

## Kapitel 6

### Planung von Experimenten in der empirischen Unterrichtsforschung

Andreas Zender

#### 6.1 Einleitung

Unterrichtsforschung ist die wissenschaftliche Disziplin, die Unterricht und die ihn bedingenden Variablen, die in ihm ablaufenden Prozesse und die Resultate (z. B. Lernzuwachs, Behaltensleistungen) wissenschaftlich erforscht (Shuell, 1996; Helmke, 2006, 2007). In dem von Dunkin und Biddle (1974) vorgeschlagenen Prozess-Produkt-Paradigma werden vier Faktoren berücksichtigt: (1) Lehrervariablen (*presage variables*) wie *fachliches* oder didaktisches Wissen, (2) Kontextvariablen (*context variables*) wie Schul- oder Klassencharakteristika, (3) Prozessvariablen (*process variables*) wie unterschiedliche methodische Aktivitäten oder Lernprozesse, und (4) Produktvariablen (*product variables*) wie aktive Lernzeit oder Lernzuwachs. Erweiterungen des Prozess-Produkt-Paradigmas adressieren grundlegende Lernbedingungen, Klassenkontexte und Mediationsprozesse (Shuell, 1996; Wellenreuther, 2007). Unterrichtsforschung auf der Grundlage des experimentellen Paradigmas (Brophy & Good, 1986; Berliner & Calfee, 1996; Graham, Harris, & Zito, 2005) als Teil der empirischen Unterrichtsforschung stellt besondere Anforderungen an die Versuchsplanung. Neben der Forderung, dass in die Versuchsplanung mehrere Faktoren (*presage, context, process variables*) einzubeziehen sind (vgl. Cronbach & Snow, 1977; O'Malley, 1988; Marjoribanks, 1988) und auch der Klassenkontext (*context variables*) zu berücksichtigen ist, wird die Verwendung von Versuchsplänen mit Messwiederholung(en) empfohlen (Lewis, 1968; Winer, 1971; Namboodiri, 1972; Klauer, 1973, 2005; Kelly & Lesh, 2000; Zender & Pfeiffer, 2009). Sie sind besonders dann indiziert, wenn die durch Unterricht bewirkten Änderungen von Schülereigenschaften (z. B. Lernzuwachs über die Zeit, Behaltensleistungen) untersucht werden sollen.

Der nächste Abschnitt befasst sich mit methodischen Anforderungen an die Unterrichtsforschung aus Sicht der Versuchsplanung und der statistischen Auswertung. Der dann folgende Abschnitt behandelt ein-, zwei- und dreifaktorielle Versuchspläne (RB-

$p$ , CR- $p$ , CRF- $p \times q$ , SPF- $p \bullet q$ , SPF- $p \times r \bullet q$ ) mit konkreten Beispielen aus der Unterrichtsforschung. Der vierte Abschnitt enthält Schlussfolgerungen.

## 6.2 Anforderungen an die Versuchsplanung

### 6.2.1 Einbeziehung mehrerer Faktoren

Diese Anforderung besagt, dass in die Versuchsplanung im Sinne des Prozess-Produkt-Paradigmas mehrere Faktoren (*presage, context, process, product variables*) einzubeziehen sind. Die Einbeziehung mehrerer Faktoren ist aus pädagogischer Sicht interessant, weil Interaktionen geprüft werden können: Ist beispielsweise die Interaktion Unterrichtsmethode  $\times$  Schülermerkmal signifikant, wirken sich unterschiedliche Unterrichtsmethoden bei Schülern mit unterschiedlichen Eigenschaften verschieden auf das Lehren und Lernen im Klassenunterricht aus (vgl. Cronbach & Snow, 1977; Marjoribanks, 1988; O'Malley, 1988).

### 6.2.2 Berücksichtigung des Klassenkontextes

Diese Anforderung berücksichtigt explizit den Faktor Schulklasse in der Versuchsplanung. Durch Einbeziehung des Faktors Schulklasse in einen zwei- oder mehrfaktoriellen Versuchsplan kann beispielsweise überprüft werden, ob eine Unterrichtsmethode in verschiedenen Schulklassen – Unterrichtsmethode  $\times$  Schulklasse – einen Effekt auf das Lehren und Lernen hat (vgl. Klauer, 1973, 2005). Ist die Interaktion signifikant, hat man nachgewiesen, dass der Effekt der Unterrichtsmethode in den Schulklassen verschieden war. In einem solchen Fall wäre eine Hypothese nicht aufrechtzuerhalten, dass eine Unterrichtsmethode allgemein überlegen ist. Andererseits meint diese Anforderung – bei entsprechender Fragestellung –, dass in die Versuchsplanung unterschiedliche Merkmale von Schulklassen (z. B. Leistungsstand der Schulklasse, Schüler aus verschiedenen Schichten, Häufigkeit des Lehrerwechsels) einzubeziehen sind, so dass Interaktionen wie etwa Unterrichtsmethode  $\times$  Schulklassenmerkmal überprüft werden können.

### 6.2.3 Verwendung eines ökonomischen Versuchsplans

Das Aufstellen eines ökonomischen Versuchsplans ist vom Hintergrundwissen und den damit zusammenhängenden Hypothesen abhängig, zudem vom Stichprobenumfang und der Anzahl der Faktorstufenkombinationen. Geht es um den Nachweis eines experimentellen Effekts ohne Verallgemeinerung auf die Population, kommt man mit relativ kleinen Stichproben aus. Auf die Forderung nach ökonomischen Versuchsplänen wird in der Literatur zur Versuchsplanung pädagogischer Experimente aufmerksam gemacht (Lewis, 1968; Cooper, 1988; Klauer, 1973, 2005).

### 6.3 Versuchspläne für die empirische Unterrichtsforschung

Dieser Abschnitt stellt fünf Typen von Versuchsplänen vor sowie Erweiterungen durch Kontrollvariablen und bezüglich mehrerer abhängiger Variablen (multivariate Versuchsplanung). Sie besitzen für die experimentelle Unterrichtsforschung einen besonderen Stellenwert, besonders zur Untersuchung von Lernerfolgen bei Schülern. Die Versuchspläne werden durch konkrete Daten aus der Unterrichtsforschung exemplifiziert, die zudem mit Boxplots, Mittelwerten und Konfidenzintervallen visualisiert werden.

In der Nomenklatur von Kirk (1996, 2012) werden die folgenden Typen von Versuchsplänen vorgestellt: RB- $p$  (Randomized Block designs), CR- $p$  (Completely Randomized designs), CRF- $p \times q$  (Completely Randomized Factorial designs), SPF- $p \bullet q$  (Split-Plot Factorial designs) und SPF- $p \times r \bullet q$  (Split-Plot Factorial designs). Zudem werden Versuchspläne mit Kovariaten und Versuchspläne mit multivariaten Daten behandelt.

#### 6.3.1 RB- $p$ Versuchspläne

RB- $p$  Versuchspläne sehen vor, dass von einem Lerner nicht nur eine, sondern  $p$  Messungen erhoben werden. Mit RB- $p$  Versuchsplänen lässt sich die Auswirkung eines  $p$ -fach gestuften Messwiederholungsfaktors auf eine abhängige Variable untersuchen.

Für die experimentelle Unterrichtsforschung sind RB- $p$  Versuchspläne durch drei Aspekte charakterisiert: (1) Unter dem Faktor  $A$  lässt sich eine Lehr-/Lernbedingung  $p$ -fach wiederholt untersuchen. (2) Lerner werden mehrmals mit der Lehr-/Lernbedingung konfrontiert, oder es werden zu verschiedenen Zeitpunkten Messungen ohne erneute Konfrontation mit der Lehr-/Lernbedingung vorgenommen (z. B. um die Behaltensleistung zu ermitteln). (3) Es wird eine Stichprobe mit  $n$  Lernern (z. B. in einer Klasse) benötigt.

Abbildung 6.1 veranschaulicht das Layout eines RB-4 Versuchsplans. Man beachte, dass die Lerner  $s_1, \dots, s_n$  unter allen  $p = 4$  Faktorstufen am Experiment teilnehmen.

$a_1$	$a_2$	$a_3$	$a_4$	$A = \text{Zeitpunkt}$ $a_1 = \text{nach 1. Unterrichtseinheit}$ $a_2 = \text{nach 2. Unterrichtseinheit}$ $a_3 = \text{nach 3. Unterrichtseinheit}$ $a_4 = \text{nach 4. Unterrichtseinheit}$
$s_1$	$s_1$	$s_1$	$s_1$	
...	...	...	...	
$s_n$	$s_n$	$s_n$	$s_n$	

Abbildung 6.1 Layout eines RB-4 Versuchsplans

Tabelle 6.1 enthält die Daten für einen RB-4 Versuchsplan mit  $n = 7$  Lernern. Von den Lernern  $s_1, \dots, s_7$  werden jeweils 4 Datensätze nach der 1., 2., 3. und 4. Unterrichtseinheit zum Lernerfolg [Note] erhoben. Angegeben sind zudem die Mittelwerte und die Standardfehler der Mittelwerte, die für die Berechnung von Konfidenzintervallen und Prüfgrößen (Teststatistiken) verwendet werden können.

Tabelle 6.1 Datensatz zum RB-4 Versuchsplan

	$a_1$	$a_2$	$a_3$	$a_4$
$s_1$	5.0	2.5	2.0	2.5
$s_2$	4.0	2.5	2.5	2.5
$s_3$	4.0	3.0	2.5	2.5
$s_4$	4.0	2.5	2.5	2.5
$s_5$	3.5	2.0	2.0	2.5
$s_6$	4.0	3.0	3.0	3.0
$s_7$	3.5	3.0	3.0	3.0
$\bar{x}$	4.0	2.6	2.5	2.6
$s_{\bar{x}}$	.19	.14	.15	.09

$A = \text{Zeitpunkt}$   
 $a_1 = \text{nach 1. Unterrichtseinheit}$   
 $a_2 = \text{nach 2. Unterrichtseinheit}$   
 $a_3 = \text{nach 3. Unterrichtseinheit}$   
 $a_4 = \text{nach 4. Unterrichtseinheit}$

Abbildung 6.2 zeigt Boxplots und Mittelwerte des Datensatzes aus Tabelle 6.1 sowie 95%-Konfidenzintervalle. Die Grafiken zeigen, dass sich Lernerfolge insbesondere zwischen der 1. und 2. Unterrichtseinheit einstellen. Die Lernerfolge werden nach Ende der 4. Unterrichtseinheit homogener, was das kleinere Konfidenzintervall im Vergleich zu den anderen Konfidenzintervallen anzeigt.

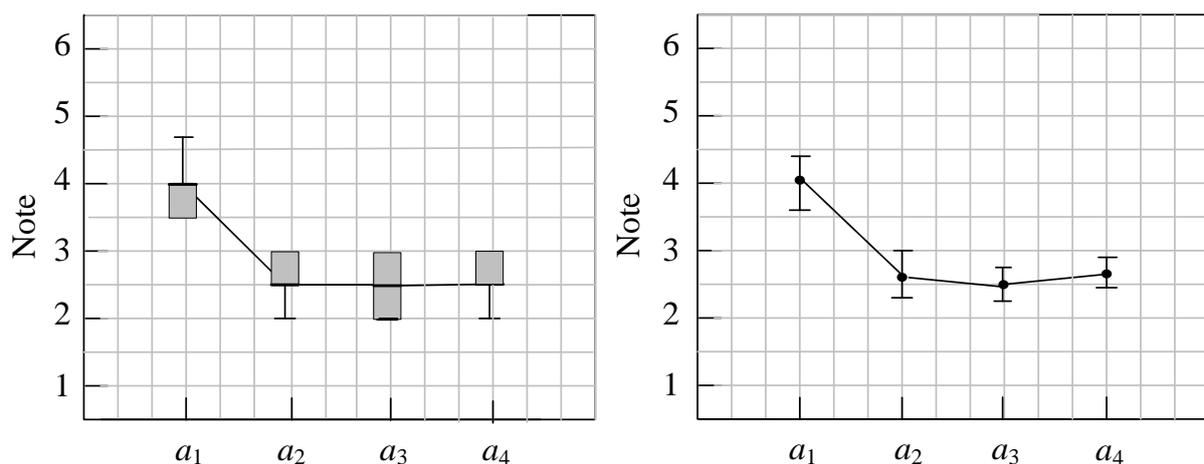


Abbildung 6.2 Boxplots, Mittelwerte und 95%-Konfidenzintervalle

### 6.3.2 CR- $p$ Versuchspläne

Mit CR- $p$  Versuchsplänen lässt sich die Auswirkung eines  $p$ -fach gestuften Faktors  $A$  auf eine abhängige Variable untersuchen.

Im Hinblick auf die experimentelle Unterrichtsforschung sind CR- $p$  Versuchspläne durch drei Eigenschaften charakterisiert: (1) Unter dem Faktor  $A$  lassen sich  $p \geq 2$  Lehr-/Lernbedingungen untersuchen; für  $p = 2$  können Halbklassen verwendet werden. (2) Lerner sind den Faktorstufen zugeordnet, so dass jeder Lerner unter einer Lehr-/Lernbedingung untersucht wird. (3) Es werden  $p$  Stichproben, z. B. Schulklassen mit Lernern benötigt.

Abbildung 6.3 veranschaulicht das Layout eines CR-3 Versuchsplans mit  $p = 3$  Faktorstufen (Klasse #1, Klasse #2, Klasse #3), das heißt, es werden 3 Stichproben  $S_1$ ,  $S_2$  und  $S_3$  herangezogen.

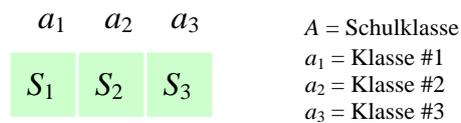


Abbildung 6.3 Layout eines CR-3 Versuchsplans

Tabelle 6.2 zeigt einen konkreten Datensatz für einen CR-3 Versuchsplan mit jeweils 7 Lernern. Von den Lernern wird jeweils 1 Datensatz zur Lernleistung [Note] erhoben. Für die Konstruktion von Konfidenzintervallen und Prüfgrößen sind die Mittelwerte und die Standardfehler der Mittelwerte berechnet.

Tabelle 6.2 Datensatz zum CR-3 Versuchsplan

	$a_1$	$a_2$	$a_3$	
	3.0	3.0	3.0	$A = \text{Schulklasse}$
	3.5	2.5	4.0	$a_1 = \text{Klasse \#1}$
	3.0	2.5	3.0	$a_2 = \text{Klasse \#2}$
	3.0	2.5	2.5	$a_3 = \text{Klasse \#3}$
	3.5	3.5	4.0	
	3.0	2.0	3.0	
	2.5	2.0	2.0	
$\bar{x}$	3.1	2.6	3.1	
$s_{\bar{x}}$	.13	.20	.28	

Abbildung 6.4 veranschaulicht Boxplots und Mittelwerte des Datensatzes aus Tabelle 6.2 mit 95%-Konfidenzintervallen. Die Schulklassen unterscheiden sich, was die Lernerfolge angeht, nur gering. Bemerkenswert sind die homogenen Lernerfolge (kleinere Konfidenzintervalle) der Klasse #1.

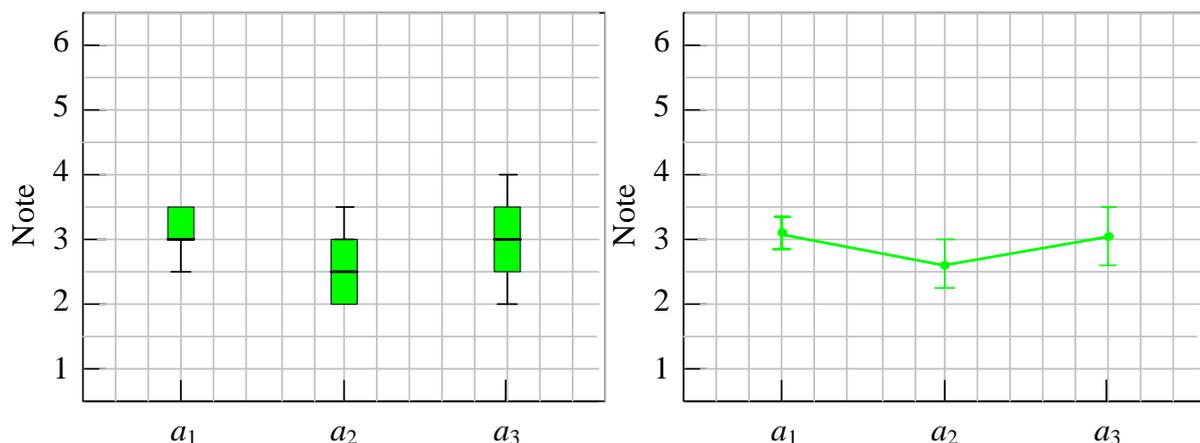


Abbildung 6.4 Boxplots, Mittelwerte und 95%-Konfidenzintervalle

### 6.3.3 CRF- $p \times q$ Versuchspläne

Mit CRF- $p \times q$  Versuchsplänen lässt sich allgemein überprüfen, wie eine abhängige Variable von zwei Faktoren  $A$  und  $B$  beeinflusst wird. Der Faktor  $A$  sei  $p$ -fach, der Faktor  $B$   $q$ -fach gestuft.

CRF- $p \times q$  Versuchspläne sind für die experimentelle Unterrichtsforschung wichtig, denn sie erlauben neben der Überprüfung zweier Haupteffekte, die Überprüfung von pädagogisch interessanten Interaktionseffekten. Somit lassen sich Interaktionen wie Unterrichtsmethode  $\times$  Schulklasse, Unterrichtsmethode  $\times$  Schulklassenmerkmal oder Unterrichtsmethode  $\times$  Schülermerkmal prüfen.

Für die experimentelle Unterrichtsforschung sind CRF- $p \times q$  Versuchspläne durch vier Aspekte charakterisiert: (1) Unter dem Faktor  $A$  lassen sich  $p \geq 2$  Bedingungen des Lehrens und Lernens gruppieren; für  $p = 2$  bietet es sich an, Halbklassen zu verwenden. (2) Unter dem Faktor  $B$  können  $q \geq 2$  Schulklassen (bzw. Schulklassenmerkmale oder Schülermerkmale) berücksichtigt werden. (3) Lerner werden zufällig den Faktorstufenkombinationen zugeordnet. (4) Es werden  $p \times q$  Stichproben mit Lernern benötigt.

Abbildung 6.5 veranschaulicht das Layout eines CRF- $3 \times 2$  Versuchsplans mit  $p = 3$  Unterrichtsmethoden (Computersimulation, Lernen durch Lehren, direkte Instruktion) und  $q = 2$  Faktorstufen für den Faktor Schulklasse (Klasse #1, Klasse#2); insgesamt werden  $3 \times 2 = 6$  Stichproben benötigt.

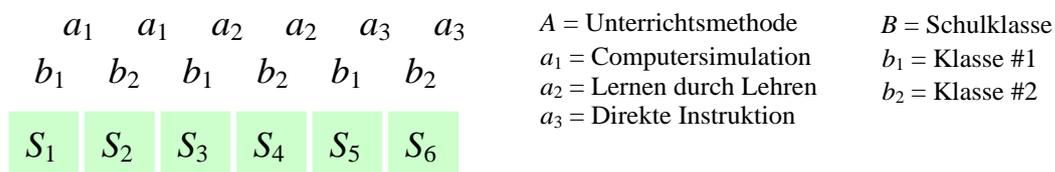


Abbildung 6.5 Layout eines CRF- $3 \times 2$  Versuchsplans

Tabelle 6.3 enthält den Datensatz für einen CRF- $3 \times 2$  Versuchsplan mit jeweils 7 Lernern. Von den Lernern wird 1 Datensatz zum Lernerfolg [Note] erhoben. Wieder sind die Mittelwerte und die Standardfehler der Mittelwerte berechnet; sie werden benötigt für die Berechnung von Konfidenzintervallen und Prüfgrößen.

Tabelle 6.3 Datensatz zum CRF- $3 \times 2$  Versuchsplan

	$a_1$	$a_1$	$a_2$	$a_2$	$a_3$	$a_3$	
	$b_1$	$b_2$	$b_1$	$b_2$	$b_1$	$b_2$	
	3.0	3.0	2.5	2.0	2.0	1.0	
	2.5	3.0	2.5	3.0	1.0	1.0	
	3.0	2.5	3.0	2.0	3.0	3.0	
	2.5	2.5	2.5	3.0	3.0	3.0	
	3.5	3.5	2.0	2.0	3.0	2.0	
	3.0	3.0	3.0	2.5	3.0	2.0	
	2.5	4.0	2.0	2.5	3.0	2.0	
$\bar{x}$	2.9	3.1	2.5	2.4	2.6	2.0	
$s_{\bar{x}}$	.22	.22	.24	.21	.22	.22	

$A =$ Unterrichtsmethode
$a_1 =$ Computersimulation
$a_2 =$ Lernen durch Lehren
$a_3 =$ Direkte Instruktion
$B =$ Schulklasse
$b_1 =$ Klasse #1
$b_2 =$ Klasse #2

Abbildung 6.6 visualisiert die Datensätze (Boxplots, Mittelwert, 95% Konfidenzintervalle) aus Tabelle 6.3. Aus der Abbildung lässt sich entnehmen, dass sich die Lernwirksamkeit der Unterrichtsmethoden unterscheidet: Mit der Unterrichtsmethode der Computersimulation wird der niedrigste, mit Lernen durch Lehren

der zweithöchste, mit der direkten Instruktion der höchste Lernerfolg bei den Lernern erzielt – und zwar in beiden Schulklassen.

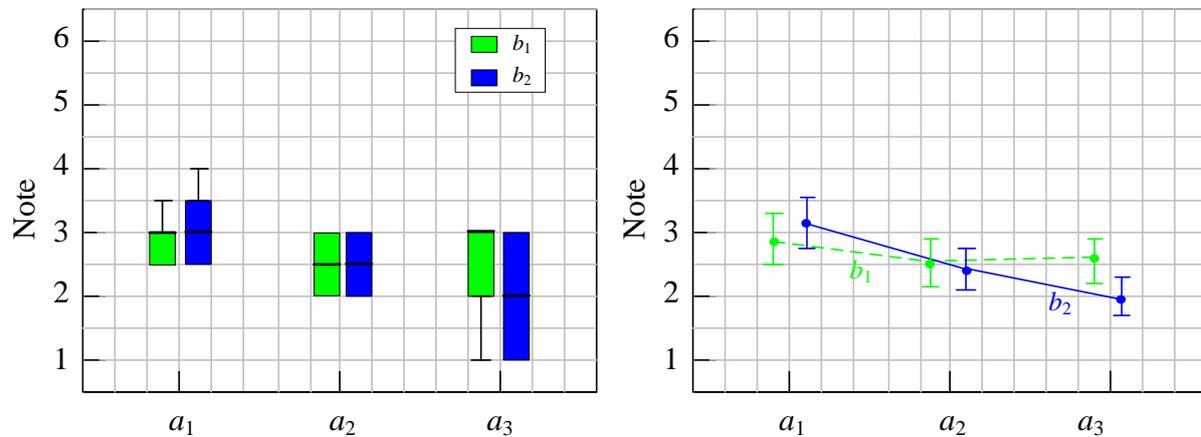


Abbildung 6.6 Boxplots, Mittelwerte und 95%-Konfidenzintervalle

### 6.3.4 SPF- $p \bullet q$ Versuchspläne

SPF- $p \bullet q$  Versuchspläne überprüfen, wie eine abhängige Variable von einem Blockfaktor  $A$  und einem Messwiederholungsfaktor  $B$  beeinflusst wird. Der Faktor  $A$  sei  $p$ -fach, der Faktor  $B$   $q$ -fach gestuft.

SPF- $p \bullet q$  Versuchspläne haben für die experimentelle Unterrichtsforschung eine sehr große Bedeutung, weil sie hervorragend geeignet sind, Lernprozesse in Bezug zu einem Blockfaktor im Sinne des Prozess-Produkt-Paradigmas (z. B. eine Lehrervariable, eine Kontextvariable, eine Prozessvariable) zu untersuchen.

Für die experimentelle Unterrichtsforschung sind SPF- $p \bullet q$  Versuchspläne durch vier Eigenschaften bestimmt: (1) Unter dem Faktor  $A$  lassen sich  $p \geq 2$  Bedingungen des Lehrens und Lernens gruppieren; für  $p = 2$  bietet es sich an Halbklassen zu verwenden. (2) Unter dem Faktor  $B$  können  $q \geq 2$  Messwiederholungen (Lernleistung über die Zeit) berücksichtigt werden. (3) Lerner aus den einzelnen Blöcken werden mit derselben Lehr-/Lernbedingung  $q$ -fach konfrontiert. (4) Es werden  $p$  Stichproben mit Lernern benötigt.

Abbildung 6.7 veranschaulicht das Layout eines SPF- $3 \bullet 2$  Versuchsplans mit  $p = 3$  Unterrichtsmethoden (Computersimulation, Lernen durch Lehren, direkte Instruktion) und  $q = 2$  Faktorstufen für zwei Lernzeitpunkte ( $b_1, b_2$ ). insgesamt werden 3 Stichproben benötigt mit den Lernern  $s_1, \dots, s_n, s_{n+1}, \dots, s_{2n}, s_{2n+1}, \dots, s_{3n}$ . Jeder Lerner wird zu zwei Lernzeitpunkten nach der 1. und 2. Unterrichtseinheit untersucht.

	$b_1$	$b_2$	
$a_1$	$s_1$	$s_1$	$A$ = Unterrichtsmethode $a_1$ = Computersimulation $a_2$ = Lernen durch Lehren $a_3$ = Direkte Instruktion  $B$ = Zeitpunkt $b_1$ = nach 1. Unterrichtseinheit $b_2$ = nach 2. Unterrichtseinheit
	...	...	
	$s_n$	$s_n$	
$a_2$	$s_{n+1}$	$s_{n+1}$	
	...	...	
	$s_{2n}$	$s_{2n}$	
$a_3$	$s_{2n+1}$	$s_{2n+1}$	
	...	...	
	$s_{3n}$	$s_{3n}$	

Abbildung 6.7 Layout eines SPF-3•2 Versuchsplans

Tabelle 6.4 zeigt den Datensatz für einen SPF-3•2 Versuchsplan mit jeweils 7 Lernern für die  $p = 3$  Unterrichtsmethoden. Von den Lernern werden 2 Datensätze zum Lernerfolg [Note] erhoben, nach der 1. und 2. Unterrichtseinheit. Für die Konstruktion von Konfidenzintervallen und Prüfgrößen sind die Mittelwerte und die Standardfehler der Mittelwerte berechnet.

Tabelle 6.4 Datensatz zum SPF-3•2 Versuchsplan

	$b_1$	$b_2$	
$s_1$	4.5	3.0	$A = \text{Unterrichtsmethode}$
$s_2$	4.0	2.0	$a_1 = \text{Computersimulation}$
$s_3$	3.0	2.0	$a_2 = \text{Lernen durch Lehren}$
$a_1$ $s_4$	4.5	2.5	$a_3 = \text{Direkte Instruktion}$
$s_5$	4.5	3.0	$B = \text{Zeitpunkt}$
$s_6$	4.0	3.0	$b_1 = \text{nach 1. Unterrichtseinheit}$
$s_7$	3.0	2.5	$b_2 = \text{nach 2. Unterrichtseinheit}$
$\bar{x}$	3.9	2.6	
$s_{\bar{x}}$	.21	.18	
	$b_1$	$b_2$	
$s_8$	5.0	4.0	
$s_9$	3.0	3.0	
$s_{10}$	4.0	3.5	
$a_2$ $s_{11}$	4.0	4.0	
$s_{12}$	4.0	3.5	
$s_{13}$	4.5	4.0	
$s_{14}$	4.0	3.0	
$\bar{x}$	4.1	3.6	
$s_{\bar{x}}$	.21	.18	
	$b_1$	$b_2$	
$s_{15}$	3.5	1.0	
$s_{16}$	4.0	1.0	
$s_{17}$	4.0	2.0	
$a_3$ $s_{18}$	4.0	2.0	
$s_{19}$	4.5	2.0	
$s_{20}$	4.0	1.0	
$s_{21}$	4.0	2.0	
$\bar{x}$	4.0	1.6	
$s_{\bar{x}}$	.21	.18	

Abbildung 6.8 visualisiert die Datensätze aus Tabelle 6.4 mit Boxplots, Mittelwerten und 95%-Konfidenzintervallen. Die Abbildung zeigt, dass alle drei Unterrichtsmethoden lernwirksam waren, mit deutlichen Vorteilen für die direkte Instruktion, vor Lernen durch Lehren und Computersimulation.

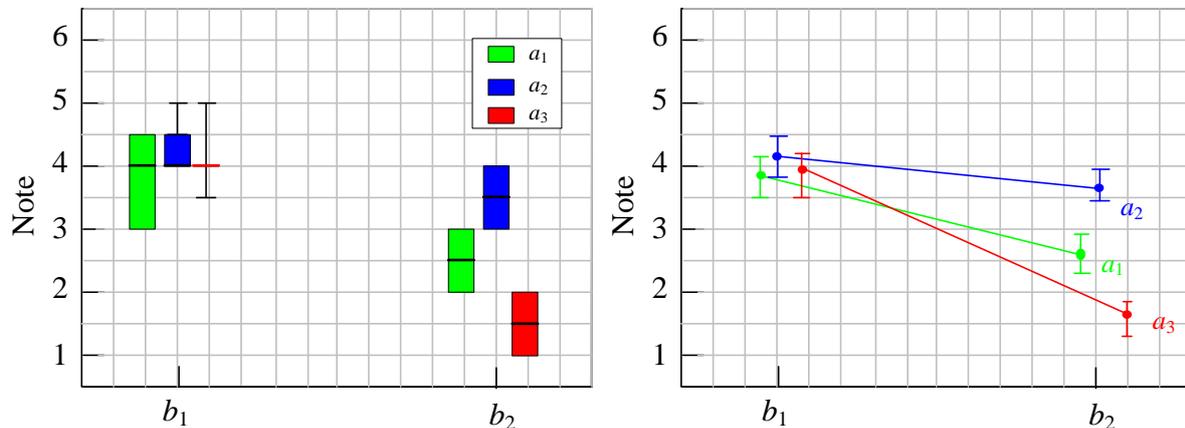


Abbildung 6.8 Boxplots, Mittelwerte und 95%-Konfidenzintervalle

### 6.3.5 SPF- $p \times q \cdot r$ Versuchspläne

SPF- $p \times r \cdot q$  Versuchspläne überprüfen, wie eine abhängige Variable von zwei Blockfaktoren  $A$ ,  $C$  und einem Messwiederholungsfaktor  $B$  beeinflusst wird.  $A$  sei  $p$ -,  $C$   $q$ - und  $B$   $r$ -fach gestuft.

SPF- $p \times r \cdot q$  Versuchspläne sind für die experimentelle Unterrichtsforschung aus folgenden Gründen sehr interessant: (1) Unter dem Faktor  $A$  lassen sich  $p \geq 2$  Bedingungen des Lehrens und Lernens gruppieren; für  $p = 2$  bietet es sich an Halbklassen zu verwenden. (2) Unter dem Faktor  $C$  kann man kombiniert mit Faktor  $A$  weitere  $r \geq 2$  Bedingungen des Lehrens und Lernens gruppieren (z. B. Schulklassen zur Kontrolle des Klassenkontextes, Geschlecht der Schüler). (3) Unter dem Faktor  $B$  können  $q \geq 2$  Messwiederholungen (Lernleistung über die Zeit) berücksichtigt werden. (4) Lerner aus den einzelnen Blöcken werden mit derselben Lehr-/Lernbedingung  $q$ -fach konfrontiert. (5) Es werden  $p \times r$  Stichproben mit Lernern benötigt.

Abbildung 6.9 veranschaulicht das Layout eines SPF- $2 \times 2 \cdot 3$  Versuchsplans (Messwiederholungen für Faktor  $B$ ) mit  $p \times r = 2 \times 2 = 4$  Stichproben und  $q = 3$  Stufen für Faktor  $B$ ; insgesamt werden 4 Stichproben benötigt mit den Lernern  $s_1, \dots, s_n, s_{n+1}, \dots, s_{2n}, s_{2n+1}, \dots, s_{3n}, s_{3n+1}, \dots, s_{4n}$ . Jeder Lerner wird zu drei Lernzeitpunkten nach der 1., 2. und 3. Unterrichtseinheit untersucht.

	$b_1$	$b_2$	$b_3$	
$a_1c_1$	$s_1$	$s_1$	$s_1$	$A = \text{Unterrichtsmethode}$ $a_1 = \text{Direkte Instruktion}$ $a_2 = \text{Lernen durch Lehren}$
	...	...	...	
	$s_n$	$s_n$	$s_n$	
$a_1c_2$	$s_{n+1}$	$s_{n+1}$	$s_{n+1}$	$B = \text{Zeitpunkt}$ $b_1 = \text{nach 1. Unterrichtseinheit}$ $b_2 = \text{nach 2. Unterrichtseinheit}$ $b_3 = \text{nach 3. Unterrichtseinheit}$
	...	...	...	
	$s_{2n}$	$s_{2n}$	$s_{2n}$	
$a_2c_1$	$s_{2n+1}$	$s_{2n+1}$	$s_{2n+1}$	$C = \text{Geschlecht}$ $c_1 = \text{Jungen}$ $c_2 = \text{Mädchen}$
	...	...	...	
	$s_{3n}$	$s_{3n}$	$s_{3n}$	
$a_2c_2$	$s_{3n+1}$	$s_{3n+1}$	$s_{3n+1}$	
	...	...	...	
	$s_{4n}$	$s_{4n}$	$s_{4n}$	

Abbildung 6.9 Layout eines SPF-2×2•3 Versuchsplans

Tabelle 6.5 enthält den Datensatz für einen SPF-2×2•3 Versuchsplan mit jeweils 7 Lernern für die  $p = 2$  Unterrichtsmethoden (direkte Instruktion, Lernen durch Lehren) und  $r = 2$  Gruppen von Lernern (Jungen, Mädchen). Von den Lernern werden 3 Datensätze zum Lernerfolg [Note] erhoben, nach der 1., 2. und 3. Unterrichtseinheit. Für die Konstruktion von Konfidenzintervallen und Prüfgrößen sind die Mittelwerte und die Standardfehler der Mittelwerte berechnet.

Tabelle 6.5 Datensatz zum SPF-2×2•3 Versuchsplan

	$b_1$	$b_2$	$b_3$	
$s_1$	4.5	4.0	3.0	A = Unterrichtsmethode $a_1$ = Direkte Instruktion $a_2$ = Lernen durch Lehren  B = Zeitpunkt $b_1$ = nach 1. Unterrichtseinheit $b_2$ = nach 2. Unterrichtseinheit $b_3$ = nach 3. Unterrichtseinheit
$s_2$	4.0	3.0	2.0	
$s_3$	3.0	2.0	2.0	
$a_1c_1$ $s_4$	4.5	4.0	2.0	
$s_5$	4.5	4.0	2.0	
$s_6$	4.0	3.0	2.0	
$s_7$	3.0	2.5	2.0	
$\bar{x}$	3.9	3.2	2.1	
$s_{\bar{x}}$	.21	.25	.16	
	$b_1$	$b_2$	$b_3$	
$s_8$	3.5	3.0	3.0	C = Geschlecht $c_1$ = Jungen $c_2$ = Mädchen
$s_9$	4.0	3.0	2.0	
$s_{10}$	4.0	3.0	3.0	
$a_1c_2$ $s_{11}$	4.0	3.0	3.0	
$s_{12}$	3.5	2.5	3.0	
$s_{13}$	3.0	2.0	2.5	
$s_{14}$	4.0	2.0	2.0	
$\bar{x}$	3.7	2.6	2.6	
$s_{\bar{x}}$	.21	.25	.16	
	$b_1$	$b_2$	$b_3$	
$s_{15}$	5.0	4.0	2.0	
$s_{16}$	3.0	3.0	1.5	
$s_{17}$	4.0	3.0	2.0	
$a_2c_1$ $s_{18}$	4.0	4.0	2.0	
$s_{19}$	4.0	3.0	1.0	
$s_{20}$	4.5	3.5	2.0	
$s_{21}$	4.0	3.0	2.0	
$\bar{x}$	4.1	3.4	1.8	
$s_{\bar{x}}$	.21	.25	.16	
	$b_1$	$b_2$	$b_3$	
$s_{22}$	3.5	3.0	2.0	
$s_{23}$	4.0	3.0	2.0	
$s_{24}$	4.0	3.0	1.5	
$a_2c_2$ $s_{25}$	3.0	2.0	1.5	
$s_{26}$	4.5	3.0	2.0	
$s_{27}$	4.0	1.0	1.0	
$s_{28}$	4.0	2.0	2.0	
$\bar{x}$	3.8	2.4	1.7	
$s_{\bar{x}}$	.21	.25	.16	

Abbildung 6.10 visualisiert die Datensätze aus Tabelle 6.5 mit Boxplots, Mittelwerten und 95%-Konfidenzintervallen. Die Abbildung zeigt, dass die beiden Unterrichtsmethoden über die Zeit lernwirksam waren. Allerdings sind die Lernerfolge mit direkter Instruktion bei den Mädchen nach der 3. Unterrichtseinheit schlechter als bei den Jungen.

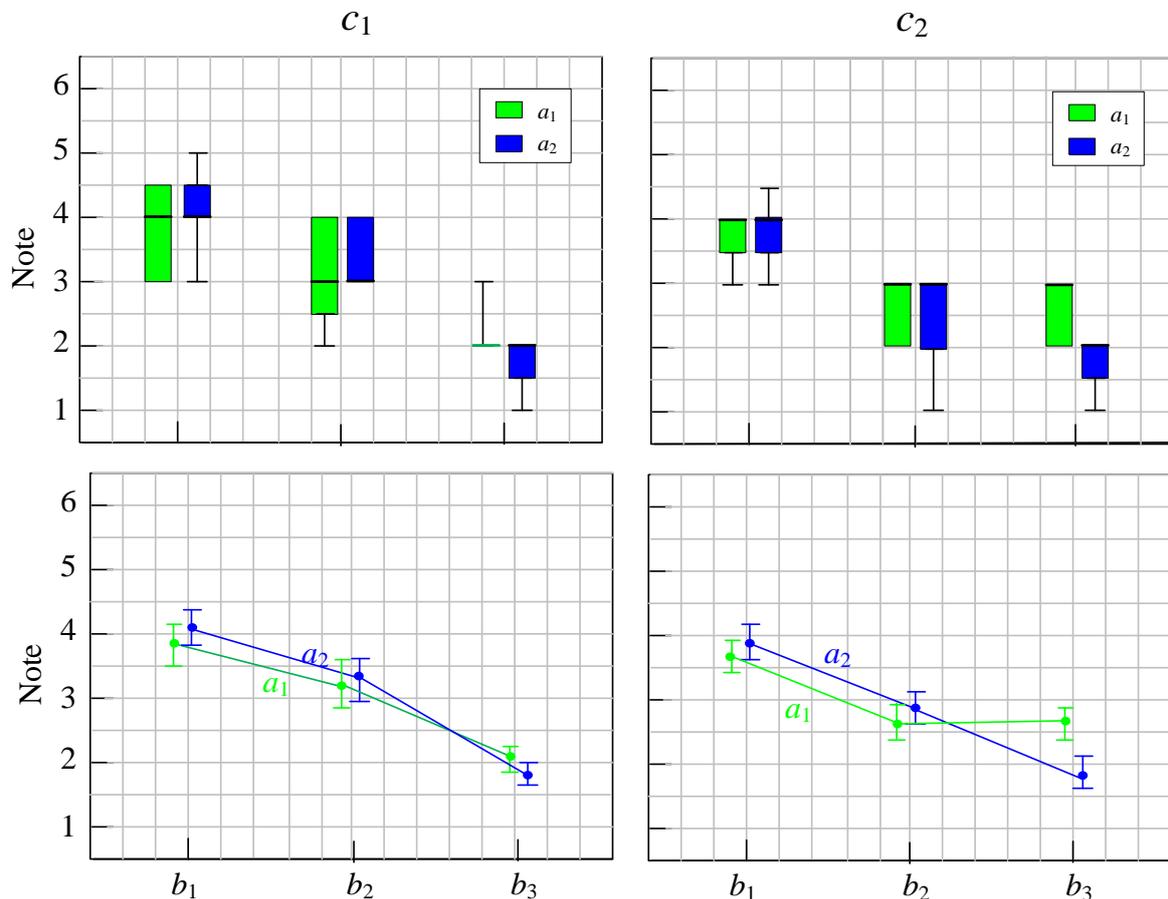


Abbildung 6.10 Boxplots, Mittelwerte und 95%-Konfidenzintervalle

### 6.3.6 Kovariante Versuchspläne

Kovariante Versuchspläne (auch: kovarianzanalytische Versuchspläne) sind Versuchspläne, die Kontrollvariablen mit in die Versuchsplanung einbeziehen. Solche Versuchspläne sind für die experimentelle Unterrichtsforschung besonders dann von Interesse, wenn *a priori*-Unterschiede zwischen Lernern (z. B. (Vor-)Wissensstand, Intelligenz, Alter) berücksichtigt werden sollen. Kontrollvariablen müssen metrisch skaliert sein.

Tabelle 6.6 zeigt den Datensatz für den CR-3 Versuchsplan mit 1 Kovariante (siehe Abbildung 6.3) mit jeweils 7 Lernern. Von den Lernern wird jeweils 1 Datensatz zur Lernleistung (Note im Test) erhoben. Zudem sind Daten für die Kovariante verfügbar (Zeugnisnote vom Vorjahr). Für die Konstruktion von Konfidenzintervallen und Prüfgrößen

sind die Mittelwerte und die Standardfehler der Mittelwerte berechnet, sowohl für die Lernleistung [Note im Test] als auch für die Zeugnisnote vom Vorjahr.

Tabelle 6.6 CR-3 Versuchsplan mit 1 Kovariate

	$a_1$		$a_2$		$a_3$	
	$X$	$Y$	$X$	$Y$	$X$	$Y$
	2.5	3.0	3.0	3.0	2.5	3.0
	2.0	3.5	3.5	2.5	2.0	4.0
	3.0	3.0	3.0	2.5	2.5	3.0
	2.5	3.0	3.0	2.5	2.5	2.5
	2.0	3.5	3.5	3.5	3.5	4.0
	3.0	3.0	3.5	2.0	2.5	3.0
	2.5	2.5	4.0	2.0	2.5	2.0
$\bar{x}$	2.5	3.1	3.4	2.6	2.6	3.1
$s_{\bar{x}}$	.15	.13	.14	.20	.17	.28

$A$  = Schulklasse  
 $a_1$  = Klasse #1  
 $a_2$  = Klasse #2  
 $a_3$  = Klasse #3  
 $X$  = Zeugnisnote vom Vorjahr (Kovariate)  
 $Y$  = Note im Test (abhängige Variable)

Abbildung 6.11 veranschaulicht die Mittelwerte des Datensatzes aus Tabelle 6.6 und die 95%-Konfidenzintervalle. Die Schulklassen unterscheiden sich hinsichtlich der Lernerfolge (Note im Test): Die Schüler der Klasse #2 haben die besten Noten im Test. Dies muss allerdings noch relativiert werden bezüglich der Zeugnisnote vom Vorjahr. Berücksichtigt man nämlich, dass die Schüler der Klasse #2 (durchschnittlich) schlechtere Zeugnisnoten im Vorjahr hatten, sind die Noten bemerkenswerter.

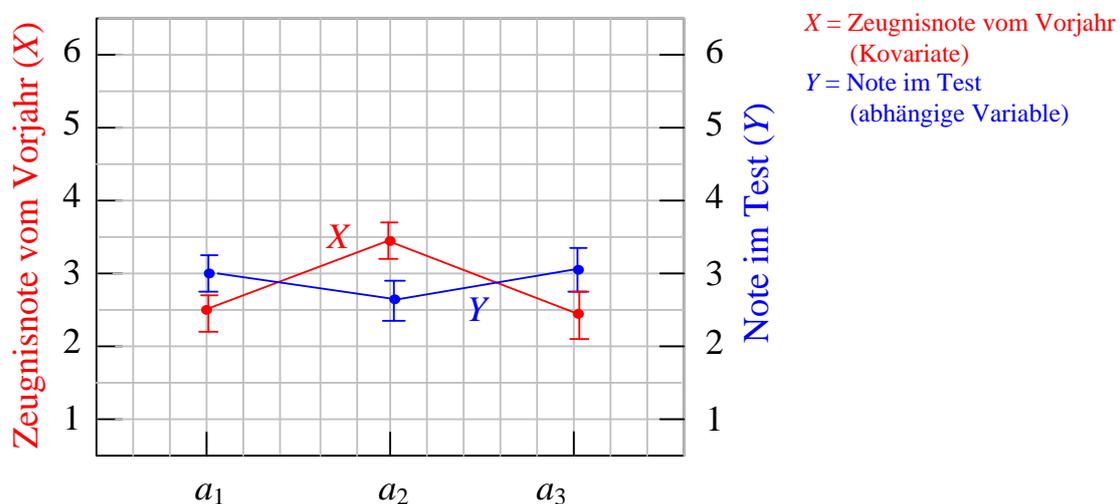


Abbildung 6.11 Mittelwerte und 95%-Konfidenzintervalle

## Anmerkungen

(1) Die Berücksichtigung von Kontrollvariablen ist für alle vorgestellten Versuchspläne möglich. (2) Auch ist die Berücksichtigung von mehreren Kontrollvariablen möglich. (3) Beispiele zur Versuchsplanung lassen sich in Kirk (2012, Kapitel 13) finden.

### 6.3.7 Multivariate Versuchspläne

Multivariate Versuchspläne sind Versuchspläne, die mehrere abhängige Variablen in die Versuchsplanung einbeziehen. Prinzipiell verwendet die multivariate Versuchsplanung dieselben Versuchspläne wie die univariate Versuchsplanung (vgl. Abschnitte 6.3.1 bis 6.3.5); sie unterscheidet sich von ihr in der Datenerhebung.

In der univariaten Versuchsplanung wird eine abhängige Variable (z. B. Lernzuwachs) untersucht, wohingegen in der multivariaten Versuchsplanung mehrere abhängige Variablen untersucht werden.

Tabelle 6.7 zeigt den multivariaten Datensatz für einen vier-variablen CR-3 Versuchsplan (siehe Abbildung 6.3) mit jeweils 7 Lernern. Von den Lernern wird jeweils 1 Datensatz zur Lernleistung erhoben mit den Teilnoten #1, #2, #3 und #4. Für die Konstruktion von Konfidenzintervallen und Prüfgrößen sind die Mittelwerte und die Standardfehler der Mittelwerte für alle Teilnoten berechnet.

Tabelle 6.7 Multivariater Datensatz zum CR-3 Versuchsplan

	$a_1$				$a_2$				$a_3$				
	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_1$	$Y_2$	$Y_3$	$Y_4$	
	2.5	3.0	2.5	3.0	3.0	2.0	1.0	3.0	2.0	2.0	2.0	1.0	$A = \text{Schulklasse}$
	3.0	3.5	2.0	3.0	2.5	2.5	2.0	3.0	1.0	1.0	1.0	1.0	$a_1 = \text{Klasse \#1}$
	3.0	3.0	2.5	3.5	3.0	3.0	2.0	3.5	1.0	1.0	1.0	2.0	$a_2 = \text{Klasse \#2}$
	2.5	3.0	2.5	2.5	2.5	3.0	1.0	2.5	1.0	2.0	2.5	2.0	$a_3 = \text{Klasse \#3}$
	3.0	3.5	3.5	3.0	2.5	2.0	2.0	3.0	2.5	2.0	2.0	1.0	$Y_1 = \text{Teilnote \#1}$
	3.0	3.0	2.5	3.0	2.0	2.0	2.5	2.0	1.0	2.0	1.0	3.0	$Y_2 = \text{Teilnote \#2}$
	2.5	2.5	2.5	4.0	2.0	2.5	2.5	2.0	2.5	2.5	1.0	2.0	$Y_3 = \text{Teilnote \#3}$
$\bar{x}$	2.8	3.1	2.6	3.1	2.5	2.4	1.9	2.7	1.6	1.8	1.5	1.7	$Y_4 = \text{Teilnote \#4}$
$s_{\bar{x}}$	.10	.13	.17	.17	.15	.17	.24	.21	.27	.21	.24	.29	

Abbildung 6.7 zeigt die Mittelwerte des Datensatzes aus Tabelle 6.7 und gibt 95%- Konfidenzintervalle an. Die Schulklassen unterscheiden sich hinsichtlich der Lernerfolge (Note im Test): Die Schüler der Klasse #2 haben die besten Noten im Test.

## Anmerkungen

(1) Die Berücksichtigung multivariater Daten ist für alle vorgestellten Versuchspläne möglich. (2) In die Versuchsplanung können mehrere abhängige Variablen und mehrere Kontrollvariablen (Kovariaten) einbezogen werden. (3) Beispiele zur Versuchsplanung mit multivariaten Daten finden sich in Bortz und Schuster (2010, Kapitel 26).

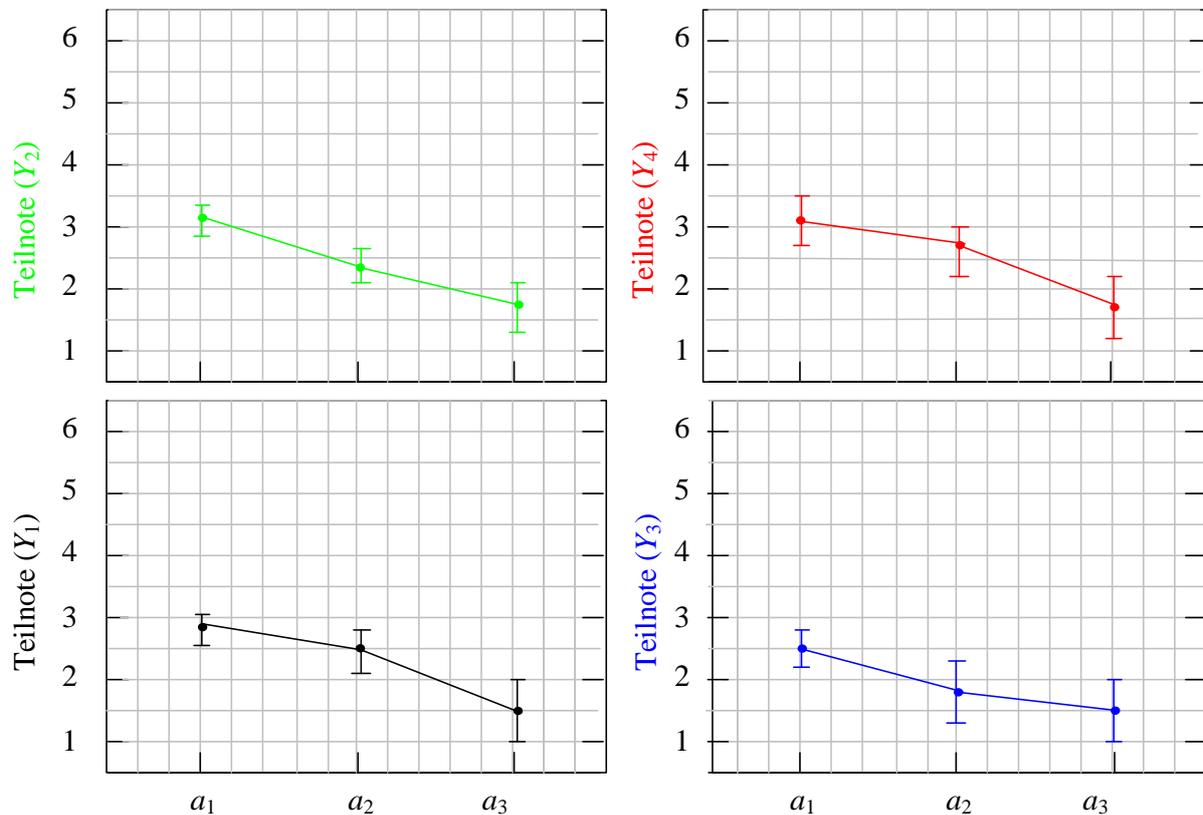


Abbildung 6.12 Mittelwerte und 95%-Konfidenzintervalle

## 6.4 Zusammenfassung und Ausblick

Dieses Kapitel behandelte fünf Typen von Versuchsplänen (RB- $p$ , CR- $p$ , CRF- $p \times q$ , SPF- $p \bullet q$ , SPF- $p \times r \bullet q$ ) sowie Erweiterungen durch Kovariaten und hinsichtlich mehrerer abhängiger Variablen, die für die experimentelle Unterrichtsforschung, besonders zur Untersuchung von Lernerfolgen bei Schülern, wichtig sind. Die Versuchspläne wurden durch konkrete Daten exemplifiziert und mit Boxplots, Mittelwerten sowie Konfidenzintervallen visualisiert.

Für die Unterrichtsforschung spielen Versuchspläne mit Kovariaten eine wichtige Rolle, weil sich mit ihnen der (Vor-)Wissenstand von Schülern kontrollieren lässt. So lassen sich beispielweise mit einem kovarianzanalytischen CRF- $2 \times 2$  Versuchsplan der Einfluss von Unterrichtsmethoden (Faktor A) und Klassenkontext (Faktor B) auf den Lernerfolg

(abhängige Variable) untersuchen, wenn als Kovariate der (Vor-)Wissenstand der Schüler erhoben worden ist.

Über die vorgestellten Versuchspläne hinaus gibt es weitere Möglichkeiten, einzelne Schulklassen ungeteilt einzubeziehen, etwa mit der wichtigen Gruppe der hierarchischen faktoriellen Versuchspläne (vgl. Kirk, 2012). Sie eignen sich besonders gut, komplexe unterrichtsrelevante Fragestellungen ökonomisch behandeln zu können, wenn differenzierte Forschungshypothesen vorliegen. Nachfolgende Arbeiten sollten solche Versuchspläne wie auch Lateinische Quadrate und Griechisch-Lateinische Quadrate berücksichtigen.

Wenn Schulklassen oder Schulen untersucht werden sollen und insbesondere der Klassenkontext zu berücksichtigen ist, dann spielen (multilevel) Versuchspläne auf der Grundlage des *Hierarchical Linear Model* (HLM) eine besondere Rolle (Bryk & Raudenbush, 1992; Hox, 2000; Hedeker, 2008).

## 6.5 Literatur

- Berliner, D. C., & Calfee, R. (Hrsg.) (1996). *Handbook of educational psychology*. New York: Macmillan.
- Bortz, J., & Schuster, C. (2010). *Statistik*. Berlin: Springer.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Cooper, M. (1988). Nonparametric statistics. In J. P. Keeves (Hrsg.), *Educational research, methodology, and measurement: An international handbook* (S. 705–710). Oxford: Pergamon Press.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on aptitude-treatment interactions*. New York: Irvington.
- Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching*. New York: Holt, Rinehart & Winston.
- Graham, S., Harris, K. H., & Zito, J. (2005). Promoting internal and external validity: A synergism of laboratory-like experiments and classroom-based self-regulated strategy development research. In G.D. Phye, D.H. Robinson, & J.R. Levin (Hrsg.), *Empirical methods for evaluating educational environments* (S. 235–265). New York.
- Hedeker, D. (2008). Multilevel models for ordinal and nominal variables. In J. de Leeuw & E. Meijer (Hrsg.), *Handbook of multilevel analysis* (S. 237–275). New York: Springer.
- Helmke, A. (2006). Unterrichtsforschung. In K.-H. Arnold, U. Sandfuchs, & J. Wiechmann (Hrsg.), *Handbuch Unterricht* (S. 56–65). Bad Heilbrunn: Julius Klinkhardt.
- Helmke, A. (2007). *Unterrichtsqualität erfassen, bewerten, verbessern* (5. Auflage). Seelze: Kallmeyersche Verlagsbuchhandlung.
- Hox, J. J. (2000). Multilevel analysis of grouped and longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Hrsg.), *Modeling longitudinal and multilevel data* (S. 15–32). London: Lawrence Erlbaum Associates.

- Kelly, A. E., & Lesh, R. (2000). Trends and shifts in research methods. In A. E. Kelly, & R. A. Lesh (Hrsg.), *Handbook of research design in mathematics and science education* (S. 35–44). London: Lawrence Erlbaum.
- Kirk, R. E. (1996). *Experimental design* (3. Auflage). Pacific Grove, CA: Brooks/Cole Publishing Company.
- Kirk, R. E. (2012). *Experimental design* (4. Auflage). Pacific Grove, CA: Brooks/Cole Publishing Company.
- Klauer, K. J. (1973). *Das Experiment in der pädagogischen Forschung*. Düsseldorf: Schwann.
- Klauer, K. J. (2005). *Das Experiment in der pädagogisch-psychologischen Forschung* (Reprint). Münster: Waxmann.
- Lewis, D. G. (1968). *Experimental design in education*. London: University of London Press.
- Marjoribanks, K. M. (1988). Interaction effects. In J. P. Keeves (Hrsg.), *Educational research, methodology, and measurement: An international handbook* (S. 664–672). Oxford: Pergamon Press.
- Namboodiri, N. K. (1972). Experimental design in which each subject is used repeatedly. *Psychological Bulletin*, 77(1), 54–64.
- O'Malley, P. M. (1988). Detection of interaction. In J. P. Keeves (Hrsg.), *Educational research, methodology, and measurement: An international handbook* (S. 625–628). Oxford: Pergamon Press.
- Shuell, T. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. Calfee (Hrsg.), *Handbook of educational psychology* (S. 726–764). New York: Macmillan.
- Wellenreuther, M. (2007). *Lehren und Lernen – aber wie?* Baltmannsweiler: Schneider.
- Winer, B. J. (1971). *Statistical principles in experimental design*. New York: Wiley.
- Zendler, A., & Pfeiffer, T. (2009). Methodische Befunde zu durchgeführten Unterrichtsexperimenten. *Empirische Pädagogik*, 23(2), 208–221.