „ Raumsinn	0,00	0,25
„ „ Auswendiglernen	— 0,07	0,25
Tonunterscheidung und Raumsinn	+ 0,29	0,18
„ „ Auswendiglernen	+ 0,17	0,20
Raumsinn „ Auswendiglernen	— 0,13	0,19

† Der erforderliche w. F. der durchschnittlichen Koeffizienten läfst sich nicht genau berechnen. Als ein Annäherungswert dafür ist hier der w. F. der „auf faktischem Wege partiell ergänzten“ Koeffizienten (S. 78) eingetragen worden.

Bevor wir weiter gehen, müssen wir erst sehen, ob diese hohen Korrelationen nur dann bestehen, wenn beide Fähigkeiten an dem selben Tage und von dem selben Versuchsleiter geprüft werden. Denn nehmen wir an, dafs die Korrelation zwischen irgend zwei Fähigkeiten A und B beinahe oder ganz verschwindet, wenn A an einem Tage von dem einen Versuchsleiter, und B eine Woche später von dem anderen Versuchsleiter geprüft wird, dann wäre die durch einen einzigen Versuchsleiter ermittelte Korrelation offenbar blofs auf die verschiedenen momentanen Stimmungen der Beobachter zurückzuführen; sie rührten daher, dafs diejenigen Beobachter, die zufällig gut disponiert waren, sich auf beiden verglichenen Gebieten, sowohl A als B, mehr oder weniger auszeichneten. Aber diese Hypothese findet keine Unterstützung in den tatsächlichen Ergebnissen. Denn das Mittel aller durch einen einzigen Versuchsleiter ermittelten Werte für die drei sichergestellten Korrelationen beträgt 0,68; und das Mittel aller anderen Werte (für diese drei Korrelationen), wo also die zwei verglichenen Messungsreihen von verschiedenen Versuchsleitern herrühren, beläuft sich auf 0,67, hat also fast denselben Wert.[1] Hiermit wird gleichzeitig der Verdacht widerlegt, daſs die Messungen durch „Autosuggestion" seitens der Versuchsleiter beeinfluſst worden seien; denn keiner der beiden Versuchsleiter hat irgend etwas über die Ergebnisse des anderen erfahren, bis zum Schlusse der ganzen Untersuchung.

Eine andere naheliegende Frage ist die, ob die Korrelationswerte bei den zweiten Messungsreihen (Spearman) einen anderen Verlauf zeigen, als bei den ersten (Krueger).

Es ist namentlich Binet[2] zu der Ansicht gekommen, daſs die Zusammenhänge zwischen solchen einfachen Leistungen bei jeder neuen Prüfung stark herabgesetzt würden, und schlieſslich sogar verschwinden könnten. Bei unseren Versuchen ist keine Tendenz in dieser Richtung zutage getreten. Der Mittelwert der drei in Betracht kommenden Korrelationen war bei Krueger 0,61 und bei Spearman 0,75.

2. Anwendung der „Ergänzungsformel".

Um von der Ergänzungsformel Gebrauch machen zu können, muſs man für jede gemessene Leistungsfähigkeit den „Zuverlässigkeitskoeffizienten" der Messungsmethode bestimmen; dies geschieht in unserem Falle, wie gesagt, dadurch, daſs man die Korrelationen zwischen den Kruegerschen und den Spearmanschen Messungen für je eine Fähigkeit berechnet.

Daraus ergeben sich die Werte in Tabelle V.

Es muſs hier ein Punkt berücksichtigt werden, der zwar im gegenwärtigen Falle keine praktische Bedeutung hat, unter anderen Umständen aber wichtig sein könnte. Es wurden nämlich unsere Zuverlässigkeitskoeffizienten aus je zwei an verschiedenen Tagen gewonnenen Messungsreihen berechnet. Mit ihnen sind also nur diejenigen Korrelationen zwischen der einen und der anderen Fähigkeit streng vergleichbar, wobei die zwei

---

[1] Der Mittelwort sämtlicher überhaupt durch einen einzigen Versuchsleiter ermittelten Korrelationskoeffizienten beträgt 0,25; während das Mittel sämtlicher anderen Korrelationskoeffizienten, wo also die zwei verglichenen Messungsreihen von verschiedenen Versuchsleitern herrühren, sich auf 0,26 beläuft.

[2] *Année psychologique*, 1900, 6, S. 395.

Tabelle V.

Unsere „Zuverlässigkeitskoeffizienten".

Korrelationen zwischen den Kruegerschen und den Spearmanschen Messungen für:	Koeffizienten	Wahrscheinl. Fehler
Addieren	0,76	± 0,07
Kombinationsmethode	0,75	0,09
Tonunterscheidung	0,87	(0,04)*
Raumsinn	0,42	0,15
Auswendiglernen	0,92	(0,03)*

* Diese sich aus der Formel ergebenden w. F. müssen als erheblich zu klein betrachtet werden (s. S. 74).

unter sich verglichenen Messungsreihen ebenfalls an verschiedenen Tagen gewonnen wurden; d. h. wir dürften hier nicht die Krueger-Kruegerschen oder die Spearman-Spearmanschen Korrelationen verwerten, sondern nur die Krueger-Spearmanschen und die Spearman-Kruegerschen. Dann vereinfacht sich die Ergänzungsformel zu folgender:

$$\mathrm{AB} = \frac{M\,(A_1 B_2,\ A_2 B_1)}{M\,(A_1 A_2,\ B_1 B_2)} \tag{d}$$

Wenn wir diese vereinfachte Ergänzungsformel anwenden, und zwar beispielsweise auf die Korrelation zwischen dem Addieren und der Unterscheidungsfähigkeit für Töne, so bekommen wir die Gleichung:

$$\mathrm{AdT} = \frac{M(Ad_1 T_2,\ Ad_2 T_1)}{M(Ad_1 Ad_2,\ T_1 T_2)} = \frac{M(0{,}76\ \ 0{,}55)}{M(0{,}76\ \ 0{,}87)} = 0{,}80.$$ [1]

Der letztere Wert ist also als die durch Ergänzung gewonnene reine Korrelation zwischen dem Addieren und der Tonunterscheidung anzusehen. In ganz analoger Weise bekommen wir für die Korrelationen zwischen dem Addieren, bzw. der Tonunterscheidung, mit den Kombinationsversuchen die ergänzten

[1] Wenn wir die vollständige Ergänzungsformel benutzen, so ergibt sich

$$\mathrm{AdT} = \frac{M(Ad_1 T_1,\ Ad_1 T_2,\ Ad_2 T_1,\ Ad_2 T_2)}{M(Ad_1 Ad_2,\ T_1 T_2)} = \frac{M(0{,}72\ \ 0{,}58\ \ 0{,}76\ \ 0{,}67)}{M(0{,}76\ \ 0{,}87)} = 0{,}83$$

also in diesem Falle beinahe derselbe Wert. Es sei daran erinnert, daß M den Mittelwert bedeutet (s. S. 57).

(also reinen) Werte von 0,93 bzw. 0,81. Der Durchschnittswert dieser drei Korrelationen ist also durch Anwendung der Ergänzungsformel von 0,68 bis auf 0,85 gestiegen; der letztere Wert wäre nämlich tatsächlich zu erwarten, wenn die Fähigkeit jeder Versuchsperson durch sehr ausführliche und oft wiederholte Messungen ganz fehlerfrei festgestellt würde.

Eine gewisse Annäherung an diese theoretisch vollständige Ergänzung läfst sich dadurch erzielen, dafs wir zwischen den Messungsergebnissen von Krueger und denen von Spearman Mittelwerte ziehen, und die Korrelationen nun zwischen diesen Mittelwertsreihen berechnen. Daraus ergeben sich die drei Korrelationskoeffizienten: 0,73, 0,84 und 0,64; der Durchschnittswert ist bis auf 0,74 gestiegen. Dieser auf faktischem Wege partiell ergänzte Wert erreicht also in der Tat eine Gröfse etwa halbwegs zwischen dem unergänzten und dem durch die Formel völlig ergänzten Werte, gerade wie es die Theorie erfordert.

Alles in allem scheint sichergestellt zu sein, dafs die Unterscheidungsfähigkeit für Töne tatsächlich eine hohe Korrelation mit den scheinbar so grundverschiedenen Fähigkeiten besitzt, wie die zum Addieren und die, welche durch die Ebbinghaussche Kombinationsmethode beansprucht wird. Zwischen den beiden letzten Fähigkeiten zeigt sich ebenfalls eine hohe Korrelation.

Jetzt wenden wir uns schliefslich zu den Korrelationen, die verschwindend klein ausgefallen sind. Dies sind ersichtlich alle diejenigen, in denen entweder der Raumsinn oder das Auswendiglernen eine der beiden verglichenen Fähigkeiten bildet. Nun ist ein solcher scheinbarer Mangel an Korrelation zunächst, wie oben auseinandergesetzt, zweideutig; er kann wirklich sein, oder aber er kann blofs von übergrofsen zufälligen Fehlern herrühren. Der zweite Fall wird sich in einem kleinen Zuverlässigkeitskoeffizienten, d. h. in einer sehr kleinen Korrelation zwischen den Kruegerschen und den Spearmanschen Werten für dieselbe Fähigkeit kund geben (vgl. Tabelle V).

Wir finden, dafs in dieser Hinsicht der Raumsinn und das Auswendiglernen ein merkwürdig entgegengesetztes Verhalten zeigen. Im Gebiete des Raumsinnes decken sich die zwei Messungsreihen von Spearman bzw. Krueger so unvollständig, dafs die Korrelation zwischen ihnen nur 0,42 beträgt. Hierin haben wir schon einen genügenden Grund für die scheinbar so kleinen Korrelationen des Raumsinnes mit den anderen Fähigkeiten; denn wenn man diesen Wert (0,42) in die Ergänzungsformel einsetzt und auch noch den durch den wahrscheinlichen Fehler angedeuteten Spielraum des Zufalls in Betracht zieht, so findet man, daſs der Raumsinn möglicherweise ziemlich hohe reine Korrelationen mit den anderen Fähigkeiten besitzen könnte. Die gegenwärtigen Versuche haben keine Entscheidung darüber herbeizuführen vermocht.

Diese übergroſse Diskrepanz zwischen den Messungsreihen der beiden Versuchsleiter kann zwei Gründe haben. Entweder liegt sie daran, daſs die zwei Messungen derselben Versuchsperson an verschiedenen Tagen durchgeführt worden sind, und deshalb anderen Bedingungen in bezug auf Disponiertheit, Ermüdung, Temperatur usf. unterlagen. Oder aber die Methode der Schwellenbestimmung war ungenügend für den von uns verfolgten Zweck. Da nun die Möglichkeit einer solchen zweideutigen Diskrepanz von vornherein unschwer vorauszusehen war, so haben wir zur Kontrolle die Schwellen sowohl des rechten wie auch des linken Jochbeins bestimmt (s. S. 69); wenn die oben besprochene Diskrepanz nur von Veränderungen der Bedingungen am zweiten Versuchstage herrührte, so müſsten immer noch die beiden Jochbeine demselben Versuchsleiter am selben Tage gut unter sich übereinstimmende Werte ergeben. Dies ist aber nicht der Fall; die Korrelation zwischen den Sensibilitäten des rechten und des linken Jochbeins bei demselben Versuchsleiter beläuft sich nur auf 0,36 (Durchschnittswert zwischen Kr. und Sp.). Wir haben also hier ein Beispiel, wie leicht eine Messungsmethode ungenügend sein kann, um irgend eine Korrelation, sei sie auch wirklich vorhanden, hervortreten zu lassen.[1]

Ganz anders verhält sich die Sache bei dem Auswendiglernen; denn hier beläuft sich der Korrelationskoeffizient zwischen

[1] Es sei daran erinnert, daſs es bei Korrelationen nicht auf die absolute Gröſse der zufälligen Messungsfehler, sondern auf das Verhältnis dieser Gröſse zur mittleren individuellen Variation ankommt. Diese Variation ist beim Raumsinn besonders klein, wie aus Tabelle II zu sehen ist. Die absolute Diskrepanz zwischen den Messungen für das rechte bzw. linke Jochbein beträgt in unseren Ergebnissen durchschnittlich nur 2 mm; die Messungsmethode ist also für viele Zwecke vollständig ausreichend.

der KRUEGERschen und der SPEARMANschen Messungsreihe auf nicht weniger als 0,92, ist also beinahe vollkommen. In diesem Falle vermag die Ergänzungsformel keine merkliche Erhöhung der Korrelationen zwischen dieser und den anderen Fähigkeiten hervorzubringen; wenn hier keine Korrelation zum Vorscheine kommt, so ist auch in Wirklichkeit keine vorhanden. Daraus ergibt sich das wichtige Resultat, dafs das Auswendiglernen von Zahlenreihen tatsächlich keine beträchtliche Korrelation mit irgend einer anderen der geprüften Fähigkeiten besitzt.

3. Anwendung der Korrektionsformel.

Schliefslich müssen wir sehen, inwieweit alle diese Korrelationswerte, trotz der oben genannten Vorsichtsmafsregeln, als durch konstante ungehörige Faktoren gestört anzusehen sind.

Fast jeder dieser Faktoren gestattete, nach ihrer von Fall zu Fall wechselnden Gröfse, die ungezwungene Herstellung einer Rangordnung; entweder auf Grund objektiver Messungen (Temperatur u. dgl.), oder auf Grund der im Protokoll enthaltenen Bemerkungen.

Wir wollen zuerst den Fall betrachten, wo der nicht zugehörige Faktor einen Zusammenhang mit beiden verglichenen Fähigkeiten haben könnte. Dieser Fall ist besonders gefährlich, weil er die Korrelation vergröfsert und einen Zusammenhang vortäuschen kann, wo keiner existiert. Hier scheinen besonders die Nationalität, das Alter, die allgemeine Gesundheit, der Eifer, und die Ermüdung oder sonstiges momentanes Angegriffensein der Versuchspersonen in Betracht zu kommen.

Am bedrohlichsten ist der Faktor der Nationalität. Denn sowohl das Addieren wie die Tonunterscheidung zeigen beträchtliche Korrelationen, von 0,49 bzw. 0,68, mit der Nationalität, und zwar in dem Sinne, dafs die Deutschen den Ausländern überlegen sind. Durch Anwendung der Ergänzungsformel steigt der erste Wert auf $\frac{0{,}49}{M(0{,}76,\ 1{,}00)} = 0{,}56$, der zweite Wert auf $\frac{0{,}68}{M(0{,}87,\ 1{,}00)} = 0{,}74$. Hier wäre zunächst eine bedenkliche illusorische Vergröfserung zu vermuten; denn wenn wir die Korrektionsformel anwenden (S. 59), so bekommen wir:

$$\overline{\mathrm{AdT}} = \frac{\mathrm{AdT} - \mathrm{AdN}\cdot\mathrm{TN}}{\sqrt{(1-\mathrm{AdN}^2)\,(1-\mathrm{TN}^2)}}$$

$$= \frac{0{,}80 - 0{,}56 \times 0{,}74}{\sqrt{(1-0{,}56^2)\,(1-0{,}74^2)}} = 0{,}69.$$

Danach wäre der wahrscheinlich wirkliche Wert **AdT** nicht 0,80, sondern nur 0,69. Aber wenn wir die Ergebnisse für die Ausländer ausschalten, und jetzt die Korrelationskoeffizienten für ausschliefslich deutsche Versuchspersonen berechnen, so ergibt sich nicht die geringste Verkleinerung. Daraus sollte man schliefsen, dafs die Verschiedenheit der Leistungen wegen verschiedener Nationalität eigentlich kein neues zu dem untersuchten Zusammenhang zufällig hinzukommendes Moment ist, sondern vielmehr selbst erst von diesem Zusammenhang herrührt.[1] Diese Frage kann jedoch jetzt nur andeutungs- und anregungsweise berührt werden, da die fraglichen Zusammenhänge zwischen der Nationalität und den untersuchten Leistungsfähigkeiten den sehr grofsen wahrscheinlichen Fehler von 0,28 haben (dies kommt daher, dafs hier ein Glied der Korrelation, nämlich die Nationalität, nur zwei Stufen darbot, „deutsch“ und „ausländisch“; deshalb war die r-Methode unbrauchbar, und eine andere, viel weniger zuverlässigere mufste herangezogen werden. Vgl. *Am. J. Psych.* 15, S. 85).

Das Alter (20—38 Jahre) zeigte keine merkliche Korrelation mit irgend einer der fünf geprüften Fähigkeiten (Maximum 0,22; Durchschnitt aller positiven und negativen Werte 0,02).

In bezug auf Gesundheit haben die Versuchspersonen, wie es scheint, nur in einer einzigen Hinsicht irgend etwas zu wünschen gelassen: etwa die Hälfte von ihnen gab an, mehr oder weniger „nervös“ oder „neurasthenisch“ zu sein. Es stellte sich aber heraus, dafs diese Angaben über Neurasthenie u. dgl. keine merkliche Korrelation mit irgend einer der Leistungen besafsen; der durchschnittliche Wert aller fünf Korrelationen war gleich Null (hier war dieselbe Rechnungsweise nötig, wie für die Nationalität).

Ebensowenig zeigte irgend eine der Fähigkeiten eine merk-

[1] Es wäre nämlich nach dieser Hypothese der „Zentralfaktor“ (s. S. 88) etwas besser entwickelt bei den deutschen, als bei den anderen hier in Betracht kommenden Rassen.

liche Korrelation mit dem scheinbaren Eifer der Versuchspersonen. Der durchschnittliche Wert der fünf Korrelationen dieser Art war —0,03. Dieses Resultat ist um so auffallender, als die KRUEGERschen und die SPEARMANschen Beobachtungen über den scheinbaren Eifer der Versuchspersonen, obgleich an verschiedenen Tagen angestellt, doch die ziemlich hohe Korrelation miteinander von 0,66 aufwiesen.

Ermüdung und sonstiges Angegriffensein haben wir von vornherein ausgeschaltet, indem wir die Versuchspersonen niemals in diesem Zustande prüften.

Demnach scheinen die Korrelationen jedenfalls nicht illusorisch vergröſsert zu sein. Wenden wir uns zu dem Falle, wo ein Zusammenhang zwischen einem fremden Faktor und nur einer der verglichenen Fähigkeiten nicht ausgeschlossen scheinen könnte. In diesem Falle wäre die Korrelation, wie früher gesagt, als in ungehöriger Weise verkleinert anzusehen.

In bezug auf den Raumsinn haben wir besonders an eine mögliche Beeinflussung durch physikalische Bedingungen gedacht, und deshalb notierten wir die Temperatur des Zimmers (sie schwankte zwischen 12,7° und 17,7° C), diejenige im Freien (—1,6° bis +6,1°), die Luftfeuchtigkeit (40—76) und die Barometerhöhe (731—764). Aber keiner dieser Umstände zeigt einen merklichen Zusammenhang mit dem Raumsinne (ihre Korrelationskoeffizienten mit dem letzteren waren 0,04, bzw. —0,04, —0,12 und 0,06).

Belanglos scheint ebenfalls der Zusammenhang zwischen dem Addieren und dem Grade der Beschäftigung der Versuchspersonen mit Mathematik gewesen zu sein. Denn hier bekamen wir einen Koeffizienten von nur 0,22, also nicht zweimal gröſser als der wahrscheinliche Fehler; auſserdem bekommen wir, wenn wir diesen Wert in die Korrektionsformel einsetzen, nach der Formel (c), S. 60.

$$\overline{AB} = \frac{AB}{\sqrt{1-0{,}22^2}} = \frac{AB}{0{,}98}.$$

Die reine Korrelation würde also durch diesen fremden Faktor nur um 2 % verkleinert werden, was bei unseren Versuchen zu vernachlässigen ist.

Bedeutungsvoller zeigte sich der Koeffizient zwischen der Unterscheidungsfähigkeit für Tonhöhen und der Beschäftigung der Versuchsperson mit Musik; denn hier ergab sich, trotz unserer erwähnten Ausschliefsung der beiderlei extremsten Klassen von Versuchspersonen, immer noch eine Korrelation von 0,46. Wenn wir diesen Wert in die Korrelationsformel einsetzen, so bekommen wir

$$AB = \frac{AB}{\sqrt{1-0{,}46^2}} = \frac{AB}{0{,}89}.$$

Danach wären die Korrelationen der Fähigkeit zur Tonunterscheidung mit den anderen Fähigkeiten um etwa 11 % in ungehöriger Weise herabgesetzt. Dieser Betrag könnte selbst bei so kurzen Versuchsreihen eine gewisse Bedeutung besitzen; er ist aber, wie sich zeigen wird, für die Gedankenentwicklung in der gegenwärtigen Abhandlung ohne Belang. Es scheint jedenfalls mehr Grund vorhanden, die angegebenen Korrelationen der Tonunterscheidung mit den anderen Fähigkeiten zu niedrig als zu hoch anzusehen.

4. Die „Zentralwerte" der verschiedenen Leistungsfähigkeiten.

Jetzt erst, nach dieser langwierigen Prüfung der Reinheit der beobachteten Korrelationen, ist es gestattet, sie zu weiteren wissenschaftlichen Überlegungen zu verwerten.

Es fragt sich nun vor allem, ob die Korrelation zwischen je zwei Fähigkeiten als jeweils ein isoliertes Phänomen aufzufassen ist, oder aber ob nicht alle solche Korrelationen von einer einzigen Ursache herrühren können (es sei darauf hingewiesen, dafs ähnliche Korrelationen in einer früheren Arbeit auch zwischen neun anderen geistigen Leistungsfähigkeiten aufgedeckt worden sind).[1]

Zunächst wollen wir uns der naheliegenden und wenigstens zu heuristischen Zwecken zulässigen Hypothese bedienen, dafs irgend zwei zusammenhängende geistige Leistungsfähigkeiten als teilweise von einem gemeinsamen Faktor abhängig angesehen werden dürfen. Diese Hypothese läfst noch Spielraum für die verschiedensten späteren näheren Bestimmungen; die zwei Fähigkeiten könnten z. B. teilweise durch dasselbe physiologische Organ vermittelt werden; oder ihre Organe könnten unter einem gemeinsamen Einflusse stehen; oder der gemein-

[1] *Am. J. Psych.*, 13, S. 202—293.

same Faktor könnte einen noch indirekteren oder abstrakteren Charakter haben.

Nehmen wir an, daſs die zwei Fähigkeiten P_1 und P_2 den gemeinsamen Faktor **P**, und daſs zwei andere Fähigkeiten Q_1 und Q_2 den gemeinsamen Faktor **Q** haben. Wenn die Korrelationen P_1P_2 und Q_1Q_2 von einer und derselben Ursache herrühren, so ist eben damit gesagt, daſs P_1 und P_2 denselben gemeinsamen Faktor wie Q_1 und Q_2 haben, oder mit anderen Worten, daſs **P** dasselbe bedeutet wie **Q**. In diesem Falle müssen die Wirkungen von **P** und **Q** sich vollständig decken, also muſs die Korrelation $\mathbf{PQ} = 1$ sein. Wir haben somit das Problem auf eine mathematische Aufgabe reduziert; es handelt sich nur noch darum, die Korrelation **PQ** zu ermitteln und zu sehen, ob sie tatsächlich immer $= 1$ ist.

Nun aber wird man eine ausgezeichnete Leistung in der Fähigkeit P_1 als einen gewissen Hinweis darauf ansehen dürfen, daſs bei dieser Person auch **P** nicht schlecht entwickelt ist (da ja nach der Hypothese P_1 teilweise von **P** abhängt). Die Messung von P_1 ergibt also ebenfalls eine, allerdings möglicherweise mit sehr groſsen zufälligen Fehlern behaftete Messung von **P**.[1] In gleicher Weise ergibt auch P_2 eine gewisse Messung von **P**, während Q_1 und Q_2 je eine von **Q** liefern.

Die zufälligen Fehler, seien sie noch so groſs, lassen sich durch unsere Ergänzungsformel kompensieren. Kurz, wenn alle Korrelationen von einer einzigen Ursache herrühren, so muſs jedesmal

$$\mathbf{PQ} = \frac{M\,(P_1Q_1,\ P_1Q_2,\ P_2Q_1,\ P_2Q_2)}{M\,(P_1P_2,\ Q_1Q_2)} = 1\ [2]$$

---

[1] Eine und dieselbe Wertreihe kann also als eine Messungsreihe von P_1 sowohl als von P oder auch von irgend einem anderen Merkmal aufgefaſst werden, von welchem P_1 teilweise oder gänzlich abhängt. Dabei aber bleibt nichts der Willkür überlassen. Denn sobald zwei Wertreihen zusammen genommen und als Messungen eines und desselben Merkmals angesehen werden, so ist dieses eindeutig bestimmt; es umfaſst nämlich alles — weder mehr noch weniger, — wovon die beiden Wertreihen eine gemeinsame Abhängigkeit haben.

[2] Hier bedeutet M, wie öfters erwähnt, den Mittelwert. Die Messungsreihen P_1 und P_2 (bzw. Q_1 und Q_2) müssen im allgemeinen als ungleich genau betrachtet werden. Deshalb hat man, streng genommen, nicht für den Nenner, sondern auch für den Zähler den geometrischen Mittelwert zu nehmen. Aber bei den kurzen hier in Betracht kommenden Reihen sind doch die arithmetischen Mittel vorteilhafter, s. Anm. 1, S. 57.

sein. Da P_1, P_2, Q_1, Q_2 beliebige Fähigkeiten darstellen, so wollen wir sie durch die Buchstaben A, B, C, D ersetzen. Dann hat man:

$$\frac{M\,(AC,\ AD,\ BC,\ BD)}{M\,(AB,\ CD)} = 1 \tag{e}$$

oder schliefslich

$$\boldsymbol{M\,(AB,\ CD) = M\,(AC,\ AD,\ BC,\ BD)} \tag{f}$$

wo A, B, C, D irgendwelche (nicht allzusehr verwandte) miteinander verglichene Leistungsfähigkeiten darstellen.

Diese Gleichung läfst sich nun ohne weiteres auf unsere Ergebnisse anwenden. Dabei können wir im gegenwärtigen Falle zu den r o h e n Werten der Korrelationskoeffizienten zurückkehren. Denn die auf Seite 77—78 vorgenommenen Ergänzungen sowie die auf Seite 82 besprochene Korrektion würden blofs beide Seiten der Gleichung mit denselben Faktoren multiplizieren. Korrektionen nach der Formel (b) S. 60 würden freilich die Gleichung beeinflussen, haben sich aber bei unseren Ergebnissen als unnötig erwiesen. Durch dieses Zurückgehen auf die rohen Werte selbst wird nicht nur jede Unsicherheit der Korrigierungsvorgänge umgangen, sondern es wird ermöglicht, auch noch den Raumsinn in Betracht zu ziehen.

Wenn man die fünf geprüften Fähigkeiten durch ihre Anfangsbuchstaben darstellt, so müfste z. B. gelten — wenn die Hypothese des einzigen gemeinsamen Faktors zutrifft —

$$M\,(AdR,\ KT) = M\,(AdK,\ AdT,\ RK,\ RT)$$

oder, wenn man die tatsächlich gewonnenen Werte einsetzt,

$$M\,(0{,}19,\ 0{,}59) = M\,(0{,}79,\ 0{,}67,\ 0{,}00,\ 0{,}29).$$

Die rechte Seite hat den Wert 0,39 und die linke den Wert 0,46, so dafs die Diskrepanz sich auf 0,07 beläuft. Mit sämtlichen in Tabelle IV aufgeführten Korrelationswerten sind 15 solche Gleichungen herstellbar; die durchschnittliche Diskrepanz ist 0,08 und die maximale 0,15. Da die wahrscheinliche blofs zufällige Diskrepanz schon etwa 0,10 beträgt, so sind unsere Ergebnisse, so weit sie reichen, in guter Übereinstimmung mit der Theorie, dafs alle Korrelationen von einem e i n z i g e n g e m e i n s a m e n F a k t o r — nennen wir ihn kurz „Z e n t r a l f a k t o r" — herrühren.

Bei kurzen Versuchsreihen läſst sich aber diese Hypothese eines einzigen Zentralfaktors besser und anschaulicher durch eine andere aus der obigen leicht ableitbare mathematische Beziehung prüfen. Wenn irgend eine Korrelation AF gröſser als BF ist, so muſs dann — nach unserer Hypothese — auch jede andere Korrelation AG (bzw. AH, AI, . . .) gröſser als BG (bzw. BH, BI, . . .) ausfallen; im Falle, daſs BG (bzw. BH, BI, . . .) $= 0$ ist, muſs auch AG (bzw. AH, AI, . . .) $= 0$ sein.[1] Durch diese Verhältnisse müssen die sämtlichen Korrelationswerte sich in eine derartige hierarchische Anordnung bringen lassen, daſs jeder Wert (auſser den Nullwerten) gröſser ist als alle nach rechts oder nach unten stehenden; und wenn ein einziger Nullwert vorkommt, so muſs die ganze betreffende Linie, entweder in vertikaler oder in horizontaler Richtung, lauter Nullwerte zeigen. Wir wollen dies an unseren Ergebnissen prüfen (und dabei alle Korrelationen gleich 0 setzen, die nicht mindestens 2 mal gröſser als der wahrscheinliche Fehler sind). Zum Vergleich sollen auch noch die Zuverlässigkeitskoeffizienten herangezogen werden *(Kursivschrift)*. Dann bekommen wir folgendes Bild:

Tabelle VI.

	Addieren	Komb.-Meth.	Töne	Raumsinn	Ausw. Lern.
Addieren	*0,75*	0,70	0,68	0,00	0,00
Komb.-Meth.	0,70	*0,90*	0,64	0,00	0,00
Tonuntersch.	0,68	0,64	*0,87*	0,00	0,00
Raumsinn	0,00	0,00	0,00	*0,49*	0,00
Ausw. Lern.	0,00	0,00	0,00	0,00	*0,92*

Bei der Herstellung dieser Tabelle haben wir uns (wie bei der Anwendung der Ergänzungsformel) nur derjenigen Korrelationskoeffizienten bedient, welche Messungsreihen an verschiedenen Tagen ergeben haben; denn nur solche sind, wie gesagt, streng vergleichbar mit den Zuverlässigkeitskoeffizienten. Diese letzteren machen hier die diagonale Wertreihe aus, die so stark gegen den Verlauf der anderen Korrelationen (zwischen verschiedenen Fähigkeiten) kontrastiert.

Man könnte vielleicht denken, daſs solche hierarchische Anordnungen auch bei ganz zufälligen Kombinationen von Werten leicht herstellbar wären. Dies ist jedoch nicht der Fall. Eine der-

[1] All dies gilt natürlich nur innerhalb der durch den w. F. bestimmten zufälligen Schwankungsbreite.

artige Anordnung der obigen Werte z. B. wäre ganz unmöglich gewesen, wenn irgend eine der drei — nach unseren Ergebnissen — hohen Korrelationen = 0 wäre; und ebenso wenn umgekehrt irgend einer unserer Nullwerte durch eine hohe Korrelation ersetzt wäre.[1]

Zum Vergleich wollen wir die ähnlichen, aber umfangreicheren Ergebnisse einer früheren Untersuchung heranziehen. Damals dienten 36 Schüler als Versuchspersonen. Vier der verglichenen Leistungsgebiete waren die Schulfächer, nämlich klassische Sprachen, Französisch, Englisch und Mathematik; die relativen Fähigkeiten der Schüler wurden durch ihre Rangordnung in den Schulprüfungen bestimmt. Eine fünfte verglichene Fähigkeit war die musikalische Begabung; die Rangordnung der Schüler in dieser Hinsicht wurde vom Musiklehrer aufgestellt.[2] Und sechstens wurde auch ihre verschiedene Fähigkeit der Tonunterscheidung geprüft. Daraus ergaben sich folgende Korrelationswerte:[3]

Tabelle VII.

	Kl. Spr.	Franz.	Engl.	Math.	Töne	Musik
Klass. Spr.	*0,87*	0,83	0,78	0,70	0,66	0,63
Französ.	0,83	*0,84*	0,67	0,67	0,65	0,57
Engl.	0,78	0,67	*0,89*	0,64	0,54	0,51
Mathem.	0,70	0,67	0,64	*0,88*	0,45	0,51
Tonunterscheidung	0,66	0,65	0,54	0,45		0,40
Musik	0,63	0,57	0,51	0.51	0,40	

Man sieht, daſs die oben besprochene hierarchische Anordnung sich wiederum durchweg herstellen läſst (die einzige kleine Ausnahme liegt wohl innerhalb der zufälligen Schwankungsbreite).[4]

[1] Setzt man 3 positive Werte und 7 Nullwerte voraus, wie sie hier vorkommen, so wird die Möglichkeit einer solchen hierarchischen Anordnung durch bloſsen Zufall nur 1 mal von 12 vorkommen, wie sich leicht berechnen läſst.

[2] Zur Methode der Aufstellung einer solchen Rangordnung vgl. S. 108.

[3] *Am. J. Psych.* 15, 1904, S. 75.

[4] Diese Schüler waren jung, etwa 9—11 Jahre alt. Der eine von uns hat die Korrelationen zwischen den verschiedenen Studienfächern auch in anderen Schulen unter älteren Kindern untersucht (noch nicht veröffentlicht). Meistens bewährt sich die hierarchische Anordnung nicht. Es bilden sich nämlich, aus manchen wohl begreifbaren Gründen, spezielle Verbindungen zwischen zwei oder mehr Studienfächern; wer sich z. B. mit den Naturwissenschaften besonders erfolgreich abgibt, wird wahrscheinlich auch der Mathematik ein besonderes Interesse schenken.

Mit Hilfe des hypothetischen Zentralfaktors und der Ergänzungsformel läfst sich ferner ein Wert von besonders hohem Interesse ermitteln. Sobald man nämlich für irgend eine Fähigkeit **A** die zwei Messungsreihen A_1 und A_2, und für irgend zwei andere Fähigkeiten je eine Messungsreihe B bzw. C gewonnen hat, so kann man für **A** einen konstanten Wert, nämlich seine Korrelation mit dem Zentralfaktor, kurz seinen „Zentralwert", angeben. Wenn man dann **A** zu irgend einer anderen Zeit und bei irgend einer anderen Messungszuverlässigkeit mit irgend zwei weiteren Fähigkeiten **D** und **E** vergleicht, so sollte man — nach der Hypothese des Zentralfaktors — den genannten Zentralwert immer wieder bekommen.

Man verfährt in folgender Weise. Die Wertreihen B und C lassen sich als Messungen des Zentralfaktors, sagen wir **Z**, auffassen (vgl. Anm. 1, S. 84); und die Wertreihen A_1 und A_2 sind nach Voraussetzung Messungen von **A**. Dann wird der gesuchte konstante Korrelationswert **AZ** durch folgende Gleichung bestimmt:

$$\mathbf{AZ} = \frac{M(AB,\ AC)}{M(A_1 A_2,\ BC)} \tag{g}$$

Die zwei Tabellen VI und VII setzen uns instand, auf diese Formel sogleich die Probe zu machen. Denn sowohl in den gegenwärtigen Versuchen wie in der älteren unter völlig anderen Bedingungen durchgeführten Untersuchung an Schulkindern war eine der verglichenen Fähigkeiten die Unterscheidungsfähigkeit für Tonhöhen. Wenn wir nun die Gleichung (g) benutzen und die Werte von Tabelle VI einsetzen, so bekommen wir für die gegenwärtigen Versuche:

$$\mathbf{TZ} = \frac{M(TK,\ TAd)}{M(T_1 T_2,\ KAd)} = \frac{M(0{,}65,\ 0{,}66)}{M(0{,}87,\ 0{,}71)} = 0{,}83$$

In den älteren Versuchen ist die Ermittlung von **TZ** dadurch erschwert, dafs damals nur eine Messungsreihe vorgenommen wurde, weshalb die Gröfse $T_1 T_2$ nicht genau feststellbar ist; wir müssen uns der freien Schätzung von $T_1 T_2$ bedienen, die schon damals vollzogen und in die Ergänzungsformel eingeführt wurde (Näheres s. *A. J. P.* 15, 1904, S. 65); diese Schätzung betrug 0,64. Wenn wir nun diesen Wert und die damaligen Korrelationen zwischen den klassischen Sprachen bzw. dem Englischen und der Tonunterscheidung (Tabelle VII) in die Gleichung (g) einsetzen, so bekommen wir:

$$\mathsf{TZ} = \frac{M(TKl,\ TE)}{M(T_1T_2,\ Kl\,E)} = \frac{M(0{,}66,\ 0{,}54)}{M(0{,}64,\ 0{,}78)} = 0{,}83$$

Setzen wir, statt der klassischen Sprachen und des Englischen, das Französische und die Mathematik (wiederum Tabelle VII) in die Gleichung ein, so ergibt sich:

$$\mathsf{TZ} = \frac{M(TF,\ TM)}{M(T_1T_2,\ FM)} = \frac{M(0{,}65,\ 0{,}45)}{M(0{,}64,\ 0{,}67)} = 0{,}84$$

Solche genaue Übereinstimmung des gegenwärtigen mit den beiden älteren Zentralwerten für die Tonunterscheidung ist natürlich, unter den vorliegenden Umständen, zufällig.[1] Wichtig ist aber, daſs diese Anwendungen der Gleichung (g) wenigstens keine sich entschieden widersprechenden Werte ergeben. Und noch wichtiger ist die Tatsache, daſs diese Gleichung irgend welchen späteren viel ausgedehnteren Versuchsreihen ein präzises und sehr leicht anwendbares Kontrollmittel an die Hand gibt.

Wenn wir jetzt die Gleichung auch auf die gegenwärtigen Ergebnisse in bezug auf das Addieren bzw. die Kombinationsmethode anwenden, so bekommen wir noch gröſsere Zentralwerte: **AdZ** = 0,97; **KZ** = 0,97. Solche groſse Zentralwerte deuten darauf hin, daſs die betreffenden Leistungsfähigkeiten zu dem Zentralfaktor in sehr enger Abhängigkeit stehen; dadurch scheinen sie besonders geeignet zu sein, die nähere Beschaffenheit dieses merkwürdigen, an allen bisher untersuchten Fähigkeiten (auſser dem Auswendiglernen) mehr oder weniger stark beteiligten Zentralfaktors etwas näher zu beleuchten. Darauf werden wir später wieder zurückkommen.

IV. Die Ergebnisse von Oehrn.

1. Oehrns Untersuchungsmethoden.

Der Zusammenhang zwischen verschiedenen einfachen Leistungsfähigkeiten ist schon früher ziemlich zahlreichen experimentellen Untersuchungen unterzogen worden. Da nun die meisten oben gewonnenen Schlüsse einen sehr allgemeinen

[1] Die auf S. 83 als möglich gefundene Störung der gegenwärtigen Ergebnisse durch die verschiedene Beschäftigung der Versuchspersonen mit Musik spricht nicht gegen diese Übereinstimmung, da eine ähnliche Störung auch in den älteren Werten angenommen werden müſste.

Charakter haben, so sollten sie sich — könnte man meinen — auch bei den älteren Ergebnissen bewähren, wenn man nur diese derselben mathematischen Behandlung unterwürfe.

Aber leider sind in diesen älteren Untersuchungen selten die ursprünglichen Ergebnisse mitgeteilt; wir bekommen zwar manche umfangreiche und ingeniöse Tabellen, aber gerade diejenigen Daten fehlen, die den Grad der Korrelation genau zu bestimmen gestatten. Ferner sind diese Ergebnisse gröfstenteils durch allerlei offenbar höchst störende Nebenfaktoren getrübt.[1]

Eine glänzende Ausnahme in beiden Hinsichten bildet die allererste solcher Untersuchungen, die schöne Arbeit von Oehrn zur „Individual-Psychologie".[2] Er teilt seine rohen Ergebnisse ausführlich mit, und ist von vornherein darauf bedacht gewesen, „gröfsere Differenzen der Versuchspersonen in bezug auf Alter, Bildungsgrad, sowie Gewöhnung an die Versuchsbedingungen" auszuschalten.[3] Seine Versuche scheinen mit einer Sorgfalt durchgeführt zu sein, wie sie von seinen Nachfolgern auf diesem Gebiete wohl selten wieder erreicht worden ist.

Wir wollen jetzt also Oehrns Ergebnisse derselben Behandlung wie die unsrigen unterwerfen. Sollte es uns gelingen, aus diesen seit 17 Jahren allgemein bekannten, höchst zuverlässigen Angaben manche neue und wichtige Gesetzmäfsigkeiten abzuleiten, so würden unsere Methoden eine entscheidende Probe bestanden haben. Und sollten diese Gesetzmäfsigkeiten eine durchgehende Übereinstimmung mit den Ergebnissen unserer eigenen Versuche aufweisen, so wäre damit für diese eine bemerkenswerte Bestätigung gewonnen.

Zuerst mögen Oehrns Versuchsmethoden kurz beschrieben werden.[4] Jede Versuchsperson wurde allein geprüft.

A. „Um die individuelle Beschaffenheit des Wahrnehmungsvorgangs kennen zu lernen", wurde den Ver-

[1] Die in Frage stehenden Untersuchungen sind bei einer früheren Gelegenheit aufgeführt und kritisiert worden. *Am. J. Psych.* **13**, S. 6. 1904.

[2] Experimentelle Studien zur Individual-Psychologie, Dorpater Diss., 1889. Auch in „Kraepelins Psychologische Arbeiten", Bd. I.

[3] Psychol. Arbeiten I, S. 105. Seine 10 Versuchspersonen waren zwischen 21 und 33 Jahre alt. Darunter waren 5 Drs. med., 3 Stud. med., 1 Stud. jur., und 1 „Fräulein".

[4] Für eine ausführliche Beschreibung mufs auf das Original verwiesen werden.

suchspersonen die Aufgabe gestellt, Buchstaben zu zählen. Jede Versuchsperson erhielt ein Exemplar desselben Buches mit der Weisung, von einem bezeichneten Absatz an mit gröfstmöglicher Geschwindigkeit Wort für Wort und Zeile für Zeile die Buchstaben zu zählen. Bei der Verwertung der Ergebnisse wurde nur die Geschwindigkeit, nicht die Genauigkeit des Zählens, in Betracht gezogen.

a) In der einen Versuchsreihe wurden die aufeinander folgenden Buchstaben jeder einzeln gezählt. Jedesmal wenn 100 gezählt waren, wurde an der betreffenden Stelle ein Bleistiftzeichen angebracht und wieder mit 1 angefangen.

b) Es stellte sich als unmöglich heraus, zu zählen, ohne die Zahlen, wenn auch nur in Gedanken, auszusprechen. Um die dadurch bedingte Verzögerung nach Möglichkeit zu eliminieren, hat Oehrn in einer zweiten Versuchsreihe das Buchstabenzählen in der Weise modifiziert, dafs er Gruppen von 3 Buchstaben zählen liefs. Jedes Bleistiftzeichen bezeichnete also 100 Gruppen zu je 3 Buchstaben.

B. „Zum Studium des Gedächtnisses" hat Oehrn seine Versuchspersonen

a) sinnlose Silbenreihen lernen lassen. Das Lernen geschah in der Weise, dafs die Reihen so lange von Anfang bis zu Ende durchgelesen wurden, bis sie einmal fehlerlos hergesagt werden konnten. Bei jeder unvollständigen Reproduktion wurde der fehlende Teil der Reihe noch einmal durchgelesen. Dann wurde die Anzahl der in je 5 Minuten gelernten Silben berechnet.

b) Eine andere Versuchsreihe wurde in ganz ähnlicher Weise durchgeführt, aber mit Zahlen- statt mit Silbenreihen.

C. „Der Assoziationsvorgang." Zu dessen Untersuchung wurde das Addieren einstelliger Zahlenreihen gewählt. Die Zahlenreihen waren in Hefte eingetragen. Die Versuchspersonen wurden angewiesen, mit gröfstmöglicher Geschwindigkeit zu addieren. Die der Versuchsperson zum Bewufstsein kommenden Fehler mufsten korrigiert werden; auf eine sonstige Kontrolle der Genauigkeit wurde (wegen technischer Schwierigkeiten) verzichtet.

D. „Motorische Funktionen."

a) Schreiben nach Diktat. Es wurde ein Abschnitt eines deutschen Buches von Oehrn diktiert und von der Versuchsperson mit gröfstmöglicher Schnelligkeit geschrieben. Bei der Verwertung der Ergebnisse wurde nur die Schnelligkeit berücksichtigt.

b) Lesen. Zu diesen Versuchen benutzte OEHRN FREYTAGS „Bilder aus der deutschen Vergangenheit", die sehr deutlich und mit deutschen Lettern gedruckt waren. Es wurde maximale Schnelligkeit des Lesens gefordert; dabei durften aber die Silben nicht verschluckt, sondern mufsten vollständig ausgesprochen werden (halblaut oder mit Flüsterstimme). Alle 5 Minuten, wenn das Signal gegeben wurde, mufste die betreffende Stelle bezeichnet werden.

2. Die Korrelationshöhe bei Übung und Ermüdung.

OEHRN selbst hat seine experimentell gewonnenen Werte mit ungewöhnlicher Sorgfalt und Ausführlichkeit behandelt. Da er jedoch in der herkömmlichen Weise ohne rechnerische Korrelationsbestimmungen verfuhr, so konnte er nur zu den allgemeinsten Schlüssen gelangen. Sein ganzer Befund in bezug auf die Zusammenhänge zwischen den verschiedenen Leistungen beschränkt sich auf folgende Sätze:

„Wir sehen, dafs die untersuchten Funktionen sich in bezug auf die absolute Dauer in zwei Gruppen scheiden: während im Buchstabenzählen, Addieren und in den motorischen Funktionen die Leistungen der einzelnen Versuchspersonen einander ziemlich proportional sind, zeigen dieselben im Auswendiglernen ein wesentlich anderes Verhalten. Wenn wir die Versuchspersonen nach dem Quantum der geleisteten Arbeit ordnen, so ist in ersteren Funktionen die Reihenfolge derselben mit nur geringen Abweichungen eingehalten, während dieselbe im Auswendiglernen eine ganz andere wird." [1]

Versuchen wir jetzt, mit Hilfe der Korrelations- und Ergänzungsformel, dieselben experimentellen Daten ausgiebiger zu verwerten.

Vor allem wollen wir daraus Gewinn ziehen, dafs OEHRN die Prüfung jeder Person nach jeder der oben beschriebenen Methoden zwei Stunden lang ununterbrochen fortgesetzt hat.

Auf S. 134—135 seiner Arbeit sind die Ergebnisse für jede

[1] Psychol. Arbeiten, I, S. 146.

Tabelle VIII.

Rangordnungen für jede Viertelstunde, nach Oehrns Ergebnissen.

Versuchspn.	1. Viertelstunde					2. Viertelstunde					3. Viertelstunde					4. Viertelstunde				
	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen
A	4	3	5	6	4	4	2	5	6	4	4½	2	6	6	7	6½	2	7	7	7
B	8	7	6	10	5	8	8	7	10	8	8	9	7	10	9	8	9	6	10	9
C	5	6	10	9	10	6½	5	10	9	9	7	5	10	9	8	6½	5	10	8	8
D	2	9	7	1	2	2	10	6	1	2	1½	10	5	1	2	1½	10	5	1	2
E	8	10	9	8	3	9	9	9	8	3	9	8	9	8	3	9	8	9	9	3
F	6	2	2	2	1	6½	3	2	2	1	6	3	2	2	1	4	4	3	2	1
G	3	4	4	3	8	3	4	1	4	6	3	4	1	5	4	3	3	1	5	6
H	1	8	1	4	9	1	7	4	3	10	1½	7	3	4	10	1½	6	4	4	10
I	8	5	3	5	7	5	6	3	5	7	4½	6	4	3	6	5	7	2	3	4
J	10	1	8	7	6	10	1	8	7	5	10	1	8	7	5	10	1	8	6	5

Versuchspn.	5. Viertelstunde					6. Viertelstunde					7. Viertelstunde					8. Viertelstunde				
	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen	Zählen	Ausw. lern.	Addieren	Schreiben	Lesen
A	5	2	7	6	6	3	1	7	6	6	3½	2	6	7	6	6	1	7	6	6
B	8	10	6	10	8	8	9	6	10	9	8	10	7	10	8	8	10	6	10	8
C	7	4	10	8	10	7	6	10	9	7	5½	4	10	8	10	3½	5	10	9	9
D	1	8	5	1	2	1½	10	5	1	2	1	8	5	1	3	1	9	5	1	2
E	9	7	9	9	3	9	8	9	8	3	9½	9	8	9	2	9	6	9	8	3
F	3	3	4	2	1	5	3	4	2	1	7	3	3	2	1	6	4	2	3	1
G	5	5	1	4	5	5	5	1	5	5	5½	5	1	4	4	6	5	1	4	4
H	2	6	2	5	9	1½	4	2	4	10	2	6	2	6	9	2	3	4	5	10
I	5	9	3	3	7	5	7	3	3	8	3½	7	4	3	7	3½	8	3	2	7
J	10	1	8	7	4	10	2	8	7	4	9½	1	9	5	5	10	2	8	7	5

Viertelstunde der Prüfung wiedergegeben.[1] Diese reellen Werte haben wir, wie bei unseren eigenen Versuchen (S. 71), zunächst in Ordnungszahlen übergeführt; damit bekamen wir die Rangordnung aller Versuchspersonen für jede Viertelstunde in jedem Leistungsgebiete. Sodann haben wir die zwei Ordnungszahlen jeder Versuchsperson für das Zählen einzelner Buchstaben, bzw. das Zählen nach Gruppen, zusammen addiert; die so erhaltenen Summen sind dann wiederum in die entsprechenden Ordnungszahlen übergeführt worden; auf diese Weise gewannen wir eine einzige Rangordnung der Versuchspersonen für das Zählen von Buchstaben überhaupt. Die Rangordnungen für das Lernen der Silbenreihen haben wir zunächst beiseite gelassen, da Oehrn nach dieser Methode nur sechs Viertelstunden, nicht acht, wie nach allen anderen, gearbeitet hat. Wir gewannen also schließlich für jede der acht Viertelstunden fünf Rangordnungen, und zwar je für folgende Leistungen: das Zählen von Buchstaben, das Auswendiglernen von Zahlenreihen, das Addieren, das Schreiben und das Lesen; sie sind in Tabelle VIII aufgeführt.

Zwischen diesen fünf Rangordnungen haben wir dann alle Korrelationen für jede Viertelstunde berechnet, und gelangten — Tabelle IX — zu den gesuchten Korrelationskoeffizienten

Tabelle IX.

Viertelstündige (rohe) Korrelationen nach den Ergebnissen von Oehrn.

Verglichene Rangordnungen (Tab. VIII)	Korrelationskoeffizienten 1.	2.	3.	4.	5.	6.	7.	8. Viertelstunde
Schreiben u. Addieren	+ 0,50	+ 0,68	+ 0,72	+ 0,65	+ 0,64	+ 0,68	+ 0,55	+ 0,68
„ „ Zählen	+ 0,58	+ 0,67	+ 0,70	+ 0,75	+ 0,81	+ 0,71	+ 0,54	+ 0,58
„ „ Lesen	+ 0,32	+ 0,42	+ 0,51	+ 0,53	+ 0,48	+ 0,38	+ 0,42	+ 0,47
„ „ Ausw. lern.	+ 0,10	— 0,02	— 0,03	— 0,03	+ 0,03	+ 0,02	+ 0,25	— 0,08
Addieren „ Zählen	+ 0,37	+ 0,56	+ 0,69	+ 0,67	+ 0,64	+ 0,59	+ 0,50	+ 0,31
„ „ Lesen	+ 0,01	+ 0,14	+ 0,24	+ 0,18	+ 0,05	— 0,18	+ 0,22	+ 0,26
„ „ Ausw. lern.	+ 0,22	+ 0,24	— 0,09	— 0,02	— 0,26	0,00	— 0,07	— 0,13
Zählen „ Lesen	— 0,17	— 0,16	— 0,04	+ 0,05	+ 0,14	— 0,10	— 0,26	— 0,21
„ „ Ausw. lern.	— 0,24	— 0,22	— 0,27	— 0,23	— 0,15	— 0,02	— 0,16	— 0,21
Lesen „ Ausw. lern.	— 0,05	+ 0,07	+ 0,08	— 0,10	+ 0,19	+ 0,05	+ 0,03	— 0,10

[1] Die dort mitgeteilten Zahlen sind zwar nicht selbst die ursprünglichen; sie sind vielmehr durch eine arithmetische Operation wesentlich verändert worden. Aber diese Operation läßt sich einfach, obwohl mühsam, rückgängig machen; damit gewinnen wir alle Werte, wie sie sich ursprünglich aus den Versuchen müssen ergeben haben.

nebst wahrscheinlichem Fehler. Da die OEHRNschen, wie unsere, Versuchspersonen wenig zahlreich waren, so werden wir wiederum die Gröfse des wahrscheinlichen Fehlers beständig im Auge behalten müssen, um uns nicht mit Details abzugeben, die leicht blofs zufällig sein könnten.

Der Anschaulichkeit halber haben wir dieselben Ergebnisse auch graphisch dargestellt (Fig. 1). Es fällt sofort auf, dafs die

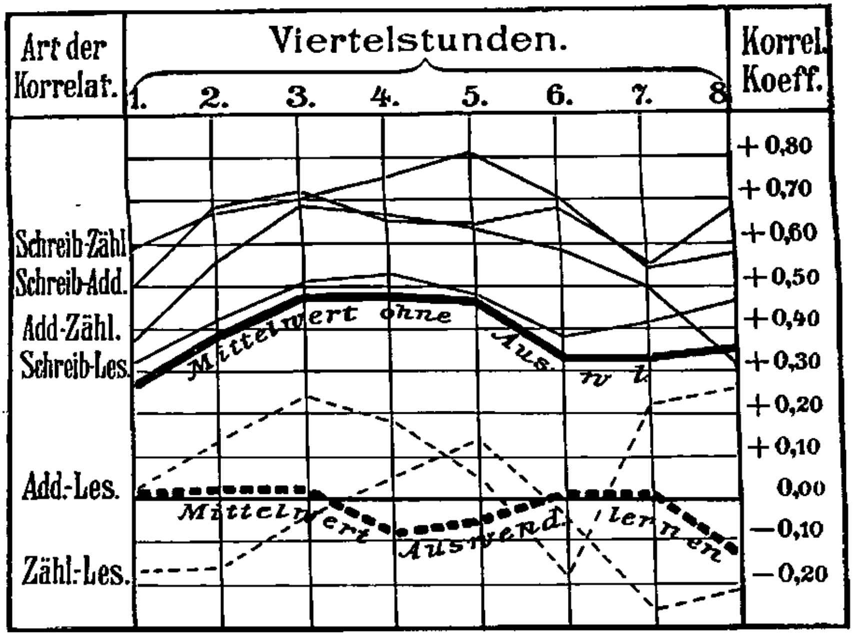

Fig. 1.

Lernversuche wiederum verschwindend kleine Korrelationen ergeben; die dicke unterbrochene Kurve, welche die Mittelwerte sämtlicher das Lernen als Glied enthaltender Korrelationen darstellt, weicht nirgends erheblich von der Nullinie, also von vollkommener Unabhängigkeit, ab.

Von den drei weiteren Korrelationen, in welche das Lesen mit eingeht, sind zwei durchweg noch nicht zweimal gröfser als der wahrscheinliche Fehler; die dritte überschreitet zwar diese Grenze, bleibt aber hinter den drei noch übrigen erheblich zurück.

Die drei noch übrigen Korrelationen schliefslich, welche nur das Schreiben, das Addieren und das Zählen als Glieder enthalten, haben sehr hohe Werte und betragen sämtlich über das Fünffache des wahrscheinlichen Fehlers.

Trotz dieser Verschiedenheiten der Korrelationskurven in bezug auf absolute Höhe, ist doch ihr allgemeiner relativer Verlauf von Viertelstunde zu Viertelstunde überall ein sehr ähnlicher (abgesehen natürlich von der Lernkurve). Durchweg ergibt die zweite Viertelstunde höhere Korrelationen als die erste, und die dritte Viertelstunde wiederum höhere als die zweite. Ferner tritt stets in der sechsten, oder spätestens der siebenten, Viertelstunde eine deutliche Herabsetzung der Korrelationshöhe ein, um dann am Schlusse fast immer einem zweiten Aufsteigen Platz zu machen. Im ganzen finden wir also in den Korrelationswerten die wohlbekannte Form der Übungs-Ermüdungskurve wieder.[1]

Binet hat, wie schon gesagt, die Meinung ausgesprochen, daſs derartige Korrelationen nur bei ungewohnten Versuchsbedingungen deutlich hervortreten; mit zunehmender Übung sollen sie rasch kleiner werden, manchmal sogar verschwinden. Der eine von uns wurde jedoch zu dem entgegengesetzten Schlusse geführt, daſs bei genauer Untersuchungsmethode die Übung (zum mindesten in ihren ersten Stadien) die Korrelationen sogar vergröſsert.[2]

Die soeben mitgeteilten Ergebnisse sprechen für die letzte Ansicht. Mit zunehmender Gewöhnung und Übung werden die Korrelationen ausnahmslos gröſser. Die Ungewohntheit ist demnach so weit davon entfernt, die Korrelationen zu verursachen, daſs sie sie im Gegenteil, direkt oder indirekt, beträchtlich herabzusetzen scheint.

Ein möglicher Grund für diese Herabsetzung liegt sehr nahe. Es pflegen nämlich bekanntermaſsen die Ergebnisse bei noch ungewohnten Versuchsbedingungen mit gröſseren zufälligen Fehlern behaftet zu sein; und dadurch wird, wie wir gesehen haben, eine Korrelation vermindert. Aber solche Verminderung läſst sich durch die „Ergänzungsformel" genau messen: sie zeigt sich im

---

[1] Man sieht, daſs hier, wie sonst, die ganz kleinen Korrelationskoeffizienten unregelmäſsiger sind, als die gröſseren (dementsprechend haben auch die ersteren, nach der Formel, einen gröſseren wahrscheinlichen Fehler).

[2] *Am. J. Psych.* 15, 1904, S. 278.

gegenwärtigen Falle als bei weitem nicht ausreichend, die beobachteten Größenunterschiede zu erklären.[1]

Wir müssen also, wie es scheint, die Ungewohntheit als einen direkt störenden Faktor auffassen. D. h. die eine Versuchsperson leidet mehr darunter als die andere, und dadurch werden die Leistungswerte von der sonst bestehenden Korrelation zwischen den betreffenden Fähigkeiten etwas abgelenkt.

„In der zweiten Stunde", bemerkt Oehrn,[2] „machte sich die Ermüdung meist schon deutlich geltend". Hierin hätten wir also einen störenden Faktor von ganz analoger Wirkung, wie die Ungewohntheit, da auch jener die verschiedenen Versuchspersonen ungleich beeinflußt. Er tritt auch in den Kurven klar genug zutage.

In der letzten Viertelstunde darf man den seit Kräpelin wohlbekannten „Antrieb" erwarten, wodurch die Ermüdung teilweise überwunden wird. Und dementsprechend steigen die Korrelationen in der Tat wieder.

Schließlich wollen wir die Kurve für die Lernversuche etwas näher betrachten. Trotz ihrer geringen Abweichungen von der Nullinie, weist sie doch immerhin eine kaum verkennbare Regelmäßigkeit auf, indem sie merkwürdigerweise der typischen Übungs-Ermüdungskurve geradezu entgegengesetzt verläuft. Diese paradoxe Erscheinung liefert aber genauer gesehen nur eine neue Bestätigung des oben Gesagten. Denn gesetzt — gemäß allen bisherigen zuverlässigen Ergebnissen — daß die Korrelation zwischen dem Lernen und den anderen Fähigkeiten

[1] So sehen wir z. B., daß die Korrelation zwischen dem Addieren und dem Zählen für die erste Halbstunde 0,50, für die zweite 0,69 beträgt. Nun ist aber der Zuverlässigkeitskoeffizient für das Addieren, bzw. das Zählen, in der ersten Halbstunde 0,88, bzw. 0,91; in der zweiten Halbstunde 0,95, bzw. 0,93. Setzen wir diese Werte in die Ergänzungsformel ein, so bekommen wir als völlig ergänzte Korrelation zwischen dem Addieren und dem Zählen, für die erste Halbstunde:

$$\frac{0{,}50}{M(0{,}88,\ 0{,}91)} = 0{,}56,$$

für die zweite Halbstunde:

$$\frac{0{,}69}{M(0{,}95,\ 0{,}93)} = 0{,}73.$$

Es hat also ein gewisser, aber verhältnismäßig geringer Ausgleich zwischen den beiden Werten stattgefunden.

[2] *Psychol. Arb.* 1, S. 110.

im allgemeinen annähernd gleich Null sei, so wird trotzdem die tatsächliche Korrelation für irgendwelche kleine zufällige als Repräsentanten der ganzen Klasse herausgegriffene Gruppen von Personen offenbar doch noch eine gewisse Gröfse, und zwar beinahe ebenso oft im negativen wie im positiven Sinne erreichen können. Dann mufs diese Korrelation durch unzugehörige Einflüsse der Versuchsbedingungen, also durch Ermüdung und Ungewohntheit gestört werden; und dies geschieht, wenn die Korrelation negativ ist, ebenso sehr, als wenn sie positiv ausfällt. Im ersteren Falle aber mufs die Übungs-Ermüdungskurve etwas nach der negativen (unteren) Seite der Nullinie hin, und zwar mit umgekehrten Hebungen und Senkungen auftreten, genau wie es in der Figur 1 geschieht.[1]

Dieser durchweg gesetzmäfsige Verlauf der Korrelationskoeffizienten selbst nach den Oehrnschen Ergebnissen, wo der Verfasser von solchen Koeffizienten gar nichts wufste, dürfte wohl jeden Zweifel über die Anwendbarkeit dieser Berechnungen (bei sonstiger angemessener Handhabung) auf so kurze Versuchsreihen beheben.

3. Anwendung der Ergänzungsformel.

Um die Ergänzungsformel anwenden zu können, bedürfen wir zweier Messungsreihen für jedes Leistungsgebiet. Diese lassen sich am einfachsten dadurch gewinnen, dafs wir die mittleren Ergebnisse der ersten Hälfte der Versuche als die ersten Messungsreihen ansehen, und ebenso die mittleren Ergebnisse der letzten Hälfte der Versuche als die zweiten Messungsreihen. Dadurch bekommt man die Rangordnungen, die in Tabelle X wiedergegeben sind.[2] Und aus diesen gewinnt man dann die Korrelationskoeffizienten, deren Mittelwerte in Tabelle XI enthalten sind.

---

[1] Dafs die Kurve für die Lernversuche viel regelmäfsiger als die zwei anderen Kurven verläuft, die sich in der Nähe der Nullinie befinden, erklärt sich leicht; denn diese stellen Einzel-, jene aber Mittelwerte dar.

[2] Diese Rangordnungen sind in ganz analoger Weise berechnet, wie diejenigen für unsere eigenen Ergebnisse (s. S. 71). Der Vollständigkeit halber haben wir die Ergebnisse für das Lernen von Zahlen und von Silben zusammengeschlossen, in derselben Weise, wie wir es für das Zählen zu 1 und zu 3 Buchstaben getan haben (S. 94).

Tabelle X.

Rangordnungen für die zwei Versuchshälften nach OEHRNS Ergebnissen.

Versuchsp.	Zählen von Buchstaben		Addieren		Schreiben		Lesen		Auswendig-lernen	
	I	II	I	II	I	II	I	II	I	II
A	7	8	5	$4^1/_2$	5	$4^1/_2$	5	$6^1/_2$	2	2
B	2	$2^1/_2$	4	$4^1/_2$	1	1	3	4	8	9
C	5	4	1	1	2	2	2	$1^1/_2$	5	$4^1/_2$
D	10	10	6	6	10	10	9	9	10	9
E	3	$2^1/_2$	2	2	3	3	8	8	9	9
F	6	$6^1/_2$	9	8	9	9	10	10	3	3
G	8	$6^1/_2$	10	10	7	7	6	5	4	$4^1/_2$
H	9	9	7	7	8	7	1	$1^1/_2$	7	6
I	4	5	8	9	6	7	4	3	6	7
J	1	1	3	3	4	$4^1/_2$	7	$6^1/_2$	1	1

Tabelle XI.

Mittlere auf Tab. X gegründete (rohe) Korrelationen.

Verglichene Rangordnungen			Durchschnittl. Korrelations-koeffizient	Wahrscheinliche Fehler*
Schreiben	und	Addieren	+ 0,70	± 0,09
„	„	Zählen	+ 0,64	0,09
„	„	Lesen	+ 0,41	0,16
„	„	Auswendiglernen	— 0,11	0,21
Addieren	„	Zählen	+ 0,46	0,14
„	„	Lesen	+ 0,06	0,21
„	„	Auswendiglernen	— 0,03	0,21
Zählen	„	Lesen	0,00	0,21
„	„	Auswendiglernen	+ 0,11	0,21
Lesen	„	Auswendiglernen	+ 0,02	0,21

* Für die Berechnungsweise des w. F. s. die Anm. Tab. IV (S. 75).

Es ergibt sich offenbar ein ganz ähnliches allgemeines Bild wie bei unseren eigenen Versuchen. Es sind sechs Korrelationen (die klein gedruckten) nicht über zweimal gröſser als ihr wahrscheinlicher Fehler, können also vernachlässigt werden. Zwei Korrelationen stehen etwas über der so gewählten Grenze. Und zwei wieder sind sehr groſs. Der Durchschnittswert der vier in

Betracht zu ziehenden Korrelationswerte beträgt 0,55. Wenn wir die zufälligen Fehler nach derselben theoretischen Methode wie früher vollständig eliminieren [1], so steigt dieser Wert auf 0,62.[2]

Man beachte, dafs bei so zuverlässigen Werten wie den Oehrnschen die Ergänzungsformel keine beträchtliche Änderung des Korrelationskoeffizienten herbeiführt. Aber dennoch ist die Formel auch in diesen Fällen nötig, und zwar eben um nachzuweisen, dafs eine solche genügende Zuverlässigkeit erreicht worden ist.

Wenn wir nun nähere Vergleiche zwischen den Oehrnschen und unseren Ergebnissen anstellen wollen, so finden wir nur eine Korrelation, die den beiden Untersuchungen gemeinsam ist. Es ist diejenige zwischen dem Addieren und dem Auswendiglernen; und beidemal hat sich auch in der Tat derselbe Korrelationswert ergeben, nämlich Null.

Gemeinsam sind ferner zwei von den „Zuverlässigkeitskoeffizienten", die für das Addieren und für das Auswendiglernen. Für das Addieren bei den Oehrnschen Versuchen beträgt dieser Wert 0,88,[3] bei unseren nur 0,76. Dies würde darauf hindeuten, dafs unsere Prüfungsmethode weniger zuverlässig als die Oehrnsche wäre. Der Grund dieser Verschiedenheit läfst sich ohne weiteres ermitteln. Denn es gab eigentlich nur einen einzigen wesentlichen Unterschied zwischen den beiden Methoden; nämlich den, dafs bei unseren Versuchen, aufser der Geschwindigkeit, auch noch die Genauigkeit des Addierens mit in die Rechnung hineingezogen wurde, während Oehrn ausschliefslich die Geschwindigkeit berücksichtigt hat. Aber nichts hindert uns, auch bei unseren Ergebnissen ausschliefslich die Geschwindigkeitsmessungen in die Rechnung eingehen zu lassen, und damit jeden Grund zu ent-

[1] Dazu benutzen wir wiederum die vereinfachte Formel (d) (s. S. 77).

[2] Wie früher (S. 78) lassen sich diese Fehler auch faktisch, aber dann nur unvollständig eliminieren. Zu diesem Zwecke bestimmen wir jedesmal den Mittelwert aus den beiden Messungen für dieselbe Versuchsperson in demselben Leistungsgebiete. Zwischen den so entstandenen Messungsreihen berechnet man wiederum die Korrelationen. Dann steigt der Mittelwert der oben in Betracht gezogenen Korrelationen tatsächlich auf 0,61, liegt also wiederum zwischen dem ganz unergänzten und dem vollständig ergänzten Wert, wie es die Theorie erfordert.

[3] Als vergleichbarster Wert erschien die Korrelation zwischen einer Leistungsfähigkeit in der ersten Viertelstunde und derselben Leistungsfähigkeit in der zweiten Viertelstunde.

fernen, weshalb die eine Methode zuverlässiger als die andere sein sollte; und jetzt beträgt unserer Zuverlässigkeitskoeffizient 0,89[1], ist also in der Tat fast identisch mit dem von OEHRN.

Interessanter ist der Vergleich zwischen den zwei Methoden für Auswendiglernen. Denn das OEHRNsche Verfahren gründete sich auf die Zahl von Wiederholungen der Eindrücke, die nötig waren, um eine durchweg richtige Reproduktion zu ermöglichen; unseren Versuchen dagegen lag der Grad der Richtigkeit der Reproduktion nach einmaliger Einwirkung zugrunde. Trotz dieser Verschiedenheit der zwei Methoden war der Zuverlässigkeitskoeffizient in beiden Fällen beinahe derselbe, und zwar beidemal sehr hoch; bei OEHRN 0,95, bei uns 0,92.

Der Zuverlässigkeitskoeffizient für die OEHRNschen Silbenreihen war im Gegenteil sehr gering, nur 0,49. Gerade durch Berechnung der Zuverlässigkeitskoeffizienten für OEHRNS Ergebnisse, konnten wir die viel gröfsere Zuverlässigkeit der Methode mit Zahlenreihen von vornherein konstatieren, und diese demnach für unsere eigenen Versuche wählen. Silbenreihen, trotz ihrer anderweitigen bekannten Vorteile, hätten uns ebenso zweideutige Ergebnisse geliefert, wie wir sie für den Raumsinn tatsächlich erhalten haben; dann wäre unsere ganze Untersuchung verfehlt gewesen (es waren nämlich unzweideutige Korrelationswerte für mindestens vier Leistungsgebiete unbedingt erforderlich). Man sieht also, welche faktische Bedeutung der Zuverlässigkeitskoeffizient haben kann.

4. Die „Zentralwerte" nach den OEHRNschen Ergebnissen.

Wenn wir nun die OEHRNschen Korrelationen in die beschriebene hierarchische Anordnung (S. 86) zu bringen versuchen, so gelingt dies sofort. Wir bekommen, indem wir die Zuverlässigkeitskoeffizienten mit heranziehen (*Kursivschrift*),

---

[1] Die Zuverlässigkeit ist demnach gröfser ohne als mit Berücksichtigung der Genauigkeit. Dies liegt daran, dafs die Genauigkeit durch zu wenig zahlreiche Daten bestimmt werden mufste. Wenn wir die Genauigkeitsmessungen allein betrachten, so zeigen sie den sehr niedrigen Zuverlässigkeitskoeffizienten von 0,43, also ungefähr denselben wie die Raumschwellenbestimmungen.

Tabelle XII.[1]

	Schreiben	Addieren	Zählen	Lesen	Ausw. lernen
Schreiben	0,92	0,69	0,63	0,44	0,00
Addieren	0,69	0,95	0,45	0,00	0,00
Zählen	0,63	0,45	0,83	0,00	0,00
Lesen	0,44	0,00	0,00	0,85	0,00
Ausw. lernen	0,00	0,00	0,00	0,00	0,87

Man sieht dafs jeder Wert (aufser den Nullwerten) gröfser als alle nach rechts oder nach unten stehenden ist.

Hier haben wir Gelegenheit zu sehen, wie unmöglich eine solche hierarchische Anordnung wird, sobald zwei Fähigkeiten eine Verwandtschaft aufser dem gemeinsamen Zentralfaktor haben. Denn OEHRN hat, wie gesagt, die Fähigkeit des Zählens auf zwei Weisen geprüft: erstens Additionen von aufeinander folgenden Einzelbuchstaben, und zweitens Additionen von je drei Buchstaben. Gestützt auf Verschiedenheiten, die sich in sehr sorgfältiger Selbstbeobachtung ergaben, hat er diese zwei Arten von Leistungen durchgängig ebenso getrennt voneinander wie von den anderen Leistungsarten behandelt; zusammen geworfen worden sind sie erst von uns. Wenn wir sie jetzt wieder trennen, so wird die hierarchische Anordnung sofort unmöglich; wir bekommen

Tabelle XIII.

	Schreiben	Addieren	Zählen 1	Zählen 3	Lesen	Ausw. lernen
Schreiben	0,92	0,69	0,66	0,59	0,44	0,00
Addieren	0,69	0,95	0,56	0,35	0,00	0,00
Zählen 1	0,66	0,56	0,93	**0,74**	0,00	0,00
Zählen 3	0,59	0,35	**0,74**	0,89	0,00	0,00
Lesen	0,44	0,00	0,00	0,00	0,85	0,00
Ausw. lernen	0,00	0,00	0,00	0,00	0,00	0,87

[1] Dabei bedienen wir uns wiederum nur derjenigen Korrelationskoeffizienten, bei denen das eine Glied von der ersten Hälfte, das andere Glied von der zweiten Hälfte der Ergebnisse herrührt (es fallen also weg die Korrelationswerte, bei welchen beide Glieder von der ersten, oder beide von der zweiten Hälfte stammen). Die dadurch erlangten Werte weichen ersichtlich sehr wenig von denen in Tabelle X ab, sind aber immerhin strenger vergleichbar mit den Zuverlässigkeitskoeffizienten. Diese letzteren bilden wieder die diagonale Wertreihe und kontrastieren stark gegen den Verlauf der anderen Korrelationen.

Das Buchstabenzählen zu 1 und das zu 3 zeigen unter sich eine Korrelation, die offenbar viel zu hoch ist, um in die allgemeine Hierarchie zu passen. Der Überschufs über das hinaus, was mit der Hierarchie vereinbar ist, weist auf einen speziellen Zusammenhang hin, der hier zu dem Zentralfaktor noch hinzukommt.

Dasselbe ergibt sich, und zwar noch auffallender, wenn wir — wie Oehrn — das Lernen von Silben und das von Zahlen getrennt behandeln. Denn dann bekommen wir Tabelle XIV, wo die Korrelation zwischen dem Silben- und dem Zahlenlernen aufser allem Verhältnis ist zu ihrer Stelle in der Tabelle.

Die zwei Arten des Auswendiglernens weisen, so wenig Korrelation sie mit anderen Leistungen haben, doch miteinander einen Korrelationswert auf, der schon unergänzt den hohen Betrag von 0,85 hat. Nach theoretischer Ergänzung steht dieser Korrelationswert nicht merklich hinter Vollkommenheit zurück. Dieser so enge spezielle Zusammenhang zwischen dem Zahlen- und dem Silbenlernen deutet auf die Möglichkeit hin, dafs vielleicht eine ziemlich grofse Gruppe von Leistungen nahe genug verwandt sind, um als eine mehr oder weniger einheitliche Leistungsfähigkeit unter den Begriff des „Auswendiglernens“ zusammengefafst werden zu dürfen.[1]

Tabelle XIV.

	Schreiben	Addieren	Zählen	Lesen	Ausw. lern. Silb.	Ausw. lern. Zahl.
Schreiben	*0,92*	0,69	0,63	0,44	0,00	0,00
Addieren	0,69	*0,96*	0,45	0,00	0,00	0,00
Zählen	0,63	0,45	*0,88*	0,00	0.00	0,00
Lesen	0,44	0,00	0,00	*0,85*	0,00	0,00
Ausw.lern. Silb.	0,00	0,00	0,00	0,00	*0,47*	**0,73**
Ausw.lern. Zahl.	0,00	0,00	0,00	0,00	**0,73**	*0,86*

Gewöhnlich konstatiert man im Gegenteil eine überraschende Abwesenheit jedes speziellen Zusammenhanges zwischen zwei Arten von Leistungen, selbst wenn diese zunächst einander sehr ähnlich erscheinen mögen.

---

[1] Merkwürdigerweise sind die meisten Psychologen im Gegenteil gerade auf diesem Gebiete sehr geneigt — ohne sich dabei auf quantitative experimentelle Ergebnisse stützen zu können — überall voneinander unabhängige Fähigkeiten oder „Spezialgedächtnisse“ anzunehmen.

Oehrn z. B. hat das schnelle Schreiben und das schnelle Lesen, auf Grund sehr sorgfältiger Selbstbeobachtung, als beidesteils „motorische Funktionen“ eng zusammengefaſst. Infolgedessen zog er aus der Tatsache, daſs beide Leistungen einen deutlichen Zusammenhang mit dem Addieren ergaben, den allgemeinen Schluſs, daſs „die motorischen Funktionen“ und „der Assoziationsvorgang“ „einander ziemlich proportional“ seien. Aber wenn man noch einmal Tabelle XIII ins Auge faſst, so bemerkt man, daſs das Schreiben gerade seine allerkleinste Korrelation mit dem Lesen hat (wir lassen das Auswendiglernen natürlich auſser Betracht); und, was noch wichtiger ist, es fällt die Korrelation zwischen dem Schreiben und dem Lesen keineswegs aus der Hierarchie heraus, wie das z. B. bei Zählen 1 und Zählen 3 der Fall ist. Dadurch daſs irgend zwei Vorgänge unter denselben Oberbegriff, wie „motorische Funktion“ u. dgl., fallen, ist man noch lange nicht berechtigt, einen entsprechenden Zusammenhang ihres tatsächlichen Verlaufs anzunehmen.

Nunmehr wollen wir feststellen, inwieweit die aus den Oehrnschen Versuchen zu gewinnenden „Zentralwerte“ mit unseren eigenen übereinstimmen. Von diesen haben wir ja gesehen, daſs sie — nach der Hypothese — unter den verschiedensten Umständen immer konstant bleiben sollten.

Es sind zwei Leistungsfähigkeiten beiden Untersuchungen gemeinsam, und liefern also vergleichbare Werte: nämlich das Auswendiglernen und das Addieren. In bezug auf das erstere haben wir ohne weiteres die erwünschte Übereinstimmung; denn aus beiden Untersuchungen ergibt sich ein Zentralwert = 0.

In bezug auf das Addieren dagegen bekommen wir von den Oehrnschen Ergebnissen, Tabelle XII, nach Formel (g) (S. 88).

$$\mathbf{Ad\,Z} = \frac{M\ (Ad\,S,\ Ad\,Z\ddot{a})}{M\ (Ad_1\ Ad_2,\ S\,Z\ddot{a})} = \frac{M\ (0{,}69,\ 0{,}45)}{M\ (0{,}95,\ 0{,}63)} = 0{,}72.$$

Hier haben wir zunächst eine bedenkliche Abweichung (0,21) von dem durch unsere eigenen Versuche gefundenen Zentralwert für das Addieren, der nämlich 0,93 betrug. Aber wir müssen uns wieder erinnern, daſs Oehrn ausschlieſslich die Geschwindigkeit, wir dagegen sowohl die Geschwindigkeit wie auch die Genauigkeit des Addierens berücksichtigten. Die Geschwindigkeit aber und die Genauigkeit einer Funktion sind zwei verschiedene Sachen. Der soeben angestellte Vergleich, den wir gemacht haben, ist also ganz unzulässig.

Wenn wir nun den Vergleich rein gestalten, indem wir auch bei unseren Versuchen ausschliefslich die Geschwindigkeitswerte in die Rechnung bringen, so bekommen wir

$$\mathbf{Ag\,Z} = \frac{M\ (Ag\,T,\ Ag\,K)}{M\ (Ag_1\ Ag_2,\ TK)} = \frac{M\ (0{,}50,\ 0{,}67)}{M\ (0{,}89,\ 0{,}64)} = 0{,}77.$$

Dieser Wert stimmt ersichtlich mit dem für die OEHRNschen Ergebnisse aufs beste.

V. Deutung aller Ergebnisse.

Wir haben jetzt heranzutreten an die wichtige und schwere Frage der Deutung aller dieser zunächst nur quantitativen Verhältnisse. Es handelt sich um die Ermittlung der gemeinsamen Ursache der festgestellten Korrelationen, oder in unserer Terminologie, um die nähere Bestimmung der Qualität des „Zentralfaktors".

Durch unsere Ergebnisse scheint es jedenfalls möglich, wenigstens eine Reihe von naheliegenden und teilweise bereits aufgestellten Erklärungen definitiv zu verwerfen.

So hat z. B. ein erfahrener Psychologe und Pädagoge in einem Briefe an den einen von uns dieser Ursache oder dem Zentralfaktor folgende Zusammensetzung zugeschrieben: „die Fähigkeit, Instruktionen zu erfassen, aufmerksam und ehrgeizig zu arbeiten, und alle verfügbaren Hilfen auszunutzen." Diese, obwohl recht naheliegende Ansicht wird ganz unhaltbar angesichts unserer Erfahrung, dafs das Schreiben nach Diktat, das Addieren, die Unterscheidung von Tonhöhen hohe Korrelationen miteinander und mit den übrigen Leistungen haben, dafs dagegen das Auswendiglernen keine merkliche Korrelation mit irgend einer der anderen Leistungen aufweist.[1] Denn es ist unmöglich zu

[1] Es sei daran erinnert, dafs ein derartiges Auswendiglernen auch keine Korrelation mit der Rangordnung in der Schule ergibt (vgl. namentlich EBBINGHAUS, *Zeitschr. f. Psychol.*, Bd. 13, S. 430). Daraus läfst sich aber nur schliefsen, dafs die Merkmale, einerseits „Auswendiglernen" und andererseits „Korrelation mit anderen Fähigkeiten", nicht zusammenhängen, keineswegs dafs sie sich ausschliefsen. Im Gegenteil könnte man wahr-

behaupten, dafs der Eifer, das schnelle Erfassen der Versuchsbedingungen, das Ausnutzen von Hilfen, und vor allem die Spannung der Aufmerksamkeit bei allen anderen Leistungen mafsgebend seien, und nur beim Auswendiglernen nicht merklich in Anspruch genommen würden! Alle bisherigen Forscher haben übereinstimmend vielmehr das Gegenteil gefunden. Nach OEHRN selbst sind z. B. beim Diktatschreiben die zentralen Prozesse dermafsen eingeübt, dafs sie „meist auch ohne Beteiligung der Aufmerksamkeit gleichsam reflektorisch vonstatten gehen."[1] Und weiter: „Was nun endlich das Auswendiglernen anbetrifft, so bedarf es wohl keiner ausführlichen Begründung, dafs dasselbe mehr als alle übrigen Funktionen eine hochgradige Anspannung der Aufmerksamkeit beansprucht."[2] Um die Zahlen- und Silbenreihen „zu behalten, müssen wir mit gespanntester Aufmerksamkeit zu erfassen suchen, was Auge und Ohr uns zuführen".[3] Das „Erfassen der Instruktionen" und das „Ausnutzen der Hilfen" spielen ebenfalls ganz sicher eine gröfsere Rolle beim Auswendiglernen, als beim Schreiben oder bei der Tonunterscheidung.

Auch BINET nimmt an, dafs alle solche einfachen Leistungen, wie die, worum es sich hier handelt, als „Prüfungen der willkürlichen Aufmerksamkeit" zu betrachten seien.[4] Er fügt jedoch hinzu, dafs er das Wort „Aufmerksamkeit" nicht ganz in der herkömmlichen Weise fasse, sondern vielmehr darunter diejenige Funktion verstehe, wodurch wir uns unter ungewohnten Bedingungen zurecht finden. Aber auch in diesem neuen Sinne des Wortes ist die Erklärung ebenso unvereinbar mit den oben besprochenen tatsächlichen Ergebnissen. OEHRN erwähnt gelegentlich, dafs die erste Prüfung aller Versuchspersonen in dem Zählen von Buchstaben bestand; es war im Laufe dieser Prüfung, sagt er, dafs „Gewöhnung an die ganze Art und Weise der Ver-

---

scheinlich in Versuche über Auswendiglernen andersartige Bedingungen leicht einführen, so dafs die Leistungen auch hier mit denen in den anderen Gebieten erhebliche Korrelationen aufwiesen.

[1] Psychol. Arbeiten, I, S. 120.

[2] a. a. O., S. 122.

[3] a. a. O., S. 123.

[4] *Année psychologique* 6, 1899. Er bezeichnet sie als „les expériences d'attention volontaire", S. 395. Er sagt, dafs er sie benutzt „pour mesurer la force de l'attention", S. 240.

suchsanordnung“ gewonnen wurde;[1] aber trotzdem ergab diese Leistung Korrelationen von nur mittlerer Gröfse. OEHRN teilt ferner mit: „Es macht sich beim Schreiben die Übung am wenigsten geltend, am meisten beim Zahlenlernen“;[2] und gerade das Schreiben wies die gröfsten, das Lernen die kleinsten Korrelationen auf. Aber, was die Hauptsache ist, wir haben tatsächlich konstatieren können, dafs die Korrelationen, statt mit zunehmender Gewöhnung und Übung sich zu vermindern, im Gegenteil auf jedem Leistungsgebiete sogar bedeutend wuchsen. Nach alledem ist, wie wir sahen, notwendig die Ungewohntheit nicht als die Ursache der Korrelationen, sondern vielmehr als ein sie störender Faktor anzusehen.

Wir haben ferner nachgewiesen, dafs die momentane „Disponiertheit“ der Versuchspersonen ebenfalls kein wesentlicher Faktor sein kann, da die Korrelationen zwischen irgend zwei Fähigkeiten nicht merklich kleiner wurden, wenn die Prüfung der einen eine Woche nach der der anderen stattfand.

Im ganzen haben wir jeden Grund anzunehmen, dafs der gesuchte Zentralfaktor zu den betreffenden Leistungen kein blofs akzidentielles Verhältnis besitzt, sondern vielmehr mit ihnen in engster funktioneller Verbindung steht.

Viel schwieriger, als diese Verwerfung der bisher versuchten oder naheliegenden Erklärungen der Korrelationen, ist die Aufstellung einer neuen stichhaltigeren Deutung.

Die anscheinend hauptsächlichsten positiven Hinweise sind die zwei folgenden: Erstens ergibt sich der merkwürdige Gegensatz zwischen dem Neuherstellen von einigen willkürlichen Zahlenassoziationen — Auswendiglernen — einerseits, wo der Zentralfaktor so gut wie keinen Einflufs zeigt, und andererseits dem Funktionieren von altgelernten und komplex verknüpften Zahlenassoziationen — Addieren —, wo der Zentralfaktor zu dominieren scheint. Zweitens fällt die überraschende psychologische Heterogeneität der Leistungen auf, die doch den engsten funktionellen Zusammenhang offenbart haben: wir fanden sehr grofse „Zentralwerte“ sowohl bei der sog. sensoriellen Leistung der Tonunterscheidung, wie bei der motorischen, beinahe reflektorischen Leistung des Schnellschreibens, und ebenfalls bei dem geistig so viel höher stehenden Erfolg in den Schulstudien.

[1] *Psychol. Arb.* I, S. 110.

[2] *Psychol. Arb.* I, S. 138.

Es sei ferner erwähnt, dafs in einigen früheren Versuchen[1] 24 Schüler nach ihrem Ruf der „Klugheit" unter ihren Mitschülern klassifiziert wurden. Ein Schüler wurde ausgewählt und gefragt: „Wen hältst du für den Klügsten deiner Kameraden?" Sodann: „Abgesehen von diesem, wer ist dann der Klügste?". Und so fort, bis eine vollständige Rangordnung sich ergeben hatte. Zur Kontrolle wurde ein zweiter Schüler in ganz derselben Weise befragt. Ebenso eine Dame, der alle betreffenden Kinder genau bekannt waren. Die auf diese Weise hergestellte Rangordnung wies eine Korrelation mit dem Zentralfaktor auf, die nicht viel hinter Vollkommenheit zurückblieb.

Versuchen wir diesen spärlichen Andeutungen nachzugehen. Der aufserordentlich hohe Zentralwert der „Klugheit" weist zwar darauf hin, dafs der Zentralfaktor stark beteiligt sein mufs an dem Rufe, den eine Person unter ihren Bekannten hinsichtlich ihrer Verstandesschärfe geniefst. Dagegen aber beweist der ebenfalls hohe Zentralwert der halb reflexartigen Schreibleistung, dafs der Zentralfaktor der wahren Intelligenz im höheren Sinne des Wortes doch recht fern steht. Sodann scheint die grofse psychische Heterogeneität der unter sich korrelationierten Leistungen zwingend darauf zu deuten, dafs der gesuchte Zentralfaktor, zunächst wenigstens, nicht als rein psychisch, sondern vielmehr als p s y c h o - p h y s i o l o g i s c h zu betrachten ist.[2] Es drängt sich die Vermutung auf, dafs irgend eine allgemeine funktionelle Qualität des Nervengewebes hier zugrunde liege. Die Wirkungsweise dieser allgemeinen Qualität liefse sich vielleicht näher als eine „p l a s t i s c h e F u n k t i o n" auffassen. Ein Nervensystem von gesteigerter plastischer Funktion würde nicht dadurch ausgezeichnet sein, dafs seine Leitungsbahnen prompter in beliebige neue Verbindungen eintreten könnten, — was etwa zur blofs rascheren Bildung irgendwelcher zufälliger Assoziationen erforderlich wäre (z. B. beim Auswendiglernen sinnloser Reihen). Wohl aber würde es imstande sein, auf allen psychophysiologischen Gebieten mit der Zeit feinere und dauerhaftere Leitungskomplexe auszugestalten, und dementsprechend präziser und konstanter (im Sinne systematischer Regelmäfsigkeit) zu funktionieren, — was namentlich in einer gröfseren Geschwindigkeit und zugleich Genauigkeit der

[1] *Am. J. Psych.* 15, 1904, S. 51.

[2] Damit soll ganz dahingestellt bleiben, ob dieser Zentralfaktor eventuell auch eine rein psychische Umdeutung zulasse oder nicht.

normalen sehr eingeübten Leistungsfähigkeiten zur Geltung käme. Ein Nervensystem, dessen Ausbildung durch eine gesteigerte plastische Funktion begünstigt wäre, würde sich in seinen Leistungen vor anderen Nervensystemen auf analoge Weise auszeichnen, wie etwa eine Maschine aus Stahl vor einer ähnlichen aus Eisen.

Am schwierigsten vielleicht fügt sich in den Rahmen dieser Hypothese der hohe Zentralwert der Unterscheidungsfähigkeit für Tonhöhen. Man mufs sich aber erinnern, dafs gerade der Vorgang, auf Grund dessen der eine von zwei Tönen jeweils als „tiefer“ oder „höher“ beurteilt wird, weitaus komplizierter ist, als er zuerst erscheinen mag.

Jedenfalls stellen wir die soeben angedeutete Hypothese nur mit der gröfsten Reserve auf; das Beobachtungsmaterial reicht noch lange nicht aus, um solche fundamentale Fragen entscheidend beantworten zu können. Die Hauptsache an dieser Arbeit sind uns die festgestellten Tatsachen. Wenn wir uns erlaubt haben, auch noch unsere vorläufige Vermutung über ihre Deutung zum Ausdruck zu bringen, so hat dies nicht sowohl den Zweck der theoretischen Formulierung, als vielmehr den der Anregung.

VI. Hauptresultate.

I. Die Leistungsfähigkeiten irgend einer Person in zahlreichen sehr verschiedenen Richtungen (Unterscheidung von Tonhöhen, Addieren von Zahlen, Ausfüllung von lückenhaften Texten, Geschwindigkeit des Schreibens, des Lesens und des Zählens) weisen hohe und konstante Korrelation untereinander auf. Auch wird eine solche Korrelation nicht merklich vermindert, wenn die eine der verglichenen Fähigkeiten von dem einen Versuchsleiter, die andere dagegen eine Woche später — und zwar ohne jede Kenntnis der ersteren Ergebnisse — von einem anderen Versuchsleiter (nach derselben Methode) geprüft wird.

II. Nach den numerischen Verhältnissen aller dieser Korrelationen, scheint man berechtigt zu sein, sie als Wirkungen eines gemeinsamen „Zentralfaktors“ aufzufassen.

III. Wenn man die Korrelationen zwischen irgend drei Leistungsfähigkeiten ermittelt hat, so ist man imstande, die Korrelation jeder dieser Fähigkeiten mit dem genannten theoretischen Zentralfaktor zu berechnen; diese Korrelation haben wir als den „Zentralwert“ der betreffenden Leistung bezeichnet. Dieser Zentralwert scheint tatsächlich, so weit unsere Erfahrungen reichen, für jede Leistung konstant zu bleiben, auch dann, wenn diese von anderen Experimentatoren bei anderen Gelegenheiten mit ganz anderen Leistungen verglichen wird.

IV. Mehrere naheliegende (und teilweise gelegentlich aufgeworfene) Erklärungen dieses Zentralfaktors haben sich als ohne allen Zweifel hinfällig erwiesen. Der Zentralfaktor läſst sich nämlich keinesfalls auf individuelle Differenzen der Versuchspersonen hinsichtlich ihres Eifers oder ihrer momentanen Disponiertheit, oder ihrer Gewöhnungsfähigkeit an die Versuchsbedingungen, oder ihrer Fähigkeit, nebenher gegebene Hilfen auszunutzen, noch selbst auf die verschieden hohe Spannung ihrer Aufmerksamkeit zurückführen.

V. Die Erklärung scheint vielmehr, zunächst wenigstens, psychophysiologisch erfolgen zu müssen. Die bisher gesammelten Erfahrungen deuten möglicherweise darauf hin, daſs das eine Nervensystem allgemein eine gesteigerte plastische Funktion besitzt gegenüber dem anderen. Diese funktionelle Tüchtigkeit wäre die Bedingung für die Ausgestaltung von präziser und konstanter funktionierenden Leitungskomplexen, was sich dann auf den verschiedensten psychophysiologischen Gebieten in einer gröſseren Genauigkeit und zugleich Geschwindigkeit der Leistung geltend machen würde. Diese Hypothese wird jedoch nur mit der gröſsten Reserve (von uns) aufgestellt, hauptsächlich als Anregung zu weiteren psychologischen und biologischen Untersuchungen.

VI. Um überhaupt eindeutige Korrelationswerte zu gewinnen, ist es unbedingt notwendig, für jedes der beiden Merkmale, deren Korrelation unter sich festgestellt werden soll, jeden Fall mindestens zweimal zu prüfen. Die sich daraus ergebende Korrelation zwischen den zwei Messungsreihen für eine und dieselbe objektive Reihe von Fällen des betreffenden Merkmals hat ferner eine ganz allgemeine Bedeutung: sie dient als ein „Zuverlässigkeitskoeffizient“ der Prüfungsmethode.

I. Anhang. Beispiel unserer Protokolle.

Protokoll.

Versuchsperson A. Versuchsleiter: KRUEGER.

1. Fragen an die Versuchsperson:
   Alter: *31; verheiratet.*
   Musik, rezeptiv: *mittelbegabt.*
   „ ausübend: *Gesangsdilettant.*
   Mathematik, Fachmann: *Als Schüler ausgezeichnete Neigung dazu.*
   Mathematik, spezielle Übung: —.
   Gesundheit, allgemeine: *gut, ein wenig nervös;* momentane: *gut disponiert.*
   Ermüdung, körperliche: —; geistige: —
   Letzter Schlaf: *gut.*
   Zerstreuung, emotionelle: —; intellektuelle: —
   Zwischenzeit seit Beendigung der letzten Mahlzeit: $1^1/_2$ *Stunden.*
   Raucher: stark, schwach[1], nicht, Trinker, Temperenzler[1]: *sehr mäfsig,* } *sehr regelmäfsiges Leben.*

2. Tonunterscheidung. Anfangszeit: 3^{15}.

Schwingungsunterschiede	1. Ton höher	2. Ton höher
30	/	/
20	//	
10	///	//
7	/////	//
6	////	//
5	////	////
4	////	//
3	/////	//
2	—0/—/	0—

Schwelle = $2^1/_2$ *Schwingungen.*

Schlufszeit: 3^{30}.

Ausländer 3. Kombinationsmethode (Vorversuch bis zur Klarheit. Hauptversuch 4 Min.).[1]

---

[1] Das hier im Drucke Unterstrichene war in unserem Protokolle durchgestrichen.

4. Raumschwellen. Anfangszeit: 3^{35}.

Linkes Jochbein			Rechte Hand			Rechtes Jochbein		
mm	2 Spitzen	1 Spitze	mm	2 Spitzen	1 Spitze	mm	2 Spitzen	1 Spitze
18	///	//	*20*	////	/—/	*18*	////	/
16	/////——/—/	//////	*15*	////	//	*16*	///	///
14	———	/	*13*	///	///	*14*	/—/——0	////
	Schwelle = 17		*10*	////	/—//		*Schwelle = 15*	
			8	—//———	/—0			
				Schwelle = 9.				

Schlufszeit: 3^{51}

5. Addieren (Vorversuch 1 Min.; 2 Hauptversuche, jeder 3 Min.).
6. Auswendiglernen (6, 8, 10 und 12 Ziffern, jede Stufe dreimal).
7. Besprechung.
   Frage (zur Kontrolle): ob Versuchsperson schon dergleichen Versuche gemacht hat, und wann).
   NB.: Keine Übung bis zur nächsten Stunde; kein Besprechen der Versuche mit zukünftigen Versuchspersonen.
8. Notanda.
   Tag: *20. I. 04.*
   Temperatur des Zimmers: *15½*; im Freien: *—2°*.
   Feuchtigkeit der Luft: *74.*
   Eifer der Versuchsperson: *grofs.*

II. Anhang. Beispiel der Kombinationstexte.

Belagerung Kolbergs. 1807.

Da der Feind fortf*uhr,* an *der* neuen Schanze am Sandwege *mit* angestr*engtem* Eifer zu .... so hatte unser neuer Kommandant gleich *in der* ersten Nacht *seines* Hierseins einen Aus*fall auf* dieselbe angeordnet, der von *einem* Trupp Grenadiere und *Jäger*, etwa hundert *Mann* stark, in mög*lichster* Stille untern*ommen* wurde. Ich schlofs mich *dem* Zuge *von* zwei in *der* Vorstadt aufgegr*iffenen* Wagen an, *um* nötigenfalls unsre T..... .....

und V...... .... aufn*ehmen* zu können. Die Überru*mpelung* erfolgte mit *gefälltem* Bajonett *und* Sturmschritt, und es lag nur *daran*, dafs die Sch...... ...... noch nicht geschlossen *war*, wenn es *der* Besatzung gelang, bis *auf* wenige *Gefan*gene zu entkommen. Wir selbst *hatten* fast keinen Verl*ust*, erbeu*teten* aber vieles Arbeitszeug, das *dazu* benutzt wurde, *den* Aufwurf *mög*lichst wieder *zu* zer*stören*, und dann *auf* meine Wagen gel*aden* und in die Festung *gebracht* wurde. Unter unseren Gefan*genen* *befand* sich ein Mensch, den anfänglich niemand in *seinem* veränderten Rocke er*kannte*, bis ich mich endlich *auf* seine *uns* nur zu wohl bekannten Gesichts*züge* besann. Es war ...... Unterof...... ......, der ...... etwa sechs Wochen, als eines heim .... Einverständnisses höchst ...., zum Feinde üb...... .... war. Ich mufs gestehen, ... mir wegen ...... ehrlosen Bu...... seither nicht wenig bange .... war. Er kannte jeden Zu...... zu unserer F...... ...... und verstand einiges vom Fortifikations...... ......, daher ...... jetzt bei den Fran...... ...... die Aufs...... bei Erbauung ...... Schanze am Sandwege .... hatte.

III. Anhang. Berechnung des Bravaisschen Korrelationskoeffizienten. Formel für den wahrscheinlichen Fehler.

Die Berechnung kann zwischen irgend zwei paarweise unter sich zugeordneten Wertreihen geschehen. Im untenstehenden Beispiele sind diese zwei Wertreihen Rangordnungen und zwar diejenigen, welche unsere Versuchspersonen bei den zwei Prüfungen ihrer Fähigkeiten zum Addieren einnahmen; sie sind in den Vertikalreihen 2 und 4 wiedergegeben. Zuerst berechnet man den Durchschnitt für jede der beiden Reihen sowie die Einzelabweichungen von demselben; diese Abweichungen, welche als x bzw. y bezeichnet werden mögen, sind in den Vertikalreihen 3 bzw. 5 aufgeführt. Es werden jetzt x^2, y^2 und xy berechnet und summiert (Vertikalreihen 6, 7 und 8). Dann ist $r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$, wo Σ das Ergebnis der Summation, und r

den gesuchten Korrelationskoeffizienten bedeutet. In unserem Beispiel ist also

$$r = \frac{83}{\sqrt{110 \times 109^1/_2}} = 0{,}75.$$

1	2	3		4	5		6	7	8	
Versuchsperson	Rangordnung nach K	x oder Abweichungen vom Durchschnitt nach K		Rangordnung nach S	y oder Abweichungen vom Durchschnitt nach S		x^2	y^2	xy	
		+	−		+	−	+	+	+	−
A	7	1		10	4		1	16	4	
B	4		2	$7^1/_2$	$1^1/_2$		4	$2^1/_4$		3
C	10	4		9	3		16	9	12	
D	1		5	1		5	25	25	25	
E	6	0		$7^1/_2$	$1^1/_2$		0	$2^1/_4$	0	
F	9	3		5		1	9	1		3
G	11	5		11	5		25	25	25	
H	3		3	2		4	9	16	12	
I	2		4	4		2	16	4	8	
J	5		1	3		3	1	9	3	
K	8	2		6	0		4	0	0	
							110 $= \Sigma x^2$	$109^1/_2$ $= \Sigma y^2$	89	6
									83 $= \Sigma xy$	

Der wahrscheinliche Fehler des Korrelationskoeffizienten wird nach folgender Formel berechnet:

$$\text{w. F.} = 0{,}6745 \frac{1 - r^2}{\sqrt{n\,(1 + r^2)}},[1]$$

also in unserem Beispiele

$$= 0{,}6745 \frac{1 - 0{,}75^2}{\sqrt{11\,(1 + 0{,}75^2)}} = 0{,}07.$$

[1] Siehe aber S. 74.

(Eingegangen am 8. September 1906.)