



Leseprobe aus: Eid, Gollwitzer, Schmitt, Statistik und Forschungsmethoden, ISBN 978-3-621-28201-7  
© 2015 Beltz Verlag, Weinheim Basel  
<http://www.beltz.de/de/nc/verlagsgruppe-beltz/gesamtprogramm.html?isbn=978-3-621-28201-7>

## 14 Unterschiede zwischen mehreren abhängigen Stichproben: Varianzanalyse mit Messwiederholung und verwandte Verfahren

### Was Sie in diesem Kapitel lernen

- ▶ Wie testet man, ob Bedingungsmittelwerte bei Experimenten mit intraindividuellem Bedingungsvariation signifikant voneinander abweichen?
- ▶ In welche Bestandteile wird bei der Varianzanalyse mit Messwiederholung die Gesamtvariation zerlegt?
- ▶ Wie können bei der Varianzanalyse mit Messwiederholung spezifische Hypothesen über Mittelwertsunterschiede getestet werden?
- ▶ Wie werden Interaktionseffekte zwischen zwei messwiederholten Faktoren getestet?
- ▶ Wie werden Interaktionseffekte zwischen einem messwiederholten und einem nicht-messwiederholten Faktor getestet?
- ▶ Wie testet man auf der Basis eines nonparametrischen Tests, ob die Mediane mehrerer messwiederholter Bedingungen signifikant voneinander abweichen?

In Kapitel 13 haben wir die Varianzanalyse als Möglichkeit kennengelernt, Mittelwertsunterschiede aus mehreren unabhängigen Stichproben auf ihre statistische Bedeutsamkeit hin zu testen. Die Variation in den Messwerten, die ein Faktor (bzw. mehrere Faktoren und ihre Wechselwirkung) verursacht, haben wir als Effekte (Haupteffekte, Interaktionseffekte) bezeichnet und gesehen, dass solche Effekte auf der Basis eines  $F$ -Tests inferenzstatistisch abgesichert werden können. Wir haben festgestellt, dass eine notwendige Voraussetzung für die Anwendung des in Kapitel 13 beschriebenen  $F$ -Tests darin besteht, dass es zwischen Messwerten, die aus unterschiedlichen Stichproben (d.h. aus unterschiedlichen Stufen des Faktors) stammen, keinerlei gegenseitigen Abhängigkeiten gibt. Es muss sich also bei den Stufen eines Faktors jeweils um *unabhängige* Stichproben handeln: In den unterschiedlichen Faktorstufen müssen sich unterschiedliche Personen befinden.

In diesem Kapitel werden wir nun sehen, wie die Varianzanalyse funktioniert, wenn es sich bei den Faktorstufen um *abhängige* Stichproben handelt. Was mit abhängigen Stichproben gemeint ist (und was nicht mit ihnen gemeint ist), haben wir bereits in Kapitel 12 ausführlich behandelt. Abhängige Stichproben liegen z. B. vor, wenn es sich um Experimente mit intraindividuellem Bedingungsvariation handelt (sog. messwiederholte Faktoren) oder wenn unterschiedliche Versuchspersonen in den unterschiedlichen Faktorstufen einander zugeordnet werden können (aufgrund einer »natürlichen Beziehung« zwischen ihnen oder aufgrund gleicher Ausprägungen auf einer Kontrollvariablen; vgl. die Technik des Parallelisierens in Abschn. 4.3.3). Da die wiederholte Messung an den gleichen Personen den typischen Anwendungsfall abhängiger Stichproben darstellt, wird die entsprechende Auswertungsprozedur als »Varianzanalyse mit Messwiederholung« (engl. *repeated-measures analysis of variance*, kurz: RM-ANOVA oder auch *within-subjects ANOVA*) bezeichnet. Den Fall einer einfaktoriellem Varianzanalyse mit Messwiederholung werden wir in Abschnitt 14.1 detailliert behandeln.

Im Falle von mehrfaktoriellen Designs können Messwiederholungen bei keinem, einigen oder allen Faktoren vorliegen. Sind alle Faktoren messwiederholt, spricht man von einem komplett messwiederholten Design; sind nur einige Faktoren messwiederholt, andere hingegen nicht, spricht man von einem partiell messwiederholten Design. Wir werden uns in Abschnitt 14.2 auf zwei Fälle beschränken: ein zweifaktorielles Design mit vollständiger Messwiederholung und ein zweifaktorielles Design mit Messwiederholung auf einem Faktor. Abschließend werden wir in diesem Kapitel noch einen nonparametrischen Test auf Unterschiede zwischen mehreren Medianen im Falle eines messwiederholten Faktors vorstellen (Abschn. 14.3) und auf Verfahren für kategoriale Variablen verweisen (Abschn. 14.4).

## 14.1 Einfaktorielle Varianzanalyse mit Messwiederholung

Wir haben in Abschnitt 13.1 darauf hingewiesen, dass die einfaktorielle Varianzanalyse *ohne* Messwiederholung eine Erweiterung (oder Verallgemeinerung) des *t*-Tests für unabhängige Stichproben ist (vgl. Abschn. 11.1.2). Dementsprechend ist die einfaktorielle Varianzanalyse *mit* Messwiederholung eine Erweiterung (oder Verallgemeinerung) des *t*-Tests für abhängige Stichproben (vgl. Abschn. 12.1.1). Sie wird verwendet, wenn die Nullhypothese getestet werden soll, dass sich die Mittelwerte mehrerer abhängiger Stichproben (z. B. experimenteller Bedingungen) nicht voneinander unterscheiden.

Der klassische Fall von abhängigen Stichproben ist die intraindividuelle Bedingungsvariation: Von allen Personen in der Stichprobe werden wiederholt Messwerte unter unterschiedlichen experimentellen Bedingungen erhoben. Beispielsweise könnte man die kognitive Leistungsfähigkeit von Personen dreimal erheben und miteinander vergleichen: (1) ohne Stimmungsinduktion, (2) nach positiver Stimmungsinduktion und (3) nach negativer Stimmungsinduktion. Mit einem solchen Design könnte die Hypothese überprüft werden, dass sich positive Stimmung förderlich und negative Stimmung hemmend auf die kognitive Leistungsfähigkeit auswirken. Da in allen drei experimentellen Bedingungen die gleichen Personen getestet werden, handelt es sich um abhängige Stichproben: Die Messwerte werden über die drei Bedingungen hinweg miteinander kovariieren, da die kognitive Leistungsfähigkeit eben nicht nur von der Stimmung, sondern auch von personengebundenen Variablen abhängt, die über die drei Stimmungsbedingungen hinweg stabil bleiben (Intelligenz, Teilnahmemotivation etc.).

**Kovariation zwischen den Messungen.** Ein Teil der Variation in den Messwerten ist also dadurch zu erklären, dass es sich um die gleichen Personen (und damit teilweise um die gleichen personengebundenen Einflüsse auf die Messwerte) handelt. Wir haben in Kapitel 12.1 die Kovarianz als einen statistischen Kennwert kennengelernt, der angibt, wie groß der Einfluss solcher personengebundenen Merkmale ist (s. hierzu ausführlich Abschn. 16.3.1). Genauer gesagt: Die Kovarianz gibt an, wie stabil Unterschiede zwischen Personen über zwei Messungen hinweg bleiben. Liegen mehr als zwei

Messungen vor, muss diese Stabilität anders quantifiziert werden. In Abschnitt 14.1.2 werden wir sehen, wie diese »Varianz zwischen Personen« bestimmt werden kann und welche Rolle sie für die Zerlegung der Gesamtvariation spielt. Durch die Kovariation wiederholter Messungen können stabile Störvariablen, die an die Person gebunden sind (z. B. Persönlichkeitsmerkmale), statistisch kontrolliert werden. Hierdurch reduziert sich der Anteil unerklärter Varianz. Folglich gewinnt man an Teststärke (Power) und benötigt eine geringere optimale Stichprobengröße als bei einer Varianzanalyse ohne Messwiederholung. Dies ist ein großer Vorteil von Messwiederholungsdesigns.

**Sequenzeffekte.** Ein Nachteil experimenteller Designs mit intraindividuelle Bedingungsvariation besteht darin, dass die Messwerte nicht vor systematischen Verfälschungen gefeit sind. So könnte es sein, dass Versuchspersonen im Laufe der Untersuchung die Hypothese erraten, dass ihre Motivation, am Experiment teilzunehmen, zunehmend nachlässt oder dass sie Antwort- bzw. Lösungsstrategien von einer Messung auf die nächste übertragen bzw. sich an ihre früheren Antworten erinnern. All diese systematischen Verfälschungen, die typisch für Messwiederholungsdesigns sind, werden unter dem Begriff *Sequenzeffekte* zusammengefasst. Für eine Systematik solcher Sequenzeffekte und den Umgang mit ihnen (etwa die Berücksichtigung bestimmter Kontrollstrategien wie die intra- oder interindividuelle Ausbalancierung) verweisen wir auf einschlägige Lehrbücher zur Versuchsplanung (z. B. Huber, 2013; Hussy & Jain, 2002; Shadish et al., 2002).

**Veränderungsmessung.** In vielen psychologischen Fragestellungen geht es nicht um Unterschiede, die durch die Manipulation einer experimentellen Variablen hervorgerufen werden, sondern lediglich um Unterschiede im Laufe der Zeit. Anders gesagt: Oft ist die Zeit selbst die unabhängige Variable, und man interessiert sich für die Veränderung in den (durchschnittlichen) Messwerten über die Zeit hinweg. Beispielsweise interessiert man sich in der allgemeinen Entwicklungspsychologie dafür, ob und wie sich psychologische Merkmale mit zunehmendem Alter verändern. Dies erfordert die wiederholte Messung der fraglichen Merkmale. Werden über die Zeit hinweg immer wieder Messwerte von den gleichen Personen erhoben, handelt es sich um eine sog. Längsschnittuntersuchung. Auch hier kann man sich mithilfe

einer Varianzanalyse mit Messwiederholung der Frage widmen, ob die Unterschiede in den Mittelwerten zwischen den Messzeitpunkten signifikant variieren.

**Evaluationsforschung.** Interessiert man sich im Rahmen einer Untersuchung zur Wirksamkeit einer psychologischen Intervention dafür, ob und inwieweit sich durch die Intervention die Merkmalsausprägung über die Zeit hinweg verändert hat (z. B. inwieweit durch eine Psychotherapie das Ausmaß subjektiver Belastungen reduziert wurde oder inwieweit durch ein Lerntraining die Schulleistung von Kindern verbessert wurde), so sind wiederum Messwiederholungsdesigns angezeigt. Messwiederholungsdesigns sind typisch für die Evaluationsforschung, v. a. für die Prozess- und die Wirksamkeitsevaluation (vgl. Gollwitzer & Jäger, 2014).

### Datenbeispiel

Um die Grundidee der einfaktoriellen Varianzanalyse mit Messwiederholung zu veranschaulichen, beginnen wir auch hier mit einem einfachen Datenbeispiel. Stellen wir uns vor, das in Abschnitt 13.1 geschilderte Experiment zum Modelllernen sei nicht auf der Basis einer interindividuellen Bedingungsvariation, sondern vielmehr auf der Basis einer *intraindividuellen* Bedingungsvariation durchgeführt worden. Insgesamt werden fünf Personen untersucht, und zwar jeweils unter drei Bedingungen. Zunächst sehen die Probanden einen Film, in dem eine Person ihres Alters für ein bestimmtes Verhalten (z. B. Aggression) belohnt wird. Im Anschluss daran wird mit einem Selbstberichtsmaß (z. B. einer visuellen Analogskala mit Ausprägungen zwischen 0 und 100) gemessen, wie stark die Personen dazu tendieren, das vom Modell gezeigte Verhalten nachzuahmen. Anschließend sehen die gleichen Personen einen zweiten Film, in dem die Modellperson für ihr Verhalten bestraft wird; wiederum wird die Nachahmungstendenz per Selbstauskunft gemessen. Schließlich sehen die Probanden einen dritten Film, in dem das Verhalten der Modellperson ohne Konsequenzen bleibt. Die Nachahmungstendenz wird hier ein drittes Mal gemessen.

Die unabhängige Variable (UV) ist also die Verhaltenskonsequenz (Belohnung, Bestrafung, keine Konsequenz). Die abhängige Variable (AV) ist die Nachahmungstendenz. Sollte die Hypothese, die aus der Theorie des Modelllernens von Bandura (1976, 1977) abgeleitet wurde, korrekt sein, müsste die Nachah-

mungstendenz nach dem ersten Film (Belohnungsbedingung) am stärksten und nach dem zweiten Film (Bestrafungsbedingung) am schwächsten sein. Dass es sich hier zugegebenermaßen um ein ziemlich problematisches Versuchsdesign handelt, da mit erheblichen Verzerrungen durch Sequenzeffekte zu rechnen ist, dürfte klar sein. Trotzdem werden wir mit diesem Beispiel weiterarbeiten.

Die experimentellen Bedingungen erhalten einen Laufindex von  $j = 1, \dots, J$  (mit  $J = 3$  in unserem Beispiel). Die Personen erhalten einen Laufindex von  $m = 1, \dots, n$  (mit  $n = 5$  in unserem Beispiel). Den Index  $j$  können wir bei  $n$  hier weglassen, da die Anzahl der beobachteten Werte pro Bedingung aufgrund der Messwiederholung der Gesamtanzahl aller Versuchspersonen entspricht. Die (fiktiven) Rohdaten des Versuchs stehen in Tabelle 14.1. Beachten Sie, dass es sich hier um die gleichen Rohdaten handelt wie in Tabelle 13.1 (in Abschn. 13.1). Wir haben bewusst die gleichen Daten gewählt, um den Unterschied zwischen der Varianzanalyse ohne Messwiederholung und der Varianzanalyse mit Messwiederholung (und die statistischen Implikationen dieses Unterschiedes) deutlich zu machen.

### Personen als Stufen eines »zufälligen Faktors«

Der Unterschied zwischen Tabelle 13.1 und Tabelle 14.1 besteht darin, dass wir hier eine zusätzliche Spalte eingefügt haben, in der die *durchschnittliche* Nachahmungstendenz einer Person über die drei experimentellen Bedingungen hinweg angegeben ist. In Tabelle 13.1 wäre es aufgrund des interindividuellen Designs nicht sinnvoll gewesen, einen solchen Personmittelwert zu berechnen, da den unterschiedlichen experimentellen Bedingungen unterschiedliche Personen zugeordnet waren. Beim intraindividuellen Design hingegen stammen alle Messwerte innerhalb der gleichen Zeile des Versuchsplans von ein und derselben Person. Daher ist hier nicht nur der Bedingungs­mittelwert  $\bar{x}_{\cdot j}$  (Spaltenmittelwert; s. die untere Zeile in Tab. 14.1), sondern auch der Personmittelwert  $\bar{x}_m$  (Zeilenmittelwert; s. die rechte Spalte in Tab. 14.1) sinnvoll interpretierbar.

In Abschnitt 13.1.11 hatten wir die Unterscheidung zwischen »festen« und »zufälligen« Effekten (bzw. Faktoren) kennengelernt. Bei den drei experimentellen Bedingungen in unserem Beispiel handelt es sich zweifelsohne um Stufen eines festen Faktors, denn das variierte Merkmal kann nur eine bestimmte Anzahl möglicher

Ausprägungen haben, und die realisierten Faktorstufen entsprechen genau diesen Ausprägungen. Bei Messwiederholungsdesigns kann man nun auch die unterschiedlichen Personen als »Stufen« eines »Faktors« auffassen. Der Personfaktor umfasst dabei alle möglichen Unterschiede zwischen den Personen. Da die Personen eine Zufallsstichprobe darstellen und sich daher auch zufällig hinsichtlich der Personmerkmale unterscheiden, handelt es sich formal gesehen um einen zufälligen Faktor (s. Abschn. 13.1.11). Anders als der Bedingungsfaktor ist der Personfaktor von untergeordnetem Interesse für die Hypothesenprüfung, aber die Vorstellung, dass es sich bei den Personen um Stufen eines zufälligen Faktors handelt, hilft uns, die Logik der Quadratsummenzerlegung besser zu verstehen. Wir können also die einfaktorielle Varianzanalyse mit Messwiederholung als quasi-zweifaktorielles Design mit einem (festen) Faktor »Bedingung« und einem zweiten (zufälligen) Faktor »Person« auffassen.

### 14.1.1 Messwertzerlegung

14

In Tabelle 14.1 gibt es drei Quellen der Variation: (1) Unterschiede zwischen den experimentellen Bedingungen, (2) Unterschiede zwischen den Personen und (3) den Anteil der Variation, der weder auf Haupteffekte der Bedingungen noch auf Personeneffekte zurückgeführt werden kann. Dieser Restanteil (Residualanteil) geht auf Interaktionen zwischen den Personen und den Bedingungen und andere unsystematische Störeinflüsse wie z. B. den Messfehler zurück.

Der Messwert  $x_{mj}$  einer Person  $m$  in einer Bedingung  $a_j$  lässt sich daher wie folgt zerlegen:

$$x_{mj} = \bar{x} + t_j + p_m + e_{mj} \quad (\text{F 14.1})$$

In dieser Gleichung bezeichnet  $\bar{x}$  den Gesamtmittelwert aller Werte,  $t_j = \bar{x}_{\cdot j} - \bar{x}$  ist der Haupteffekt der  $j$ -ten Stufe des Faktors  $A$ , d.h. die Abweichung des Bedingungs-mittelwerts  $\bar{x}_{\cdot j}$  vom Gesamtmittelwert  $\bar{x}$ . Mit  $p_m = \bar{x}_{m\cdot} - \bar{x}$  wird der Haupteffekt der  $m$ -ten Person bezeichnet, d.h. die Abweichung des Mittelwerts  $\bar{x}_{m\cdot}$  einer Person  $m$  (über alle  $J$  Bedingungen hinweg) vom Gesamtmittelwert  $\bar{x}$ . Der Residualwert  $e_{mj} = x_{mj} - \bar{x} - t_j - p_m$  einer Person  $m$  in Bedingung  $a_j$  ist der Anteil am Wert der Person, der übrig bleibt, wenn man den Gesamtmittelwert und die Bedingungs- und Personeneffekte abzieht.

### 14.1.2 Quadratsummenzerlegung

Die Gesamtvariation (totale Quadratsumme) der 15 Messwerte berechnet sich analog zu dem in Abschnitt 13.1.4 beschriebenen Vorgehen. Sie ist definiert als die Summe der quadrierten Abweichungen der einzelnen Messwerte vom Gesamtmittelwert:

$$QS_{\text{tot}} = \sum_{j=1}^J \sum_{m=1}^n (x_{mj} - \bar{x})^2 \quad (\text{F 14.2})$$

Sie beträgt in unserem Beispiel  $QS_{\text{tot}} = 4.532$ .

Wie kann die totale Quadratsumme bei der einfaktoriellen Varianzanalyse mit Messwiederholung zerlegt werden? Um das zu veranschaulichen, machen wir von der Vorstellung Gebrauch, es handele sich bei den unterschiedlichen Personen um Stufen eines zufälligen

**Tabelle 14.1** Fiktive Daten eines Experiments zum Modelllernen (mit intraindividuellem Bedingungsvariation)

| Person $m$  | Stufe $a_j$ des Faktors |                      |                            | Personmittelwert $\bar{x}_{m\cdot}$ |
|---|-------------------------|----------------------|----------------------------|-------------------------------------|
|   | Belohnung ( $a_1$ )     | Bestrafung ( $a_2$ ) | Keine Konsequenz ( $a_3$ ) |                                     |
| 1   | 57                      | 18                   | 36                         | 37                                  |
| 2   | 45                      | 15                   | 27                         | 29                                  |
| 3   | 49                      | 13                   | 43                         | 35                                  |
| 4   | 69                      | 37                   | 29                         | 45                                  |
| 5   | 70                      | 37                   | 55                         | 54                                  |
| <b>Bedingungs-mittelwert <math>\bar{x}_{\cdot j}</math></b> | 58                      | 24                   | 38                         | $\bar{x} = 40$                      |

Faktors. Genau wie jeder andere Faktor kann also auch jede Person einen Haupteffekt, d. h. einen von der experimentellen Bedingung unabhängigen (und insofern »unbedingten«) Effekt, haben. Die Haupteffekte der Personen sind darauf zurückzuführen, dass diese sich über die unterschiedlichen Messungen hinweg konsistent (d. h. gleichbleibend) in Personmerkmalen unterscheiden, welche die AV beeinflussen. Zusätzlich zu den Haupteffekten der Personen gibt es für jede Stufe des Faktors *A* einen Haupteffekt der experimentellen Bedingung, der von der jeweiligen Person unabhängig ist. Und schließlich kann es sein, dass der Effekt der experimentellen Manipulation bei unterschiedlichen Personen unterschiedlich ausfällt. Unklar bleibt jedoch, ob es sich hierbei um einen echten Interaktionseffekt zwischen Personmerkmalen und der experimentellen Bedingung handelt oder lediglich um unsystematische Effekte wie Messfehler.

Bei der einfaktoriellen Varianzanalyse mit Messwiederholung wird die totale Quadratsumme  $QS_{\text{tot}}$  also in drei Teile zerlegt:

- ▶ eine Quadratsumme, welche die Variation zwischen Personen angibt ( $QS_{\text{zWP}}$ ),
- ▶ eine Quadratsumme, welche die Variation zwischen den Stufen des Faktors *A* angibt ( $QS_{\text{zWA}}$ ), und
- ▶ eine Quadratsumme, welche diejenige Variation angibt, die weder durch Haupteffekte der Bedingungen noch durch Haupteffekte der Personen erklärt werden kann. Diese Quadratsumme werden wir im Folgenden als *Residualquadratsumme* ( $QS_{\text{Res}}$ ) bezeichnen.

### Variation zwischen Personen

Die Variation zwischen Personen ist derjenige Teil der Gesamtvariation, der auf Unterschiede zwischen den Personen – unabhängig von den Stufen des experimentellen Faktors – zurückgeht. Es handelt sich also um diejenigen Unterschiede zwischen Personen, die sich über alle Faktorstufen hinweg konsistent zeigen. Die Variation zwischen Personen ist ein unbedingter (d. h. von Faktor *A* unabhängiger) Effekt des »Faktors« Person.

Die Variation zwischen Personen entspricht im Falle von zwei Bedingungen der Kovarianz beim *t*-Test für abhängige Stichproben (s. Abschn. 12.1.1). Sie ist ein Maß für den Einfluss jener Merkmale, die über die experimentellen Bedingungen hinweg konsistente Unterschiede zwischen den Personen produzieren. Solche konsistenten Unterschiede zwischen Personen manifestieren sich in der Variation der Personmittelwerte (s. rechte Spalte in Tab. 14.1).

Die Zwischen-Personen-Quadratsumme  $QS_{\text{zWP}}$  basiert also auf den quadrierten Abweichungen der Personmittelwerte vom Gesamtmittelwert:

$$QS_{\text{zWP}} = J \cdot \sum_{m=1}^n (\bar{x}_{m\bullet} - \bar{x})^2 \quad (\text{F 14.3})$$

Sie beträgt in unserem Beispiel  $QS_{\text{zWP}} = 1.128$ .

### Variation zwischen Faktorstufen

Die Variation zwischen den Faktorstufen ist derjenige Teil der Gesamtvariation, der auf systematische Unterschiede zwischen den experimentellen Bedingungen zurückgeführt werden kann. Diese Unterschiede manifestieren sich – genau wie bei der einfaktoriellen Varianzanalyse ohne Messwiederholung – in der Variation der Bedingungsmittelwerte (s. untere Zeile in Tab. 14.1). Die Zwischen-Quadratsumme des Faktors *A* ( $QS_{\text{zWA}}$ ) basiert also auf den quadrierten Abweichungen der Bedingungsmittelwerte vom Gesamtmittelwert:

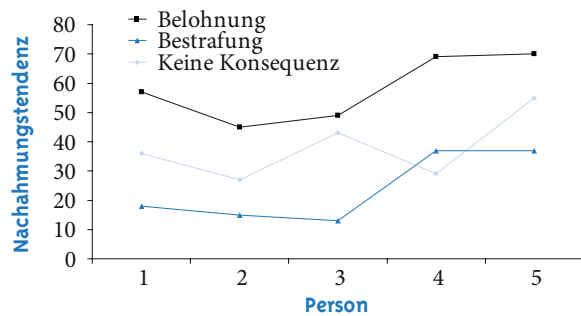
$$QS_{\text{zWA}} = n \cdot \sum_{j=1}^J (\bar{x}_{\bullet j} - \bar{x})^2 \quad (\text{F 14.4})$$

Sie beträgt in unserem Beispiel  $QS_{\text{zWA}} = 2.920$ .

### Variation zwischen Personen in Bezug auf den Effekt des Faktors

Wenn wir die Personen und die Bedingungsvariation als Faktoren auffassen, dann besteht auch die Möglichkeit, dass diese miteinander interagieren. Inhaltlich bedeutet diese Interaktion, dass sich Personen darin unterscheiden, wie groß die Unterschiede in ihren Messwerten zwischen den drei Bedingungen sind. Grafisch kann man die Idee einer solchen Interaktion zwischen Person und Bedingung veranschaulichen, wenn man die Messwerte in einem Liniendiagramm abträgt, wobei die Personen als Ausprägungen auf der Abszisse und die experimentellen Bedingungen als drei unterschiedliche Linien dargestellt werden können (s. Abb. 14.1). Man sieht deutlich, dass die Linien nicht parallel verlaufen; es gibt also (Person-)Unterschiede in den (Bedingungs-)Unterschieden.

Machen wir uns die Idee einer Interaktion zwischen Person und Bedingung an Abbildung 14.1 klar. Wir sehen z. B., dass die Messwerte von Person 2 näher beieinanderliegen als die Messwerte von Person 1. Der Effekt der experimentellen Manipulation ist bei Per-



**Abbildung 14.1** Grafische Darstellung der Interaktion zwischen Person und Bedingung (Datenbeispiel aus Tab. 14.1)

son 1 größer als bei Person 2. Möglicherweise ist dieser Unterschied zwischen den beiden Personen darauf zurückzuführen, dass Person 2 ihr Verhalten generell weniger stark an stellvertretenden Verhaltensbekräftigungen ausrichtet. Und möglicherweise ist der Effekt der experimentellen Manipulation bei Person 1 genau deshalb stärker, weil diese Person ihr Verhalten sehr stark daran ausrichtet, ob andere Menschen für das gleiche Verhalten belohnt oder bestraft werden. Diese Überlegung würde nahelegen, dass es eine echte Interaktion zwischen der experimentellen Manipulation (stellvertretende Verhaltenskonsequenzen) und Merkmalen der beobachtenden Personen (hier: Sensibilität für stellvertretende Verhaltenskonsequenzen) gibt. Das Problem ist: Wir können nicht testen, ob es sich um eine echte Interaktion handelt oder lediglich um zufällige Schwankungen bzw. Messfehler, die überhaupt nicht auf systematische Personunterschiede zurückzuführen sind. Der Grund dafür, dass wir systematische Person-Bedingungs-Interaktionen nicht testen können, ist der, dass wir in jeder Kombination von Person und Bedingung jeweils nur einen einzigen Messwert haben. Wie stark dieser Messwert von Messfehlern vs. von systematischen Person-Bedingungs-Interaktionseffekten beeinflusst ist, bleibt unbekannt. Insofern werden wir der Einfachheit halber die entsprechende Quadratsumme mit  $QS_{Res}$  bezeichnen.

Wie quantifiziert man die Residualquadratsumme  $QS_{Res}$ ? Grundsätzlich erfolgt die Berechnung nicht anders, als in Abschnitt 13.2.3 beschrieben. Es handelt sich um die Variation zwischen den Messwerten, die weder auf einen unbedingten Effekt der Person noch auf einen unbedingten Effekt der Bedingung zurückgeführt werden kann. Diese Variation manifestiert sich also in den Abweichungen der Messwerte vom Ge-

samtmittelwert, nachdem der Haupteffekt einer Person und der Haupteffekt der Bedingung  $a_j$  von diesen Abweichungen abgezogen wurden (vgl. auch Formel F 13.97a):

$$\begin{aligned}
 QS_{Res} &= \sum_{j=1}^J \sum_{m=1}^n ((x_{mj} - \bar{x}) - (\bar{x}_{\cdot j} - \bar{x}) - (\bar{x}_{m\cdot} - \bar{x}))^2 \\
 &= \sum_{j=1}^J \sum_{m=1}^n (x_{mj} - \bar{x}_{\cdot j} - \bar{x}_{m\cdot} + \bar{x})^2
 \end{aligned}
 \tag{F 14.5}$$

Sie beträgt in unserem Beispiel  $QS_{Res} = 484$ .

Die Summe aus  $QS_{zwB}$ ,  $QS_{zWA}$  und  $QS_{Res}$  entspricht der totalen Quadratsumme.

**! Additivität der Quadratsummen**

Bei der einfaktoriellen Varianzanalyse mit Messwiederholung lässt sich die totale Quadratsumme  $QS_{tot}$  in drei Teile zerlegen:

- ▶ einen Teil, der die Variation zwischen Personen ausdrückt (»Haupteffekte« der Person;  $QS_{zwP}$ ),
- ▶ einen Teil, der die Variation zwischen Bedingungen ausdrückt (Haupteffekte des Faktors A;  $QS_{zWA}$ ), und
- ▶ einen unerklärten Teil ( $QS_{Res}$ ):

$$QS_{tot} = QS_{zwP} + QS_{zWA} + QS_{Res} \tag{F 14.6}$$

**Variation zwischen und innerhalb von Personen**

Versuchen wir nun, die Quadratsummenzerlegung bei der einfaktoriellen Varianzanalyse mit Messwiederholung mit jener ohne Messwiederholung zu vergleichen. Bei der einfaktoriellen Varianzanalyse ohne Messwiederholung haben wir die totale Quadratsumme in zwei Teile zerlegt: einen, der Variation zwischen Bedingungen ( $QS_{zw}$ ), und einen, der Variation innerhalb von Bedingungen anzeigt ( $QS_{inn}$ ). Bei der einfaktoriellen Varianzanalyse mit Messwiederholung haben wir die totale Quadratsumme in drei Teile zerlegt: einen, der Variation zwischen Bedingungen ( $QS_{zWA}$ ), einen, der Variation zwischen Personen ( $QS_{zwP}$ ), und einen dritten, der unerklärte Variation anzeigt ( $QS_{Res}$ ).

**Variation zwischen Bedingungen.** Die Variation zwischen Bedingungen ist in beiden varianzanalytischen Modellen die gleiche, und sie wird auch gleich be-

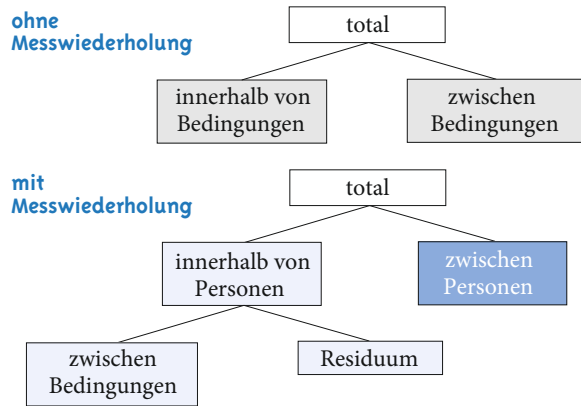
rechnet (vgl. die Formeln F 13.6b und F 14.4). Da unser Datenbeispiel in Tabelle 14.1 exakt dem in Tabelle 13.1 entspricht, resultiert für die Variation zwischen den Bedingungen in beiden Fällen der gleiche Wert, nämlich  $QS_{zWA} = 2.920$ . Das bedeutet auch, dass der Anteil der Variation, der auf die experimentelle Manipulation zurückzuführen ist (also der Anteil der  $QS_{zWA}$  an der totalen Quadratsumme), für beide Modelle gleich ist. Darauf kommen wir später zurück, wenn es um die Bestimmung der Effektgröße geht. Der einzige formale Unterschied besteht darin, dass es sich bei der  $QS_{zw}$  der Varianzanalyse ohne Messwiederholung um eine Variation *zwischen Personen* handelt, während es sich bei der  $QS_{zWA}$  der Varianzanalyse mit Messwiederholung aufgrund der intraindividuellen Bedingungsvariation um eine Variation *innerhalb von Personen* handelt.

**Variation innerhalb Bedingungen.** Bei der einfaktoriellen Varianzanalyse ohne Messwiederholung haben wir die Innerhalb-Quadratsumme wie folgt hergeleitet: Wir hatten einen Fall konstruiert, in dem es nur Unterschiede *innerhalb* der, nicht aber *zwischen* den experimentellen Bedingungen gibt (s. Abschn. 13.1.4). Die Quadratsumme dieser an ihrem jeweiligen Bedingungs-mittelwert zentrierten Messwerte betrug gemäß Formel F 13.7  $QS_{inn} = 1.612$ . Ein Vergleich mit den Quadratsummen, die wir in diesem Kapitel berechnet haben, zeigt: Dieser Wert entspricht der Summe aus der Zwischen-Personen-Quadratsumme  $QS_{zWP}$  und der Residualquadratsumme  $QS_{Res}$  ( $1.128 + 484 = 1.612$ ). Mit anderen Worten: Die  $QS_{inn}$  bei der Varianzanalyse ohne Messwiederholung entspricht der Summe aus  $QS_{zWP}$  und  $QS_{Res}$  bei der Varianzanalyse mit Messwiederholung:

$$QS_{inn} = QS_{zWP} + QS_{Res} \quad (F 14.7)$$

Bei der  $QS_{zWP}$  handelt es sich um Variation *zwischen Personen*, während es sich bei der  $QS_{Res}$  um Variation *innerhalb von Personen* handelt.

**$QS_{zWP}$  und  $QS_{innP}$ .** Während also bei der Varianzanalyse ohne Messwiederholung eine Unterscheidung in Variationsquellen zwischen den bzw. innerhalb der *Bedingungen* im Vordergrund steht, wird bei der Varianzanalyse mit Messwiederholung zunächst danach unterschieden, ob es sich um Variationsquellen zwi-



**Abbildung 14.2** Quadratsummenzerlegung bei der einfaktoriellen Varianzanalyse mit und ohne Messwiederholung

schen oder innerhalb der *Personen* handelt. Die Variation zwischen Personen drückt sich in der Zwischen-Personen-Quadratsumme  $QS_{zWP}$  aus, die Variation innerhalb von Personen in der Zwischen-Bedingungen-Quadratsumme  $QS_{zWA}$  und der Quadratsumme  $QS_{Res}$ , deren Summe die Innerhalb-Personen-Quadratsumme  $QS_{innP}$  ergibt. Sie berechnet sich in unserem Beispiel zu  $2.920 + 484 = 3.404$ . Die Quadratsummenzerlegung bei der einfaktoriellen Varianzanalyse mit und ohne Messwiederholung ist in Abbildung 14.2 grafisch veranschaulicht.

### 14.1.3 Effektgrößenmaße

Als Effektgröße bietet sich auch bei der einfaktoriellen Varianzanalyse mit Messwiederholung an, denjenigen Anteil der Varianz der Messwerte zu quantifizieren, der auf den Effekt des Faktors *A* zurückzuführen ist. Insofern stimmt die Bedeutung des Konzepts »Effektgröße« bei Varianzanalysen mit Messwiederholung genau mit der Bedeutung bei Varianzanalysen ohne Messwiederholung überein. Bei der Schätzung dieser Effektgröße aus den Daten kann das Ausmaß der stabilen Personunterschiede entweder berücksichtigt werden oder nicht. Genau das macht den Unterschied zwischen dem partiellen Effektgrößenschätzer  $\hat{\eta}_p^2$  und dem nicht-partiellen Effektgrößenschätzer  $\hat{\eta}^2$  aus. Wir werden beide Maße im Folgenden behandeln und im Anschluss daran diskutieren, ob und wann es sinnvoll ist, die Abhängigkeit der Stichproben bei der Effektgrößenschätzung mit zu berücksichtigen oder nicht.



**Nicht-partielles Effektgrößenmaß  $\hat{\eta}^2$ .** Beim sog. »nicht-partiellen« Effektgrößenmaß  $\hat{\eta}^2$  wird der Quotient aus der Quadratsumme  $QS_{zWA}$  und der totalen Quadratsumme gebildet:

$$\hat{\eta}^2 = \frac{QS_{zWA}}{QS_{tot}} = \frac{QS_{zWA}}{QS_{zWA} + QS_{zWP} + QS_{Res}} \quad (\text{F 14.8})$$

**Partielles Effektgrößenmaß  $\hat{\eta}_p^2$ .** Beim partiellen Effektgrößenmaß  $\hat{\eta}_p^2$  wird der Anteil der Gesamtvarianz, der auf stabile Unterschiede zwischen den Personen zurückgeht, nicht mit berücksichtigt. Bei der Schätzung der Effektgröße geht die  $QS_{zWP}$  also nicht mehr mit in die Gleichung ein:

$$\hat{\eta}_p^2 = \frac{QS_{zWA}}{QS_{zWA} + QS_{Res}} \quad (\text{F 14.9})$$

Da beim partiellen Effektgrößenmaß der Nenner niemals einen größeren Wert annehmen kann als beim nicht-partiellen Effektgrößenmaß, kann  $\hat{\eta}_p^2$  niemals kleiner werden als  $\hat{\eta}^2$ . Der Unterschied zwischen beiden wird umso größer, je größer der Anteil der Gesamtvariation ist, der auf stabile Personunterschiede zurückgeführt werden kann.

#### Welches Effektgrößenmaß ist informativer?

Stellen wir uns vor, das zu Beginn dieses Kapitels geschilderte Experiment zum Modelllernen mit drei experimentellen Bedingungen wäre (1) einmal mit einer »echten« Messwiederholung (intraindividuelle Bedingungsvariation: alle Personen durchlaufen alle Bedingungen des Faktors A), (2) einmal mit einer »Quasi-Messwiederholung« aufgrund einer Parallelisierung der Versuchspersonen anhand eines Vortests (interindividuelle Bedingungsvariation, wobei Personen mit gleichen Ausprägungen auf der Vortestvariablen jeweils einer der drei experimentellen Bedingungen zugewiesen werden) und (3) ein drittes Mal ohne Messwiederholung (interindividuelle Bedingungsvariation mit randomisierter Zuweisung der Personen zu einer der drei Bedingungen) durchgeführt worden. Im ersten Fall ist eine hohe Zwischen-Personen-Quadratsumme  $QS_{zWP}$  zu erwarten; im zweiten Fall dürfte die  $QS_{zWP}$  kleiner sein, da die Abhängigkeit zwischen den Messwerten hier nur noch auf eine einzige Variable (nämlich diejenige, die im Vortest gemessen wurde) zurückzuführen ist; und im dritten Fall ist die  $QS_{zWP}$  definitionsgemäß gleich 0. Die  $QS_{tot}$  wäre in allen drei Fällen die gleiche, und auch die Haupteffekte der Be-

dingungen wären in allen drei Fällen exakt identisch. Das nicht-partielle Effektgrößenmaß  $\hat{\eta}^2$  würde in allen drei Fällen also den gleichen Wert annehmen. Das partielle Effektgrößenmaß  $\hat{\eta}_p^2$  hingegen wäre im ersten Fall größer als im zweiten Fall und dort wiederum größer als im dritten Fall, weil eben die Varianz, die auf konsistente Personenunterschiede zurückgeht ( $QS_{zWP}$ ), hier nicht mit in den Ausdruck im Nenner eingeht und der Ausdruck im Nenner dementsprechend im ersten Fall kleiner ist als im zweiten Fall und dort wiederum kleiner als im dritten Fall.

Das Beispiel zeigt, dass mit dem partiellen Effektgrößenmaß  $\hat{\eta}_p^2$  ein Problem verbunden ist: Untersucht man den gleichen Effekt mit unterschiedlichen Designs (intraindividuelle vs. interindividuelle Bedingungsvariation), dann unterscheidet sich das partielle Effektgrößenmaß  $\hat{\eta}_p^2$  zwischen diesen Designs auch dann, wenn die Unterschiede zwischen den Bedingungsmitelwerten in den Designs exakt identisch sind. Unterschiedliche Untersuchungen, die zwar das Gleiche untersuchen, aber unterschiedliche Designs verwenden, sind hinsichtlich ihrer Effektgrößen nicht mehr miteinander vergleichbar. Das ist v.a. dann ein Problem, wenn man im Rahmen einer sog. Metaanalyse versucht, die Effektgrößen, die in vielen unterschiedlichen Primärstudien gefunden wurden, zusammenzufassen (Dunlap et al., 1996). Und dabei sollen standardisierte Effektgrößen ja gerade so definiert sein, dass sie auch über unterschiedliche Studien hinweg miteinander verglichen werden können.

Welches der beiden Effektgrößenmaße informativer ist, hängt auch von der konzeptuellen Bedeutung des Faktors ab, den man untersuchen will. Handelt es sich bei dem Faktor A um eine experimentell manipulierte Variable, die man sowohl mit einem intraindividuellen als auch mit einem interindividuellen Design untersuchen kann, sollte man beide Effektgrößenmaße berichten. Das nicht-partielle Effektgrößenmaß  $\hat{\eta}^2$  erlaubt dann einen besseren Vergleich mit Studien, die mit einem interindividuellen Design gearbeitet haben. Handelt es sich bei dem Faktor A hingegen um die Zeit und ist man daran interessiert, wie viel Varianz in den Messwerten durch Veränderungen in der Merkmalsausprägung über die Zeit hinweg aufgeklärt wird, so ist das partielle Effektgrößenmaß  $\hat{\eta}_p^2$  informativer, da Unterschiede in der Größe der Personeffekte für die Schätzung des Veränderungseffekts irrelevant sind.

Statistikprogramme wie SPSS geben nur das partielle Effektgrößenmaß an; will man zusätzlich das nicht-partielle Effektgrößenmaß berichten, so muss man es anhand von Formel F 14.8 selbst berechnen.

**Beispiel**

**Effektgrößenmaße für das Beispiel zum Modellernen**

Wie groß ist der Effekt der experimentellen Manipulation (Faktor A) in unserem Datenbeispiel aus Tabelle 14.1? Zunächst berechnen wir das geschätzte nicht-partielle Effektgrößenmaß nach Formel F 14.8:

$$\hat{\eta}^2 = QS_{z_{wA}} / QS_{tot} = 2.920 / 4.532 = 0,64$$

Das partielle Effektgrößenmaß beträgt nach Formel F 14.9:

$$\hat{\eta}_p^2 = QS_{z_{wA}} / (QS_{z_{wA}} + QS_{Res}) = 2.920 / 3.404 = 0,86$$

**14.1.4 Populationsmodell der einfaktoriellen Varianzanalyse mit Messwiederholung**

Das Populationsmodell der einfaktoriellen Varianzanalyse mit Messwiederholung besagt, dass ein Messwert beeinflusst wird durch

- ▶ den unbedingten Populationsmittelwert ( $\mu$ ),
- ▶ den Populationseffekt derjenigen Bedingung, unter der der Wert erhoben wurde ( $\tau_j$ ),
- ▶ Effekte, die auf Eigenschaften der jeweiligen Person zurückgehen ( $\pi_m$ ),
- ▶ den bedingten Effekt der Bedingung, gegeben eine spezifische Person (Interaktion Person  $\times$  Bedingung;  $(\pi\tau)_{mj}$ ), und
- ▶ alle unsystematischen Einflüsse einschließlich des Messfehlers ( $\varepsilon_{mj}$ ).

Formal:

$$x_{mj} = \mu + \tau_j + \pi_m + (\pi\tau)_{mj} + \varepsilon_{mj} \quad (F 14.10a)$$

Da sich die Einflussgrößen  $(\pi\tau)_{mj}$  und  $\varepsilon_{mj}$  mit diesem Design empirisch nicht trennen lassen, gehen wir im Folgenden davon aus, dass alle Interaktionseffekte  $(\pi\tau)_{mj}$  gleich 0 sind. Damit verkürzt sich Formel F 14.10a wie folgt:

$$x_{mj} = \mu + \tau_j + \pi_m + \varepsilon_{mj} \quad (F 14.10b)$$

Diese Annahme vereinfacht die weitere Ableitung des Modells.

**Haupteffekt der Bedingung  $a_j$ .** Der Koeffizient  $\tau_j$  wird – genau wie bei der einfaktoriellen Varianzanalyse ohne Messwiederholung – als Haupteffekt einer Bedingung  $a_j$  bezeichnet. Er ist definiert als die Abweichung eines Bedingungs-mittelwerts ( $\mu_{\cdot j}$ ) vom Gesamtmittelwert:

$$\tau_j = \mu_{\cdot j} - \mu \quad (F 14.11)$$

Diese Definition impliziert, dass die Summe der Haupteffekte über alle  $J$  Bedingungen hinweg immer 0 ergeben muss:

$$\sum_{j=1}^J \tau_j = 0 \quad (F 14.12)$$

Die Haupteffekte  $\tau_j$  variieren nicht über Personen hinweg; sie sind für alle Personen in der Population konstant. Insofern trägt ihre Varianz auch nichts zur Varianz der Messwerte innerhalb einer Bedingung bei.

**Haupteffekt der Person  $m$ .** Der Koeffizient  $\pi_m$  kennzeichnet den unbedingten Effekt einer Person  $m$ . Hierunter fallen alle Merkmale dieser Person, die einen Einfluss auf die abhängige Variable haben und von der experimentellen Manipulation unabhängig sind, also über die Messungen hinweg stabil bleiben. Der Haupteffekt  $\pi_m$  ist definiert als die Abweichung eines Person-mittelwerts ( $\mu_{m\cdot}$ ) vom Gesamtmittelwert:

$$\pi_m = \mu_{m\cdot} - \mu \quad (F 14.13)$$

Auch hier gilt, dass die Summe der Personen-Haupteffekte über alle  $n$  Personen hinweg immer 0 ergeben muss:

$$\sum_{m=1}^n \pi_m = 0 \quad (F 14.14)$$

Da Haupteffekte der Personen zufällige Effekte sind, lässt sich ihre Varianz ( $\sigma_\pi^2$ ) nicht von vornherein kontrollieren; vielmehr handelt es sich um einen Populationsparameter, der aus den Daten geschätzt werden muss. Wichtig ist, dass die Personen-Haupteffekte einen Teil der Unterschiede zwischen den Personen (und damit einen Teil der Varianz innerhalb einer Bedingung) erklären.

**Residuum.** Wie wir gesehen haben, setzt sich das Residuum  $\varepsilon_{mj}$  aus allen unsystematischen Einflüssen wie z. B. dem Messfehler zusammen. Es ist derjenige Teil in der Variation der Messwerte, der weder durch Bedingungs- noch durch Personeneffekte erklärt werden kann. Setzt man die Formeln F 14.11 und F 14.13 in Formel F 14.10b ein und löst nach  $\varepsilon_{mj}$  auf, ergibt sich:

$$\begin{aligned} \varepsilon_{mj} &= x_{mj} - \mu - \tau_j - \pi_m \\ &= x_{mj} - \mu - (\mu_{\bullet j} - \mu) - (\mu_{m\bullet} - \mu) \\ &= x_{mj} - \mu - \mu_{\bullet j} + \mu - \mu_{m\bullet} + \mu \\ &= x_{mj} - \mu_{\bullet j} - \mu_{m\bullet} + \mu \end{aligned} \quad (\text{F 14.15})$$

Das Residuum ist also nichts anderes als die Abweichung der einzelnen Messwerte vom Gesamtmittelwert, nachdem sowohl der Bedingungs- als auch der Personeneffekt kontrolliert (herausgerechnet) wurden. Die Varianz dieses Residuums ( $\sigma_\varepsilon^2$ ) ist ein Populationsparameter, der aus den Daten geschätzt werden muss; sie ist ein Teil der Varianz der Messwerte innerhalb der Bedingungen.

### Varianz der Messwertvariablen

Unter der Annahme, dass das Modell in Formel F 14.10b gültig ist, lässt sich die Varianz der Messwertvariablen  $X_{mj}$  anhand der siebten Rechenregel für Varianzen (vgl. Formel F 7.35) zerlegen. Die Variable  $X_{mj}$  repräsentiert die potenziellen Werte einer zufällig gezogenen Person  $m$  in einer Bedingung  $a_j$ . Der individuelle Wert  $x_{mj}$  ist die Realisierung der Variablen  $X_{mj}$  in einer konkreten Studie. Wenden wir Formel F 7.35 auf unser Problem an, erhalten wir:

$$\begin{aligned} \text{Var}(X_{mj}) &= \text{Var}(\mu + \tau_j + \pi_m + \varepsilon_{mj}) \\ &= \text{Var}(\mu) + \text{Var}(\tau_j) + \text{Var}(\pi_m) \\ &\quad + \text{Var}(\varepsilon_{mj}) + \text{Cov}(\mu, \tau_j) \\ &\quad + \text{Cov}(\mu, \pi_m) + \text{Cov}(\mu, \varepsilon_{mj}) \\ &\quad + \text{Cov}(\tau_j, \pi_m) + \text{Cov}(\tau_j, \varepsilon_{mj}) \\ &\quad + \text{Cov}(\pi_m, \varepsilon_{mj}) \end{aligned} \quad (\text{F 14.16a})$$

Der Gesamtmittelwert  $\mu$  und die Bedingungs-Haupteffekte  $\tau_j$  variieren nicht zwischen Personen; der Effekt einer Bedingung  $a_j$  ist für alle Personen in der Population gleich. Innerhalb einer Bedingung  $a_j$  gibt es also keine Varianz von  $\tau_j$ . Bei  $\tau_j$  handelt es sich somit innerhalb von  $a_j$  um eine Konstante.

Die Feststellung, dass  $\mu$  und  $\tau_j$  Konstanten sind, hat verschiedene Implikationen:

- ▶ Die beiden Ausdrücke  $\text{Var}(\mu)$  und  $\text{Var}(\tau_j)$  sind beide gleich 0.
- ▶ Alle Kovarianzen, an denen  $\mu$  und  $\tau_j$  beteiligt sind, sind gleich 0; denn wenn eine Variable eine Konstante ist, kann sie auch nicht mit anderen Variablen kovariieren (s. Abschn. 16.3.1).

Damit verkürzt sich Formel F 14.16a wie folgt:

$$\text{Var}(X_{mj}) = \text{Var}(\pi_m) + \text{Var}(\varepsilon_{mj}) + \text{Cov}(\pi_m, \varepsilon_{mj}) \quad (\text{F 14.16b})$$

### Annahmen

Bei der einfaktoriellen Varianzanalyse mit Messwiederholung werden in Bezug auf die drei Größen auf der rechten Seite von Formel F 14.16b die folgenden drei zusätzlichen Annahmen getroffen (Fahrmeir et al., 1996a):

- (1) Die zufälligen Personeneffekte  $\pi_m$  sind unabhängig und identisch normalverteilt mit  $N(0, \sigma_\pi^2)$ .
- (2) Die Residuen  $\varepsilon_{mj}$  sind unabhängig und identisch normalverteilt mit  $N(0, \sigma_\varepsilon^2)$ .
- (3) Die Kovarianz der Personeneffekte und der Residuen ist gleich 0:  $\text{Cov}(\pi_m, \varepsilon_{mj}) = 0$ .

Aus diesen Annahmen folgt, dass die Varianz von  $X_{mj}$  der Varianz  $\sigma_{X_j}^2$  entspricht mit

$$\sigma_{X_j}^2 = \sigma_\pi^2 + \sigma_\varepsilon^2. \quad (\text{F 14.17})$$

Bei Gültigkeit der Modellannahmen in der Population muss die Varianz des Merkmals in allen Faktorstufen identisch sein, nämlich  $\sigma_\pi^2 + \sigma_\varepsilon^2$ . Aufgrund der Annahme, dass die Variablen unabhängig und identisch verteilt sind, lassen wir im Folgenden bei der weiteren Betrachtung der Variablen des Populationsmodells aus Gründen der Vereinfachung den Index  $m$  für die Person weg. Mit  $X_j$  bezeichnen wir die Variable, deren Werte die individuellen Messwerte  $x_{mj}$  in einer Bedingung (bzw. einem Messzeitpunkt)  $a_j$  sind,  $\pi$  bezeichnet die Personeneffektvariable und  $\varepsilon_j$  die Residualvariable in der Bedingung  $a_j$ .

### Kovarianz zwischen den Faktorstufen

Bei Designs mit intraindividuellem Bedingungsvariatio(n) (Messwiederholung) sind die Messwerte zwischen den Faktorstufen nicht unabhängig voneinander, da sie von den gleichen Personen stammen. Das Ausmaß dieser Abhängigkeit kann über die Kovarianz quantifiziert werden. Das hatten wir bereits in Abschnitt 12.1.1

im Rahmen des  $t$ -Tests für abhängige Stichproben festgestellt, und in Abschnitt 16.3.1 werden wir die Kovarianz im Detail behandeln. Die Kovarianz ist dann positiv, wenn Individuen, die in einer Bedingung überdurchschnittliche Werte haben, auch in einer anderen Bedingung überdurchschnittliche Werte aufweisen und umgekehrt. In einem Design mit intraindividuelle Bedingungsvariation und  $J = 3$  Faktorstufen können demnach drei Kovarianzen berechnet werden:  $Cov(X_1, X_2)$ ,  $Cov(X_1, X_3)$  und  $Cov(X_2, X_3)$ . Bei intraindividuellen Designs gibt es also nicht nur Varianzen innerhalb der Faktorstufen, sondern auch Kovarianzen zwischen den Faktorstufen. Varianzen und Kovarianzen werden in Form einer Matrix (Varianz-Kovarianz-Matrix oder einfach *Kovarianzmatrix*) dargestellt. Eine Kovarianzmatrix wird mit dem griechischen Großbuchstaben  $\Sigma$  (Sigma) symbolisiert. Um deutlich zu machen, dass es sich um eine Matrix handelt, wird  $\Sigma$  fett geschrieben. Für den Fall, dass es  $J = 3$  Faktorstufen gibt, sieht die Kovarianzmatrix der Variablen  $X_j$  wie folgt aus:

$$\Sigma_X = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & Cov(X_1, X_3) \\ Cov(X_2, X_1) & Var(X_2) & Cov(X_2, X_3) \\ Cov(X_3, X_1) & Cov(X_3, X_2) & Var(X_3) \end{pmatrix} \quad (F 14.18)$$

Eine Kovarianzmatrix ist immer quadratisch und hat  $J$  Zeilen und  $J$  Spalten. In der Hauptdiagonale der Matrix stehen die Varianzen innerhalb einer jeweiligen Faktorstufe. In den restlichen Zellen der Matrix stehen die Kovarianzen zwischen zwei der  $J$  Faktorstufen.

In der einfaktoriellen Varianzanalyse mit Messwiederholung wird angenommen, dass die Residuen unterschiedlicher Faktorstufen voneinander unabhängig sind und daher alle eine Kovarianz von 0 aufweisen. Unter den Modellannahmen der einfaktoriellen Varianzanalyse mit Messwiederholung gibt es daher nur einen Grund, aus dem die Messwerte über die Bedingungen hinweg kovariieren: nämlich dass es stabile Personenunterschiede gibt. Diese sind auf die Personen-Haupteffekte  $\pi_m$  zurückzuführen; sie manifestieren sich also in der Varianz  $\sigma_\pi^2$ . Da die Personen-Haupteffekte über alle Messzeitpunkte hinweg konstant sind, müssen alle Kovarianzen gleich sein und der Personvarianz  $\sigma_\pi^2$  entsprechen. Mit dieser Feststellung und der Zerlegung der Varianzen in Formel F 14.17 lässt sich die Kovarianzmatrix in F 14.18 wie folgt reformulieren:

$$\Sigma_X = \begin{pmatrix} \sigma_\pi^2 + \sigma_\epsilon^2 & \sigma_\pi^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\epsilon^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\epsilon^2 \end{pmatrix} \quad (F 14.19)$$

**! Kovarianzstruktur des Modells der einfaktoriellen Varianzanalyse mit Messwiederholung**

Aus den genannten Annahmen der einfaktoriellen Varianzanalyse mit Messwiederholung folgt, dass alle Faktorstufen in der Population eine konstante Varianz aufweisen, die der Summe aus der Personvarianz und der Residualvarianz entspricht, und dass die Kovarianz der Messwerte zwischen zwei beliebigen Faktorstufen der Personvarianz entspricht.

**14.1.5 Schätzung der Populationsparameter**

Es gibt vier Populationsparameter, die wir aus den Stichprobendaten schätzen müssen: den unbedingten Populationsmittelwert ( $\mu$ ), den Effekt einer experimentellen Bedingung ( $\tau_j$ ), den Effekt einer Person ( $\pi_m$ ) bzw. die Personvarianz ( $\sigma_\pi^2$ ) und das Residuum  $\epsilon_{mj}$  bzw. die Populationsresidualvarianz ( $\sigma_\epsilon^2$ ).

**Unbedingter Populationsmittelwert  $\mu$ .** Den unbedingten Populationsmittelwert  $\mu$  schätzen wir wieder aus dem Gesamtmittelwert  $\bar{x}$  (s. auch Formel F 13.23):

$$\hat{\mu} = \bar{x} = \frac{\sum_{j=1}^J \sum_{m=1}^n x_{mj}}{J \cdot n} \quad (F 14.20)$$

**Bedingungs-Haupteffekte  $\tau_j$ .** Die Bedingungs-Haupteffekte  $\tau_j$  schätzen wir aus den Differenzen der Bedingungs-mittelwerte  $\bar{x}_{\cdot j}$  vom Gesamtmittelwert  $\bar{x}$  (vgl. Formel F 13.24):

$$\hat{\tau}_j = \bar{x}_{\cdot j} - \bar{x} \quad (F 14.21)$$

**Personen-Haupteffekte  $\pi_m$ .** Die Personen-Haupteffekte  $\pi_m$  schätzen wir aus den Abweichungen der Personen-mittelwerte  $\bar{x}_{m\cdot}$  vom Gesamtmittelwert  $\bar{x}$ :

$$\hat{\pi}_m = \bar{x}_{m\cdot} - \bar{x} \quad (F 14.22)$$

**Residuum  $\epsilon_{mj}$ .** Nachdem die Abweichungen  $\bar{x}_{\cdot j} - \bar{x}$  und  $\bar{x}_{m\cdot} - \bar{x}$  zur Schätzung der Populationsparameter  $\tau_j$  und  $\pi_m$  verwendet werden, bleibt nur noch eine Quelle der