

# Zum Einfluss fehlender Werte („Missing Data“) auf Ergebnisse einer Längsschnittstudie im Vorschulalter – Wie zufällig sind „fehlende Beobachtungen“?

Heinz Krombholz, Staatsinstitut für Frühpädagogik IFP München

## Zusammenfassung

Bei Längsschnitterhebungen über einen längeren Zeitraum können meist nicht alle Daten wie vorgesehen erfasst werden. Das Fehlen von Daten stellt für die Auswertung und die Interpretation der Ergebnisse ein nicht unerhebliches Problem dar, da keinesfalls vorausgesetzt werden kann, dass die Ausfälle zufällig erfolgen. Anhand der Daten einer Felduntersuchung in Kindertagesstätten über zwei Jahre wurden die Ergebnisse von „vollständigen“ Fällen mit denen von Kindern mit „fehlenden Beobachtungen“ („Missing Data“) verglichen. Vorrangiges Ziel der Studie war die Prüfung der Auswirkungen motorischer Fördermaßnahmen in Kindertagesstätten auf die motorische Entwicklung. Hierfür wurden 273 Kinder der Interventionsgruppe mit 295 Kindern der Kontrollgruppe verglichen.

Etwa 30 Prozent der teilnehmenden Kinder standen für die Abschlussuntersuchung nicht mehr zur Verfügung, da sie die Tagesstätten verlassen hatten oder nicht alle Daten wie geplant erhoben werden konnten. Die in der Stichprobe verbliebenen und die ausgeschiedenen Kindern unterschieden sich nicht hinsichtlich der Ausgangsdaten (anthropometrische Merkmale, motorische und kognitive Leistungen). Ebenfalls zeigten sich keine Zusammenhänge zwischen dem Verbleib in der Stichprobe und dem Geschlecht, der sozialen Schicht, der Teilnahme an Angeboten von Sportvereinen und der Nationalität. Allerdings schieden mehr Kinder der Kontrollgruppe als der Interventionsgruppe im Studienverlauf aus. In der Interventionsgruppe betrafen die Ausfälle insbesondere leistungsfähigere Kinder.

Konsequenzen für die Auswertung von Längsschnittuntersuchungen werden diskutiert.

Schlüsselwörter: fehlende Beobachtungen („Missing Data“), Längsschnittstudie, Vorschulalter, Motorische Entwicklung

## The Impact of Missing Data on the Results of a Longitudinal Study in Preschool Age - How coincidental are "missing observations"?

### **Summary**

Usually not all data can be collected as planned for longitudinal studies. Missing data are a pervasive problem in empirical psychological research because it cannot be assumed that missing observations happen at random. Based on data from a two-year longitudinal study in day care centers the results of "complete cases" were compared with those of children with missing data. The primary objective of the study was to examine the effects of a physical activity intervention in daycare on motor development. For this purpose, 273 children of the intervention group were compared with 295 children of the control group.

About 30 percent of the children participating were not available for the final examination because they either had left the day care centers or not all data could be collected as planned. The children remaining in the sample and the children who had dropped out did not differ in baseline data (anthropometric characteristics, motor and cognitive performance). There were also no correlations between sample retention and gender, social class, attendance at sports clubs and nationality. However, more children in the control group than in the intervention group left during the course of the study. In the intervention group, the dropping-out of more capable children was prevalent.

Implications for the statistical analyses of longitudinal studies with missing data are discussed.

**Keywords:** Missing Data, Longitudinal Study, Preschool Age, Motor Development

Bedauerlicherweise können in empirischen Untersuchungen oft nicht alle Daten wie vorgesehen erhoben werden. Die gilt für experimentelle Studien, aber noch mehr für quasiexperimentelle Felduntersuchungen mit einer hohen Zahl von abhängigen und unabhängigen Variablen und TeilnehmerInnen, besonders wenn es sich dabei um Längsschnitterhebungen über längere Zeiträume handelt. Der Ausfall von Probanden und das Fehlen von Daten stellen ein erhebliches Problem für die Auswertung und die Interpretation der Ergebnisse dar, da nicht erwartet werden darf, dass die Ausfälle zufällig („Missing at random“) erfolgen (Krombholz 1985, Diaz-Ordaz, Kenward, Cohen, Coleman, & Eldridge 2014). Das Problem „fehlende Werte“ wird in der psychologischen Forschung zunehmend, aber immer noch nicht ausreichend beachtet (Lüdtke, Robitzsch, Trautwein & Köller 2007, Fiero, Huang, Oren & Bell 2016). Selbst wenn in Beschreibungen von Studie erwähnt wird, dass nicht alle Daten erhoben werden konnten, wird selten auf das weitere Vorgehen bei der Analyse der Datensätze mit „Missing Data“ eingegangen (Diaz-Ordaz, Kenward, Cohen, Coleman, & Eldridge 2014, Lüdtke, Robitzsch, Trautwein & Köller 2007, Fiero, Huang, Oren & Bell 2016).

Können nicht von allen für eine Untersuchung ausgewählte Probanden alle vorgesehenen Daten erhoben werden, so ist die Repräsentativität der Stichprobe grundsätzlich in Frage gestellt. Verweigern z. B. ausgewählte Personen die Teilnahme an einer Studie, so gelten die Ergebnisse nur für Personen, die bereit sind, an derartigen Studien teilzunehmen. Erfolgen die Ausfälle in einer Längsschnittstudie nicht zufällig sondern gruppenspezifisch, ändert sich die Zusammensetzung der Stichprobe. Dies kann die Ergebnisse von Interventionsstudien entscheidend beeinflussen.

Meist dürfte es kaum möglich sein, anhand vorhandener und sinnvoller Kriterien „Teilnehmer“ und „Verweigerer“ zu vergleichen. Fehlen dagegen nur einige Daten, die erhoben werden sollten, indem die Teilnehmer einzelne Fragen nicht beantworten (wollen) oder – bei Panelstudien – aus einer laufenden Studie ausscheiden oder an vorgesehen Erhebungszeiten nicht erreichbar sind, so besteht zumindest die Möglichkeit, anhand der verfügbaren Daten Fälle mit vollständigen und unvollständigen Datensätzen zu vergleichen.

Von entscheidender Bedeutung ist, wie fehlende Beobachtungen, sei es hinsichtlich abhängiger oder unabhängiger Variablen, zustande kommen. Nach Rubin (1976) werden drei Arten von fehlenden Werten unterschieden: vollkommen zufällig (Missing Completely at Random MCAR, keinerlei Informationen zu möglichen Ursachen), zufällig (Missing at Random MAR, ignorierbare Nichtantwort) oder nicht zufällig (Missing not at Random MNAR, nicht ignorierbare Nichtantwort). Die Prüfung und Entscheidung, wie die fehlenden Werte zustande kommen, ist allerdings keineswegs trivial und nur selten möglich. Sind die Ausfälle zufällig (MCAR- oder MAR-Datensätze) bestehen, abgesehen davon, dass die Analysen wegen der verminderten Stichprobe weniger aussagekräftig sind, keine grundsätzlichen Probleme. Ein MNAR Ausfall führt dagegen zu nicht der Realität entsprechenden Schlussfolgerungen (Göthlich 2009).

Zur Prüfung der Zufälligkeit fehlender Werte und zum angemessenen Vorgehen bei MNAR-Ausfällen gibt es umfangreiche Empfehlungen (Little & Rubin 2002). Ziel ist es, fehlende Werte anhand vorliegender Werte möglichst zuverlässig zu schätzen. Neben den **klassischen Verfahren** stehen **imputationsbasierte und modellbasierte Verfahren** zur Verfügung (Lüdtke, Robitzsch, Trautwein & Köller 2007, Göthlich 2009).

Die **klassischen Verfahren** sind: (a) fallweiser Ausschluss (Complete Case Analysis): eine Person fällt aus den Berechnungen aus, wenn einer der vorgesehenen Werte fehlt (Beschränkung auf die vollständigen Fälle), (b) paarweiser Ausschluss (Available Case Analysis): alle vorliegenden Werte werden verwendet (möglichst optimale Nutzung der vorliegenden Daten) und (c) Gewichtungsverfahren: Datensätze mit fehlenden Werten sollen nicht berücksichtigt und vollständige stärker in die Analysen eingehen.

Bei den **imputationsbasierten Verfahren** wird versucht, die fehlenden Werte durch geeignete Schätzwerte zu ersetzen (z.B. Einsetzen des Mittelwerts, der anhand der vorliegenden Werte berechnet wurde, Schätzung durch Regression, Multiple Imputation). Es werden zunächst die fehlenden Werte geschätzt und danach mit diesen Daten die geplanten Analysen durchgeführt. Bei **den modellbasierten Verfahren** erfolgen die Schätzung der fehlenden Werte und die Schätzung der Modellparameter parallel (Göthlich 2009).

Die einfachsten und am häufigsten in der Praxis verwendeten Methoden zur Behandlung von fehlenden Werten sind der Listenweise und der Paarweise Fallausschluss und das Ersetzen fehlender Werte durch individuelle Mittelwerte. Im Statistikpaket IBM SPSS sind diese Möglichkeiten vorgesehen, darüber hinaus bietet das Erweiterungsmodul Missing Values zwei Prozeduren zur vertieften Analyse und Behandlung fehlender Werte, die Missing Values Analyse MVA und die Multiple Imputation, die die Analyse der Muster fehlender Daten ermöglicht, aber nur selten zur Anwendung kommen (Lüdtke, Robitzsch, Trautwein & Köller 2007, Baltes-Götz 2013).

Beim Listenweisen Fallausschluss werden bei der statistischen Analyse nur Fälle berücksichtigt, bei denen alle Variablen gültige Ausprägungen aufweisen. Fehlt nur eine einzige Variable wird der komplette Fall nicht berücksichtigt. Ein Nachteil ist der mitunter sehr große Informationsverlust, falls viele Fälle ausgeschlossen werden müssen; dies hat unter anderem Auswirkungen auf die Power der verwendeten statistischen Tests. Beim Paarweisen Fallausschluss werden nur einzelne Variablen mit fehlenden Werten nicht berücksichtigt, alle anderen Werte werden wenn möglich in die Berechnungen mit einbezogen, was eine größere Anzahl von Fällen zur Folge hat. Werden fehlender Werte durch individuelle Mittelwerte aus vorhandenen Items ergänzt, so erhöht sich die Anzahl der Fälle weiter.

Auf die einzelnen Verfahren und ihre jeweiligen Vor- und Nachteile kann hier nicht eingegangen werden, eine Einführung zum Umgang mit fehlenden Daten und Lösungsvorschläge finden sich in Lüdtke, Robitzsch, Trautwein & Köller 2007. Welche Vorgehensweise beim Umgang mit fehlenden Daten zu realistischeren Ergebnissen führt, ist umstritten, es scheint sich aber zu zeigen dass modellbasierte Verfahren und die multiple Imputation die „wahren Werte“ am besten repräsentieren (Göthlich 2009).

Im Folgenden soll anhand der Daten einer Längsschnittstudie im Vorschulalter geprüft werden, ob Kinder, die aus der Studie ausscheiden oder bei denen nicht alle Daten vorliegen, sich von Kindern unterscheiden, deren Daten vollständig erhoben werden konnten<sup>1</sup>. Ziel dieser Studie war

---

<sup>1</sup> Auf das Vorliegen von „Missing data“ wurde bereits in den vorliegenden Veröffentlichungen der Untersuchungsergebnisse eingegangen (Krombholz 2004b, 2005). Da es aber nicht möglich war, dieses Problem mit der notwendigen Ausführlichkeit darzustellen, soll die Analyse der „Ausfälle“ hier noch einmal ausführlicher dargestellt werden.

es zu prüfen, ob sich ein erhöhtes kindgemäßes Bewegungsangebot:

- in Kindertagesstätten verwirklichen lässt,
- welche Auswirkungen solch ein Angebot auf die körperliche, motorische und kognitive Entwicklung der Kinder hat und
- ob sich Zusammenhänge mit dem sozialen Verhalten der Kinder aufzeigen lassen.

An der Studie beteiligten sich 22 Kindertagesstätten in München, je 11 als Interventions- und Kontrolleinrichtungen. Die Teilnahme an der Studie als Interventions- bzw. Kontrolleinrichtung seitens der Kindergärten war freiwillig, eine zufällige Zuordnung auf Interventions- bzw. Kontrollgruppe war nicht in allen Fällen gewährleistet, eine Parallelisierung erfolgte anhand der Kriterien Einzugsgebiet, Größe und Ausstattung. Insgesamt bestand die Gesamtstichprobe aus mehr als 700 Kindern, etwa gleich viel Jungen und Mädchen. Da ein Teil dieser Kinder altersbedingt aus der Studie vorzeitig ausschieden, da sie eingeschult wurden, reduzierte sich die Zahl der TeilnehmerInnen auf 273 in der Interventions- und 294 in der Kontrollgruppe

In den Interventionsgruppen wurde das Bewegungsangebot erhöht – pro Woche gab es zwei „Turnstunden“ und an den anderen Tagen jeweils „Bewegungsangebote“ von mindestens 20 Minuten. Falls erforderlich wurden die Außen- und Innenbereiche der Kindertagesstätten bewegungsfreundlicher gestaltet. Daneben gab es Fortbildungsangebote für die teilnehmenden Erzieherinnen (Einzelheiten s. Krombholz 2004a, 2005). In den Kontrolleinrichtungen wurde das normale Bewegungsangebot – eine „Turnstunde“ pro Woche – beibehalten. Um die Wirkung der Intervention zu überprüfen, wurde eine Längsschnittuntersuchung über zwei Jahre mit drei Erhebungen durchgeführt.

Von Bedeutung für die Frage nach dem Einfluss fehlender Werte auf die Untersuchungsergebnisse ist primär, ob sich verbleibende und ausgeschiedene Kinder unterscheiden und ob dies auch für untersuchungsrelevante Gruppen gilt. Daher wurden die Daten der verbleibenden und der ausgeschiedenen Kindern zu Beginn der Studie verglichen und die Verteilung der Ausfälle in unterschiedlichen für die Fragestellung relevanten Gruppen analysiert. Hierbei wurden Faktoren berücksichtigt, die neben dem Treatment – einem verstärkten Bewegungsangebot – einen deutlichen Zusammenhang mit der körperlichen, motorischen und kognitiven Entwicklung im Vorschulalter erwarten lassen. Dies sind das Geschlecht, die Teilnahme an sportlichen Fördermaßnahmen, die soziale Schicht und der Migrationsstatus (Krombholz 1998).

Für die Bewertung der Interventionsmaßnahmen in Kindertagesstätten ist es nicht nur von Bedeutung, dass die Probanden der Interventions- und der Kontrollgruppe sich bei der Eingangserhebung nicht wesentlich unterscheiden, sondern auch, ob die Ausfälle von Teilnehmer/innen zufällig erfolgen oder bestimmte Leistungsgruppen betreffen. Sollten in der Kontrollgruppe vermehrt leistungsfähige Kinder ausscheiden, wäre ein möglicher positiver Effekt der Fördermaßnahmen eventuell darauf zurückzuführen.

Bei der Analyse, ob der Ausfall von teilnehmenden Kindern zufällig erfolgt, können zwei Fragestellungen unterschieden werden:

- Scheiden überwiegend leistungsstärkere oder leistungsschwächere Kinder aus und gibt es

einen Zusammenhang zwischen dem Ausscheiden und Faktoren, von denen bekannt ist, dass sie die körperliche, motorische und kognitive Entwicklung im frühen Kindesalter beeinflussen?

- Sind die Interventions- und die Kontrollgruppe gleichermaßen von Ausfällen betroffen und sind die Ausfälle von leistungsschwächeren Kindern in beiden Gruppen gleich?

Antworten auf die erste Frage sind von genereller Bedeutung für die Planung, Durchführung und Auswertung epidemiologischer Studien zur kindlichen Entwicklung, Antworten auf die zweite Frage haben Konsequenzen für die Durchführung und Bewertung von Interventionsstudien im Kindesalter.

Aufgrund der Ergebnisse einer Untersuchung zum Problem „Missing Data“ im Grundschulalter (Krombholz 1985) wird erwartet, dass die verbleibenden Kinder in der Längsschnittstudie:

- in privilegierten Lebensverhältnissen aufwachsen,
- vermehrt Angebote von Sportvereinen nutzen,
- eher oberen sozialen Schichten angehören,
- einen geringeren Migrantenanteil aufweisen und
- eine höhere motorische und kognitive Leistungsfähigkeit besitzen.

## 1 Methode

Es gab drei Erhebungszeitpunkte, zu Beginn der Intervention und am Ende des 1. und des 2. Kindergartenjahres. Für die folgende Analyse zur Bedeutung der fehlenden Beobachtungen werden jedoch nur die Ergebnisse der Eingangs- und der Abschlusserhebung herangezogen. Die Erhebung der Daten erfolgte in den Kindertagesstätten, erfreulicherweise gaben 97 Prozent der Eltern ihre Einwilligung, dass ihr Kind an der Untersuchung teilnehmen durfte.

An jedem der drei Erhebungszeitpunkten wurden verschiedene Daten der teilnehmenden Kinder erhoben, die die Bereiche körperliche Entwicklung (Körperlänge, Gewicht, Body Mass Index BMI, Hautfältendicke), motorische Leistungen (Testbatterie zur Erfassung motorischer Leistungen im Vorschulalter MoTB 3-7, Krombholz 2011, diese erfasst die Bereiche Körperkoordination, Fitness und Feinmotorik) und kognitive Leistungen (Konzentration, Wortschatz, Intelligenz) umfassten.

In Tabelle 1 sind die verwendeten Testverfahren zusammengestellt, eine ausführliche Beschreibung dieser Verfahren findet sich in Krombholz (2005, S. 35-45). Angaben zu familialen und ökologischen Lebensbedingungen der Kinder wurden durch einen Elternfragebogen erhoben, der u. a. Fragen nach der sozialen Schicht, den Wohnverhältnissen, der Muttersprache (Migrationsstatus), dem Gesundheitszustand und der sportlichen Betätigung umfasst. Das Ausmaß der sportlichen Betätigung wurde mit der Frage ermittelt, ob das Kind Angebote von Sportvereinen nutzt (nie, gelegentlich, regelmäßig), die Zuordnung zur sozialen Schicht (Oberschicht / obere Mittelschicht, mittleren Mittelschicht, untere Mittelschicht / Unterschicht) erfolgte anhand vorgegebener Berufe, die Bestimmung des Migrationsstatus erfolgte anhand der Frage nach der Muttersprache des Kindes (deutsch vs. nicht deutsch). Diesen Elternfragebogen beantworteten

allerdings nur 64 Prozent der Eltern<sup>2</sup>.

Bei der Analyse eines möglichen Einflusses von fehlenden Beobachtungen auf die Ergebnisse wurde nicht nur geprüft, ob zwischen verbleibenden und ausgeschiedenen Kindern generell Unterschiede bestehen, sondern auch, ob sich Zusammenhänge mit den Versuchsbedingungen (Interventions-, Kontrollgruppe), dem Geschlecht, der sozialen Schicht, der sportlichen Betätigung der Kinder in Sportvereinen und dem Migrationsstatus nachweisen lassen.

### **Tabelle 1: Übersicht über die Variablen und Testverfahren bei den drei Erhebungen**

#### ***Körperliche Merkmale***

Geschlecht

Alter (Monate)

Körperlänge (cm)

Körpergewicht (kg)

Hautfältendicke (Oberarm, Bauch, Rücken; Mittelwert) (mm)

Body Mass Index BMI = Körpergewicht (kg) / Körpergröße (m)<sup>2</sup>

#### ***Motorik***

Motorik-Testbatterie 3-7 (MoTB 3-7, Krombholz 2011) erfasst:

##### ***Körperkoordination***

Balancieren Vorwärts (Anzahl, max. 24)

Balancieren Rückwärts BR und Seitliches Hin- und Herspringen SH (Körperkoordinationstest für Kinder KTK, Kiphard & Schilling 2007, Anzahl)

Einbeiniges Hüpfen rechter / linker Fuß (Anzahl, max. 20)

##### ***Fitness***

Standweitsprung (cm)

Halten an der Reckstange (sec, max. 30)

Pendellauf - Strecke: 4 x 4m (sec)

##### ***Feinmotorik***

Leistung der rechten und linken Hand (Punktiertest und Leistungsdominanztest PTK - LDT, Schilling 2009, Anzahl)

#### ***Kognitive Leistungen***

Konzentration (Frankfurter Tests für Fünfjährige – Konzentration FTF-K, Raatz & Möhling 1971, Anzahl)

Wortschatz (Peabody Picture Vocabulary Test PPVT, Bondy, Cohen, Eggert & Lür 1975, Anzahl)

Intelligenz (Grundintelligenztests Skala 1 CFT 1, Weiß & Osterland 1997, 3 Untertests, T-Wert)

---

<sup>2</sup> Das Fehlen des Elternfragebogens galt nicht als Ausschlusskriteriums für vollständige Fälle. Die Vollständigkeit bezieht sich auf die am Kind erhobenen Daten.

## 2 Ergebnisse

### 2.1 Kinder mit vollständigen und unvollständigen Datensätzen

In Tabelle 2 findet sich, getrennt für Jungen und Mädchen, die Anzahl der Kinder, die für die Auswertung bei der Eingangs- und der Ausgangsuntersuchung zur Verfügung standen. Der „Schwund“ beträgt fast 30 Prozent, Jungen und Mädchen waren gleichermaßen von Ausfällen betroffen (statistisches Verfahren:  $\chi^2$ -Test:  $\chi^2 = .01$ ,  $df = 1$ ,  $p = .51$ ).

**Tabelle 2: Anzahl der an der Eingangs- und der Abschlussuntersuchungen teilnehmenden Kinder**

	Eingangserhebung		Abschlusserhebung	
	N	%	N	%
Jungen	293	51.7	206	36.3
Mädchen	274	48.3	193	34.0
Gesamt	567	100.0	399	70.4

In Tabelle 3 sind die Ergebnisse, jeweils Mittelwert M und Standardabweichung SD, bei der Eingangs- und der Abschlusserhebung von Kindern, mit vollständigen Datensätzen und von Kindern mit unvollständigen Datensätzen gegenübergestellt. Dies sind Kinder, die ausgeschieden sind oder von denen einzelne Daten nicht erhoben werden konnten. Die statistische Prüfung der Unterschiede zwischen den Gruppen für jede Variable erfolgte mittels univariater Varianzanalysen (ANOVAs), angegeben sind der F-Wert, die Irrtumswahrscheinlichkeit p und die praktische Bedeutsamkeit (partielles  $\eta^2$ ). Signifikante Unterschiede zwischen den beiden Gruppen fanden sich lediglich beim Körpergewicht, die praktische Bedeutsamkeit (Effektstärke) war aber auch hier wie bei allen überprüften Mittelwertunterschieden nur sehr gering.



**Tabelle 3: Vergleich der Ergebnisse der Kinder, die an der Abschlusserhebung teilnahmen (Vollständige Fälle, N = 399) mit denen der ausgeschiedenen Kinder („MISSING“, N = 168) bei der Eingangserhebung: jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der Varianzanalysen (F-Wert, Irrtumswahrscheinlichkeit p und Praktische Bedeutsamkeit Eta<sup>2</sup>)**

	Vollständige Fälle		MISSING				
	MEAN	SD	MEAN	SD	F	p	Eta <sup>2</sup>
<i>Köperl. Merkmale</i>							
Alter (Monate)	52.9	7.2	54.1	8.0	3.1	.08	.01
Körperläng (cm)	107.5	5.9	108.3	6.7	1.9	.17	.01
Gewicht (kg)	18.1	2.5	18.4	2.9	3.9	.05	.01
Hautfalte (mm)	4.6	1.8	4.9	2.0	3.2	.07	.01
BMI	15.6	1.5	15.9	1.7	2.7	.10	.01
<i>Motorik</i>							
Bal. Vorw. (Anzahl)	12.4	5.6	12.9	5.6	.8	.37	.01
Bal. Rückw. (Anzahl)	11.3	11.0	12.9	12.0	2.0	.16	.01
Seitl. HH-Spr. (Anzahl)	19.6	6.8	21.3	7.8	5.9	.02	.01
Standweit (cm)	75.5	21.3	78.2	20.2	1.7	.19	.01
Pendellauf (sec)	10.0	1.7	10.9	1.9	.1	.72	.01
Hüpfen Vorw R (Anzahl)	9.6	7.7	10.6	8.0	1.8	.19	.01
Hüpfen Vorw L (Anzahl)	8.7	7.2	9.9	7.7	2.8	.09	.01
Halten (sec)	19.1	8.9	18.7	9.2	.2	.64	<.01
Handgeschick. R (Anzahl)	20.1	16.2	22.5	16.4	2.2	.14	.01
Handgeschick. L (Anzahl)	6.5	7.8	7.6	9.6	1.7	.20	.01
Handgeschickl. G (Anzahl)	26.6	20.9	30.3	22.0	2.8	.20	.01
<i>Kognitive Leistungen</i>							
Konz. Test (Anzahl)	19.8	7.5	20.6	7.5	1.0	.27	.01
Wortschatz (Anzahl)	33.9	16.8	33.0	17.2	.2	.33	<.01
Intelligenz (T-Wert)	54.8	8.3	54.4	8.5	.8	.78	<.01

Auch multivariate Analysen (MANOVAs) getrennt für die körperlichen Merkmale, die Sporttests und die feinmotorischen und die kognitiven Leistungen konnten keine signifikanten Zusammenhänge mit dem Faktor Verbleib in der Stichprobe nachweisen (Tabelle 4, 5, 6). In diesen Analysen wurde das Geschlecht als zusätzlicher Faktor berücksichtigt, der sich als statistisch signifikant und praktisch bedeutsam erwies; die Wechselwirkung Geschlecht und Verbleib war jedoch in keinem Fall bedeutsam. – Im Gegensatz zu den univariaten (paarweiser Ausschluss) werden bei den multivariaten Analysen vollständige Fälle herangezogen (fallweiser Ausschluss).

**Tabelle 4: Ergebnisse MANOVA****Faktoren: teilnehmende und ausgeschiedene Kinder (Missing) und Geschlecht****AV: Körperliche Merkmale bei der Eingangserhebung (Größe, Gewicht, Hautfalten, BMI)**

Effekt	F	df	p	Eta <sup>2</sup>
Missing	1.2	4; 466	.32	.01
Geschlecht	18.3	4; 466	<.01	.14
Missing * Geschl.	.2	4; 466	.98	.01

**Tabelle 5: Ergebnisse MANOVA****Faktoren: teilnehmende und ausgeschiedene Kinder (Missing) und Geschlecht****AV: Leistungen im Motoriktest MoTB 3-7 bei der Eingangserhebung**

Effekt	F	df	p	Eta <sup>2</sup>
Missing	1.2	9; 416	.33	.02
Geschlecht	14.2	9; 416	<.01	.23
Missing * Geschl.	.4	9; 416	.93	.01

**Tabelle 6: Ergebnisse MANOVA****Faktoren: teilnehmende und ausgeschiedene Kinder (Missing) und Geschlecht****AV: Leistungen in den kognitiven Tests (Konzentration, Wortschatz, Intelligenz)**

Effekt	F	df	p	Eta <sup>2</sup>
Missing	1.1	3; 380	.36	.01
Geschlecht	3.8	3; 380	<.01	.03
Missing * Geschl.	.6	3; 380	.61	.01

## 2.2 Fallweiser oder paarweiser Ausschluss

Ob bei den Berechnungen alle vorliegenden Daten (paarweiser oder itemweiser Fallausschluss, größeres N) oder nur die „vollständigen Fälle“ (fallweiser Ausschluss, kleineres N) herangezogen werden, hat keinen signifikanten Einfluss auf die Mittelwerte und die Korrelationskoeffizienten. Bei keiner der Variablen waren die Mittelwertunterschiede zwischen paarweisem und vollständigem Ausschluss signifikant, beide Vorgehensweisen führen zu fast identischen Korrelationskoeffizienten (Mittelwerte Tab. 7, Korrelationen Tab. 8). Es bestand allerdings die Tendenz, dass die Streuungen bei den meisten Variablen geringfügig größer sind, wenn nicht nur die vollständigen Fälle berücksichtigt werden und die Fallzahlen höher sind.

**Tabelle 7: Vergleich der Ergebnisse bei fallweisem (vollständige Fälle) und paarweisem Ausschluss, jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der t-Tests (Irrtumswahrscheinlichkeit p). Die praktische Bedeutsamkeit (partielles  $\eta^2$ ) bei allen Vergleichen ist kleiner als .01).**

	Paarweise MEAN	SD	Fallweise MEAN	SD	p
Körperlänge (cm)	107.6	6.2	108.3	5.8	.10
Gewicht (kg)	18.3	3.1	18.5	2.9	.34
Hautfalte (mm)	4.7	1.8	4.7	1.8	.99
BMI	15.7	1.6	15.7	1.5	.99
<i>Motorik</i>					
Bal. Vorw. (Anzahl)	12.5	5.6	12.7	5.3	.60
Bal. Rückw. (Anzahl)	11.7	11.4	12.2	11.2	.65
Seitl. HH-Spr. (Anzahl)	20.1	7.1	20.3	7.1	.63
Standweit (cm)	76.3	21.0	77.0	20.3	.63
Pendellauf (sec)	10.9	1.8	10.8	1.6	.40
Hüpfen Vorw R (Anzahl)	9.9	7.8	10.1	7.8	.72
Hüpfen Vorw L (Anzahl)	9.0	7.3	8.8	7.1	.69
Halten (sec)	20.0	9.0	19.3	8.8	.26
Handgeschicklichkeit (Ges)	27.6	21.3	27.8	21.7	.89
<i>Kognitive Leistungen</i>					
Konz. Test (Anzahl)	20.0	7.5	20.4	7.6	.45
Wortschatz (Anzahl)	33.6	16.9	35.4	16.5	.13
Intelligenz (T-Wert)	54.7	8.5	55.0	8.6	.62

**Tabelle 8: Vergleich der Korrelationen ausgewählter Variablen bei paarweisem (oberhalb) und fallweisem Ausschluss (Complete Cases, unterhalb der Diagonalen)**

	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>
<i>1</i> Körperlänge	-	.79	.28	.37	.38	-.38	.30	.33	-.02
<i>2</i> Gewicht	.80	-	.14	.20	.22	-.20	.20	.15	-.08
<i>3</i> Bal. Rückw	.28	.11	-	.55	.38	-.41	.38	.22	.04
<i>4</i> Seidl.HH-Sp	.35	.17	.55	-	.60	-.56	.31	.24	.05
<i>5</i> Standweit	.41	.23	.44	.63	-	-.59	.14	.29	.02
<i>6</i> Pendellauf	-.36	-.21	-.39	-.57	-.63	-	-.30	-.31	-.03
<i>7</i> Handgeschickl	.31	.17	.63	.30	.17	-.28	-	.17	.08
<i>8</i> Wortschatz	.29	.12	.21	.26	.28	-.28	.17	-	.20
<i>9</i> Intelligenz	-.03	-.08	.06	.04	.05	-.01	.09	.21	-

## 2.3 Unabhängige Variablen

### 2.3.1 Soziale Schicht

Die Verteilung auf die sozialen Schichten für Jungen und Mädchen kann Tabelle 9 entnommen werden, der Zusammenhang zwischen Schicht und Verbleib in der Stichprobe Tabelle 10. Da sich die hier als Mittelschicht und Oberschicht bezeichneten Gruppen kaum unterschieden, wurden sie für diese Tabelle zusammengefasst. Der Zusammenhang zwischen sozioökonomischem Status und dem Verbleib in der Stichprobe war nicht bedeutsam ( $\chi^2 = .07$ ,  $df = 1$ ,  $p = .79$ ).

**Tabelle 9: Verteilung auf die sozialen Schichten für Jungen und Mädchen einschließlich der Kinder ohne Angabe zur Schicht (KA)**

	Jungen		Mädchen		Gesamt	
	N	%	N	%	N	%
KA	92	16.2	100	17.6	192	33.9
US	100	17.6	77	13.6	177	31.2
MS	72	12.7	77	13.6	149	26.3
OS	29	5.1	20	3.5	49	8.6
Ges	293	51.7	274	48.3	567	100

**Tabelle 10: Sozioökonomischer Status (Unterschicht US vs. Mittel- und Oberschicht MS) und Teilnahme an der Eingangs- und Abschlusserhebung**

	Eingangserhebung		Abschlusserhebung	
	N	%	N	%
US	177	47.2	153	48.0
MS	198	52.8	166	52.0
Gesamt	375	100.0	319	100.0

Die Ergebnisse der Kinder der unterschiedlichen sozialen Schichten sind in Tabelle 11 dargestellt, jeweils Mittelwert, Standardabweichung und Irrtumswahrscheinlichkeit  $p$  für univariate Varianzanalysen. In dieser Tabelle werden die ausgeschiedenen und die Kinder ohne Angabe zur Schicht zusammengefasst (MISSING) und den Kindern der Unterschicht (US) und der Mittel- und Oberschicht (MS/OS) gegenübergestellt. Zusammenhänge mit dem Faktor Schicht zeigten sich bei der Hautfältendicke und dem BMI. Post-hoc-Tests (LSD) zeigten signifikante Unterschiede zwischen den Kindern ohne Angabe zur Schicht und der Mittel- und der Oberschicht sowohl für die Hautfältendicke als auch den BMI. Kinder unterer sozialer Schichten wiesen höhere Werte auf als Kinder der Mittel- und Oberschicht, die Unterschiede zwischen Kindern ohne Angabe zur Schicht und den Unterschichtkindern waren nicht bedeutsam. Keine statistisch bedeutsamen Zusammenhänge konnten zwischen dem Faktor Schicht und den Motoriktests nachgewiesen werden.

Dagegen war dieser Zusammenhang für die kognitiven Leistungen Wortschatz und Intelligenz bedeutsam (VA:  $F = 9.4$ ,  $df = 3$ ; 531,  $p = < .001$ ,  $F = 11.3$ ,  $df = 3$ ; 501,  $p = < .001$ ): Kinder oberer sozialer Schichten erreichten höhere Leistungen als Kinder unterer sozialer Schichten und Kinder ohne Angabe zur Schicht. Aufgrund der Post-hoc-Tests (LSD) war beim Wortschatz und bei der Intelligenz lediglich der Unterschied zwischen Mittel- und Oberschicht nicht bedeutsam. Die Leistungen der Kinder ohne Angaben zur Schicht lagen jeweils unter der der Kinder der Unterschicht.

**Tabelle 11: Vergleich der Ergebnisse der sozialen Schichten (US, MS und OS, ausgeschiedene und Kinder ohne Angabe zur Schicht MISSING) bei der Eingangserhebung: jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der Varianzanalysen (Irrtumswahrscheinlichkeit p)**

	MISSING		US		MS / OS		
	MEAN	SD	MEAN	SD	MEAN	SD	p
<i>Körperl. Merkmale</i>							
Alter (Monate)	53.6	7.7	53.2	7.3	53.0	7.4	.74
Körperläng (cm)	107.6	6.8	107.1	6.1	108.2	5.7	.31
Gewicht (kg)	18.7	3.6	18.2	3.1	18.1	2.6	.18
Hautfalte (mm)	5.0	2.2	4.8	1.9	4.4	1.3	.01
BMI	16.0	1.8	15.8	1.6	15.4	1.2	.01
<i>Motorik</i>							
Bal. Vorw. (Anz)	12.5	5.6	12.6	5.9	12.5	5.4	.97
Bal. Rückw. (Anz)	11.8	11.8	11.4	10.7	12.1	11.6	.85
Seitl. HH-Spr. (Anz)	20.2	6.8	19.9	7.0	20.2	7.6	.87
Standweit (cm)	76.3	20.4	76.3	21.3	76.3	21.4	.99
Pendellauf (sec)	11.1	1.7	10.9	209	10.8	1.5	.46
Hüpf Vorw R (Anz)	10.2	7.9	9.1	7.8	10.2	7.7	.34
Hüpf Vorw L (Anz)	9.6	7.5	9.0	7.3	8.5	7.2	.35
Halten (sec)	17.7	9.0	19.3	9.1	19.9	8.6	.08
Handgesch. R (Anz)	22.1	15.7	19.3	16.4	20.7	16.5	.33
Handgesch. L (Anz)	7.3	8.7	6.3	7.9	6.9	8.4	.57
<i>Kognitive Leistungen</i>							
Konz. Test (Anz)	20.3	6.8	19.2	7.8	20.5	7.8	.25
Wortschatz (Anz)	31.2	17.5	29.8	17.3	38.3	15.0	<.01
Intelligenz (T-Wert)	51.7	7.43	53.7	8.6	57.1	8.3	<.01

### 2.3.2 Sportliche Betätigung

Der Zusammenhang zwischen der sportlichen Betätigung der Kinder und dem Verbleib in der Stichprobe ist in Tabelle 12 dargestellt, die Gruppen der regelmäßig und gelegentlich Teilnehmenden wurden zusammengefasst. Es konnte kein Zusammenhang zwischen der Teilnahme an sportlichen Angeboten und dem Verbleib in der Stichprobe nachgewiesen werden ( $\chi^2 = .05$ ,  $df = 1$ ,  $p = .82$ ).

**Tabelle 12: Sportliche Betätigung (regelmäßig und gelegentlich vs. nie) und Teilnahme an der Eingangs- und Abschlusserhebung**

	Eingangserhebung		Abschlusserhebung	
	N	%	N	%
nie	197	52.3	168	52.5
gel. / regelm.	180	47.7	152	47.5
Gesamt	377	100.0	320	100.0

Die Ergebnisse der Überprüfung des Zusammenhangs zwischen dem Besuch von Übungsangeboten von Sportvereinen und den körperlichen Merkmalen und den überprüften motorischen und kognitiven Leistungen ist in Tabelle 13 dargestellt. Es wurden auch die Kinder berücksichtigt, von denen keine Angaben hierzu vorlagen. Es zeigte sich ein bedeutsamer Zusammenhang mit körperlichen Merkmalen (Gewicht, BMI) und allen sportlichen Leistungen mit Ausnahme des Balancierens Vorwärts und Rückwärts.

Werden die einzelnen Gruppen miteinander verglichen (stat. Verfahren Post-hoc Tests LSD), so fanden sich signifikante Unterschiede zwischen Kindern, die Übungsstunden von Sportvereinen besuchen und denen, die dies nicht tun, bei der Körperlänge, dem Seitlichen Hin- und Herspringen, dem Standweitsprung, dem Pendellauf, beim Hüpfen auf dem rechten und dem linken Bein, bei der Geschicklichkeit der rechten Hand, bei der Konzentration und dem Wortschatz. Im Vergleich zu den Fällen mit Missing Data zeigten sich signifikante Unterschiede beim BMI, beim Pendellauf, beim Halten, bei der Konzentration, beim Wortschatz und bei der Intelligenz. Hierbei sind Kinder, die sportliche Übungsstunden besuchen, überlegen. Zwischen Kindern die keine Übungsstunden besuchen und denen mit Missing Data fanden sich signifikante Unterschiede beim Gewicht, bei den Hautfalten, beim BMI, beim Hüpfen rechts und links, bei der Handgeschicklichkeit und bei der Intelligenz. Bei den motorischen Leistungen sind die Kinder mit Missing Data überlegen, bei der Intelligenz die Kinder, die keine Übungsstunden besuchen.

**Tabelle 13: Vergleich der Ergebnisse der Kinder, die an sportlichen Angeboten von Sportvereinen teilnahmen (nie, gelegentlich oder regelmäßig, ausgeschiedene und Kinder ohne Angabe zur sportlichen Betätigung MISSING) bei der Eingangserhebung: jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der Varianzanalysen (Irrtumswahrscheinlichkeit p)**

	MISSING MEAN	SD	nie MEAN	SD	gel./regelm. MEAN	SD	p
<i>Körperl. Merkmale</i>							
Alter (Monate)	53.7	7.7	52.4	7.2	53.7	7.6	.17
Körperläng (cm)	107.7	6.8	107.0	5.7	108.4	6.0	.08
Gewicht (kg)	18.7	3.7	17.9	2.7	18.5	3.0	.04
Hautfalte (mm)	5.0	2.2	4.5	1.6	4.7	1.7	.08
BMI	16.0	1.8	15.5	1.4	15.6	1.5	.02
<i>Motorik</i>							
Bal. Vorw. (Anz)	12.5	5.6	12.1	5.7	12.9	5.4	.40
Bal. Rückw. (Anz)	11.9	11.7	10.9	11.0	12.5	11.4	.41
Seitl. HH-Spr. (Anz)	20.2	6.8	18.9	7.4	21.2	7.2	.02
Standweit (cm)	76.7	20.2	73.1	21.6	79.2	20.9	.03
Pendellauf (sec)	11.1	1.7	11.1	2.0	10.7	1.5	.05
Hüpf Vorw R (Anz)	10.4	7.9	8.6	7.6	10.7	7.8	.03
Hüpf Vorw L (Anz)	9.9	7.5	7.4	6.9	9.8	7.4	<.01
Halten (sec)	17.6	9.1	19.0	8.8	20.2	8.8	.03
Handgesch. R (Anz)	22.1	15.7	18.1	16.6	22.2	16.2	.04
Handgesch. L (Anz)	7.3	8.7	6.7	8.2	6.7	8.2	.70
<i>Kognitive Leistungen</i>							
Konz. Test (Anz)	20.3	6.8	18.8	7.7	21.1	7.8	.02
Wortschatz (Anz)	32.0	17.4	31.4	16.3	37.4	16.5	<.01
Intelligenz (T-Wert)	51.6	7.3	54.8	9.0	56.4	8.1	<.01



### 2.3.3 Migrationsstatus

Der Zusammenhang zwischen Migrationsstatus und Verbleib in der Stichprobe findet sich in Tabelle 14; der Zusammenhang ist nicht bedeutsam ( $\chi^2 = .73$ ,  $df = 1$ ,  $p = .39$ ).

**Tabelle 14: Migrationsstaus und Teilnahme an der Eingangs- und Abschlusserhebung**

	Eingangserhebung		Abschlusserhebung	
	N	%	N	%
deutsch	287	75.9	247	76.7
nicht deutsch	91	24.1	75	23.3
Gesamt	378	100.0	322	100.0

Die Ergebnisse der Überprüfung des Zusammenhangs zwischen dem Migrationsstatus und den fehlenden Fällen (MISSING) für die körperlichen Merkmalen und den motorischen und den kognitiven Leistungen kann Tabelle 15 entnommen werden. Es zeigt sich ein bedeutsamer Zusammenhang mit körperlichen Merkmalen (Hautfalten, BMI), dem Item Halten der Motorik-Testbatterie, dem Wortschatz und der Intelligenz.

Werden die einzelnen Gruppen miteinander verglichen (stat. Verfahren Post-hoc Tests LSD), so fanden sich signifikante Unterschiede zwischen Kindern ohne und mit Migrationsstatus beim BMI, beim Wortschatz und der Intelligenz. Kinder mit Missing Data unterschieden sich von den Kindern mit Migrationshintergrund beim Wortschatz und von den deutschen Kindern bei der Hautfalte, beim BMI, beim Halten, beim Wortschatz und bei der Intelligenz.

**Tabelle 15: Vergleich der Ergebnisse der Kinder mit (nicht deutsch) und ohne Migrationshintergrund (deutsch) und ausgeschiedene und Kinder ohne Angabe zum Migrationsstatus MISSING) bei der Eingangserhebung: jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der Varianzanalysen (Irrtumswahrscheinlichkeit p)**

	MISSING		deutsch		nicht deutsch		p
	MEAN	SD	MEAN	SD	MEAN	SD	
<i>Körperl. Merkmale</i>							
Alter (Monate)	53.6	7.7	53.2	7.3	52.5	7.4	.72
Körperläng (cm)	107.7	6.9	108.0	5.5	106.6	6.8	.18
Gewicht (kg)	18.7	3.6	18.2	2.7	18.2	3.3	.23
Hautfalte (mm)	4.9	2.2	4.5	1.5	4.8	1.8	.05
BMI	16.0	1.8	15.5	1.4	15.9	1.5	.01
<i>Motorik</i>							
Bal. Vorw. (Anz)	12.4	5.6	12.8	5.6	11.7	5.6	.29
Bal. Rückw. (Anz)	11.8	11.8	12.1	11.3	10.5	10.7	.55
Seitl. HH-Spr. (Anz)	20.2	6.8	20.4	7.3	19.1	7.2	.40
Standweit (cm)	76.4	20.2	77.1	20.2	73.4	20.7	.46
Pendellauf (sec)	11.1	1.7	10.8	1.6	11.2	2.2	.14
Hüpf Vorw R (Anz)	10.3	7.9	9.8	7.7	9.3	7.9	.62
Hüpf Vorw L (Anz)	9.8	7.5	8.4	7.2	9.4	7.4	.17
Halten (sec)	17.7	9.1	19.8	9.0	18.8	8.4	.06
Handgesch. R (Anz)	22.1	15.6	19.9	16.6	20.8	16.5	.45
Handgesch. L (Anz)	7.3	8.7	6.9	8.4	5.7	7.1	.39
<i>Kognitive Leistungen</i>							
Konz. Test (Anz)	20.2	6.7	20.4	7.9	18.5	7.5	.16
Wortschatz (Anz)	31.9	17.4	38.2	21.0	15.1	.33	.01
Intelligenz (T-Wert)	51.8	7.4	56.6	8.2	52.1	9.1	.01

### 2.3.4 Interventions- und Kontrollgruppe und Wirksamkeit der Fördermaßnahmen

Die wesentliche Fragestellung der Studie war, ob ein verstärktes Bewegungsangebot in Kindertagesstätten die motorische Leistungsfähigkeit der Kinder steigern kann. Dies konnte eindeutig belegt werden (Tabelle 16) und die Leistungssteigerungen im Untersuchungszeitraum waren erwartungsgemäß in der Interventions- stärker als in der Kontrollgruppe (Effekt Treatment \* Messwiederholung). Die Leistungen von Jungen und Mädchen unterschieden sich bedeutsam, wobei es Übungen gab, in denen Mädchen und Übungen, in denen Jungen höhere Leistungen erzielten. Die Intervention war bei den Mädchen allerdings effektiver als bei den Jungen. Ebenfalls zeigte sich die Intervention als wirkungsvoller bei den wenig leistungsfähigen und übergewichtigen Kindern gegenüber leistungsfähigeren und normalgewichtigen Kindern (eine ausführliche Darstellung dieser Befunde s. Krombholz 2012, 2016). Diese Ergebnisse sollen hier jedoch unter dem Aspekt Verbleib in der Stichprobe (Missing Data) ausführlich analysiert und diskutiert werden, wobei geprüft wird, ob ein Ausscheiden der Kinder die vorliegenden Ergebnisse beeinflusst haben könnte.

**Tabelle 16: Ergebnisse MANOVA mit Messwiederholung**  
**Faktoren: Treatment (Interventions-, Kontrollgruppe), Geschlecht, Messwiederholung (MW)**  
**AV: Leistungen im Motoriktest MoTB 3-7**

Effekt	F	df	p	Eta <sup>2</sup>
Treatment	2.0	9; 296	.04	.06
Geschlecht	15.4	9; 296	.01	.32
Treat * Geschl	.7	9; 296	.77	.02
MW	321	9; 296	<.01	.10
MW * Geschl	.9	9; 296	.57	.03
MW * Treat	2.7	9; 296	<.01	.08
MW * Ges * Treat	1.3	9; 296	.23	.04

Mehr Eltern der Kinder der Interventionsgruppe haben den Elternfragebogen EFB ausgefüllt (56 vs. 44 %,  $\chi^2 = 27.4$ ,  $df = 1$ ,  $p < .01$ ). Wird geprüft, in welchem Ausmaß die Interventions- und die Kontrollgruppe von Ausfällen betroffen sind, so zeigt sich, dass mehr Kinder der Kontroll- als der Interventionsgruppe im Untersuchungszeitraum aus der Stichprobe ausgeschieden sind (Tab. 17,  $\chi^2 = 6.5$ ,  $df = 1$ ,  $p = .01$ ).

**Tabelle 17: Versuchsbedingung (Interventions- und Kontrollgruppe) und Teilnahme an der Eingangs- und Abschlusserhebung**

	Eingangserhebung		Abschlusserhebung	
	N	%	N	%
Intervention	273	48.1	206	51.6
Kontroll	294	51.9	193	48.4
Gesamt	567	100.0	411	100.0

Die Ergebnisse der Analyse der Ausfälle in der Interventions- und der Kontrollgruppe sind in Tabelle 18 dargestellt. Interventions- und Kontrollgruppe unterschieden sich bei der Eingangserhebung nicht bedeutsam hinsichtlich der erfassten körperlichen Merkmale und der motorischen und kognitiven Leistungsfähigkeit. Es besteht allerdings zumindest die Tendenz, dass eher körperlich leistungsfähigere Kinder aus der Interventionsgruppe ausschieden. Post hoc Analysen zeigten für die Interventionsgruppe deutliche Unterschiede zwischen den in der Studie verbleibenden und den ausgeschiedenen Kindern. Signifikante Unterschiede finden sich beim Gewicht, bei den Hautfalten, beim BMI, bei den sportlichen Leistungen Balancieren Rückwärts, beim Seitlichen Hin- und Herspringen, beim Standweitsprung, beim Hüpfen vorwärts rechts und links sowie bei der Geschicklichkeit der rechten Hand; die motorischen Leistungen der ausgeschiedenen übertrafen die der verbleibenden Kinder.

Dagegen konnte für die Kontrollgruppe bei keiner der erhobenen Variablen Unterschiede zwischen ausgeschiedenen und verbliebenen Kindern nachgewiesen werden.

**Tabelle 18: Vergleich der Ergebnisse von Kindern mit vollständigen Daten (Vollst.) und ausgeschiedenen Kindern (MISSING) in der Interventions- und Kontrollgruppe bei der Eingangserhebung: jeweils Mittelwert (MEAN), Standardabweichung (SD) und Ergebnisse der Varianzanalysen (Irrtumswahrscheinlichkeit p)**

	Interventionsgruppe				Kontrollgruppe				p
	Vollst.		MISSING.		Vollst.		MISSING		
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	
<i>Körperl. Merkmale</i>									
Alter (Monate)	53.0	7.2	55.4	7.2	52.8	7.2	53.2	8.4	.08
Körperläng (cm)	108.0	5.6	108.8	6.8	106.8	6.2	107.9	6.8	.11
Gewicht (kg)	18.3	2.7	19.3	4.4	18.0	3.1	18.4	3.0	.06
Hautfalte (mm)	4.7	1.6	5.3	2.5	4.6	2.0	4.7	1.5	.08
BMI	15.6	1.4	16.1	2.2	15.6	1.6	15.8	1.3	.23
<i>Motorik</i>									
Bal. Vorw. (Anz)	12.4	5.6	13.9	5.5	12.3	5.6	12.1	5.6	.25
Bal. Rückw. (Anz)	11.5	11.1	15.0	13.1	11.1	11.1	11.3	11.0	.13
Seitl. HH-Spr. (Anz)	19.8	6.9	23.5	7.8	19.4	6.2	19.7	7.4	>.01
Standweit (cm)	75.3	22.6	83.7	19.1	75.8	19.8	74.4	20.2	.03
Pendellauf (sec)	10.9	1.6	10.6	1.6	11.1	1.8	11.1	2.0	.19
Hüpf Vorw R (Anz)	9.5	7.8	12.0	7.9	9.6	7.7	9.6	7.9	.16
Hüpf Vorw L (Anz)	8.5	7.2	10.8	7.6	8.9	7.1	8.3	7.7	.90
Halten (sec)	19.2	8.7	19.2	9.3	19.0	9.1	18.3	9.2	.06
Handgesch. R (Anz)	20.0	16.	24.9	17.4	20.2	15.8	20.9	15.5	.24
Handgesch. L (Anz)	6.2	7.8	8.3	9.6	6.9	7.8	7.2	9.7	.45
Handgesch. G (Anz)	26.2	20.1	33.6	22.3	27.1	20.9	28.0	21.7	.17
<i>Kognitive Leistungen</i>									
Konz. Test (Anz)	20.8	8.1	23.0	7.2	18.6	6.4	18.9	7.3	>.01
Wortschatz (Anz)	34.4	16.5	34.4	17.9	33.1	17.3	32.1	16.8	.72
Intelligenz (T-Wert)	55.0	8.0	55.6	10.7	54.6	8.6	54.2	9.0	.96

## 4 Diskussion

Anhand der Daten einer Studie zum Einfluss motorischer Fördermaßnahmen in Kindertagesstätten wurde geprüft, ob die Ausfälle von Kindern in einer Längsschnittuntersuchung im vorschulischen Alter zufällig erfolgten. An der Studie nahmen mehr 700 Kinder teil, die Kindergärten in München besuchten. Daten zur Entwicklung der Kinder wurden dreimal erhoben: nach Eingewöhnung in den Kindergarten (Eingangserhebung), am Ende des ersten Kindergartenjahres und zum Abschluss der Kindergartenzeit (Abschlusserhebung). Für die hier vorgestellten Ergebnisse wurden allerdings lediglich die Daten der Eingangs- und der Abschlussuntersuchung herangezogen.

Im Verlauf der Studie kam es zu einem erheblichen Ausfall an Probanden, vor allem infolge von Austritten aus dem Kindergarten, vorübergehender Abwesenheit infolge Urlaub der Eltern (der Kindergartenbesuch in Deutschland ist freiwillig) oder Krankheit am vorgesehenen Untersuchungstermin. Zwar war versucht worden, abwesende Kinder nachzuuntersuchen, dies war leider nicht immer möglich. Von den 567 Kindern, die an der Eingangsuntersuchung teilnahmen und für die vollständige Datensätze vorlagen, konnten bei der Abschlusserhebung nach 21 Monaten nur noch von 399, 206 Jungen und 193 Mädchen, alle Daten wie vorgesehen erhoben werden.

Jungen und Mädchen waren gleichermaßen von Ausfällen betroffen. Ebenfalls konnten keine Zusammenhänge zwischen dem Verbleib in der Studie und der sozialen Schicht und der Nationalität (Deutsch vs. Nicht-deutsch) nachgewiesen werden. Bei einer Studie im Grundschulalter waren dagegen mehr Jungen als Mädchen ausgeschieden, was darauf zurückgeführt werden konnte, dass mehr Jungen als Mädchen in der Grundschule scheitern (Krombholz 1985, 1988).

Entgegen der Erwartung, dass überproportional leistungsschwache ProbandInnen aus Längsschnittuntersuchungen im Kindesalter ausscheiden, konnten – anhand der vorliegenden Daten bei der Eingangsuntersuchung – keine Unterschiede zwischen Kindern, die in der Untersuchung verblieben und den ausgeschiedenen Kindern nachgewiesen werden, weder für die körperlichen Merkmale noch die motorischen und die kognitiven Leistungen. Dagegen schieden in einer Untersuchung im Grundschulalter überwiegend Kinder mit geringeren sportlichen und kognitiven Leistungen in den ersten beiden Schuljahren aus (Krombholz 1985). Dies kann damit erklärt werden, dass im Grundschulalter überproportional Kinder mit schlechten Schulleistungen die Schule verlassen oder an vorgesehenen Untersuchungsterminen nicht in der Schule angetroffen werden. Allerdings schieden in der Grundschule bedeutend weniger Kinder aus als im Kindergarten (12% vs. 30%). Wenig überraschend ist der Unterschied zwischen Ausscheidern und Verbleibern höher, wenn der Verbleib mit der Leistungsfähigkeit zusammenhängt. Im Unterschied zum Schulbesuch, bei dem die Leistung eine bedeutsame Rolle spielt und der verpflichtend ist, ist der Besuch einer Kindertagesstätte freiwillig und Eltern können ihre Kinder jederzeit abmelden oder kurzfristig und tageweise auf das Betreuungsangebot verzichten, nicht nur im Krankheitsfall.

Die Berücksichtigung lediglich der vollständigen Fälle (Complete Cases) oder aller verfügbarer Daten (Paarweiser Fallausschluss) hatte keinen Einfluss auf die gefunden Ergebnisse, weder auf die Mittelwerte noch die Korrelationen. Bei keiner der Variablen waren die Mittelwertunterschiede signifikant, keine der ermittelten Korrelationen unterscheidet sich zwischen beiden Berechnungsarten. Dagegen kamen bei einer Studie im Vorschulalter beide Vorgehensweisen zu

unterschiedlichen Ergebnissen (Krombholz 1988). Es besteht allerdings die Tendenz, dass die Streuungen bei den meisten Variablen geringfügig größer sind, wenn nicht nur die vollständigen Fälle berücksichtigt werden.

#### **4.1 Sozioökonomischer Status**

Zusammenhänge zwischen dem Verbleib in der Studie und dem Sozialstatus konnten nicht nachgewiesen werden. Dagegen schieden in den ersten beiden Grundschuljahren mehr Kinder mit geringerem sozialen Status als mit höherem sozialen Status aus, da der Schulerfolg nicht unabhängig vom Sozialstatus ist (Krombholz 1988).

Beim Vergleich der Ergebnisse der Kinder, die unterschiedlichen sozialen Schichten angehören und der ausgeschiedenen und Kindern ohne Angabe zur Schicht zeigten sich keine statistisch bedeutsamen Zusammenhänge zwischen Schicht und den Motoriktests, allerdings bei der Hautfältendicke und dem BMI. Kinder unterer sozialer Schichten wiesen im Vergleich zu Kindern oberer sozialer Schichten höhere Werte auf. Ähnlicher Ergebnisse fanden sich auch in anderen Studien (Haslam & James 2005, WHO 2006, Kurth & Schaffrath Rosario 2007). Die Unterschiede zwischen Kindern ohne Angabe zur Schicht und den Unterschichtkindern waren nicht bedeutsam.

Interessanterweise zeigte sich, dass Kinder von Eltern, die den Fragebogen zum familialen Hintergrund nicht beantworteten oder die keine Angaben zur Schicht machten, sich von denen unterschieden, deren Eltern diese Angaben machten. Entsprechende Ergebnisse fanden sich im Grundschulalter (Krombholz 1988). Obwohl kein Zusammenhang zwischen dem Ausscheiden aus der Studie und der sozialen Schicht nachgewiesen werden konnte, lässt sich am Beispiel der Schichtzugehörigkeit zeigen, dass Kinder, für die keine Angaben zur sozialen Schicht vorlagen, sich hinsichtlich gesundheitsrelevanter Parameter (Übergewicht) und der kognitiven Leistungen, nicht jedoch hinsichtlich der körperlichen Leistungsfähigkeit unterschieden. Erwartungsgemäß wiesen Kinder oberer sozialer Schichten günstigere Werte hinsichtlich des Körpergewichtes und der kognitiven Leistungen im Vergleich zur Unterschicht auf, Kindern ohne Angabe zur Schicht wiesen jedoch noch schlechtere kognitive Leistungen als Kinder der Unterschicht auf. Bei der Analyse des Zusammenhangs Schicht und kognitive Leistungsfähigkeit wäre der Zusammenhang eher unterschätzt worden, da die am wenigsten Leistungsfähigen nicht berücksichtigt wurden, hinsichtlich der Gesundheit und der körperlichen Leistungsfähigkeit wäre die Höhe des Zusammenhangs dagegen nicht betroffen.

#### **4.2 Sportverein**

Zwischen der Teilnahme an Übungsstunden in Sportvereinen und dem Verbleib in der Stichprobe konnte kein Zusammenhang gefunden werden. Kinder, die an Übungsangeboten von Sportvereinen teilnahmen sind in allen motorischen Leistungen mit Ausnahme dem Balancieren Vorwärts und Rückwärts Kindern überlegen, die dies nicht tun. Von ähnlichen Ergebnissen berichtet Schwarz 2014. Im Vergleich zu den Fällen mit Missing Data zeigten sich signifikante Unterschiede beim BMI, beim Pendellauf, beim Halten, bei der Konzentration, beim Wortschatz und bei der Intelligenz. Bezüglich der kognitiven Leistungen ist zu berücksichtigen, dass bereits bei Kindern ein positiver Zusammenhang zwischen Sozialstatus und sportlichem Engagement besteht. Bei den Kindern, die keine Übungsstunden besuchten, sind die Ergebnisse zum Missing Data schwierig zu interpretieren: Signifikante Unterschiede fanden sich beim Gewicht, bei den

Hautfalten, beim BMI, beim Hüpfen rechts und links, bei der Handgeschicklichkeit und bei der Intelligenz. Bei den motorischen Leistungen sind die Kinder mit Missing Data überlegen, bei der Intelligenz die Kinder, die keine Übungsstunden besuchen.

#### **4.3 Migrationsstatus**

Kinder mit und ohne Migrationshintergrund unterschieden sich nicht hinsichtlich des Verbleibs in der Stichprobe. Zwischen dem Migrationsstatus und einzelnen körperlichen Merkmalen (BMI), motorischen Leistungen (Item Halten) und kognitiven Leistungen (Wortschatz, Intelligenz) fanden sich bedeutsame Zusammenhänge, wobei Kinder mit deutscher Muttersprache erwartungsgemäß höhere Leistungen beim Wortschatz erzielten. Kinder mit Missing Data erreichten höhere Leistungen als Kinder mit Migrationshintergrund beim Wortschatz, aber gegenüber deutschen Kindern höhere Werte bei Hautfaldendicke und BMI und geringere Werte beim Halten, beim Wortschatz und bei der Intelligenz.

#### **4.4 Intervention**

Mehr Eltern der Interventions- als der Kontrollgruppe haben den Elternfragebogen ausgefüllt. Dies kann als Hinweis gelten, dass Personen sich stärker engagieren, wenn sie einer wie auch immer definierten herausgehobenen Gruppe angehören. Dies könnte sowohl für die Eltern als auch die beteiligten ErzieherInnen gelten. Da die Elternfragebögen etwa ein Jahr nach Beginn der Studie – aus organisatorischen Gründen – von den beteiligten ErzieherInnen den Eltern mit der Bitte um Beantwortung ausgehändigt wurden, könnten auch ein geringeres Engagement der ErzieherInnen der Kontroll- gegenüber denen der Interventionseinrichtungen eine Rolle gespielt haben.

Mehr Kinder der Kontroll- als der Modellgruppe schieden aus der Untersuchung aus. Dabei betrafen die Ausfälle in der Interventionsgruppe überwiegend körperlich leistungsfähigere Kinder, in der Kontrollgruppe war dies nicht der Fall. Die Leistungssteigerungen im Untersuchungszeitraum waren in der Interventions- stärker als in der Kontrollgruppe, die angestrebte Steigerung der motorischen Fertigkeiten und Fähigkeiten wurde erreicht. Es darf vermutet werden, dass ohne das überproportionale Ausscheiden der leistungsfähigeren Kinder in der Interventionsgruppe der Unterschied zur Kontrollgruppe am Ende der Intervention größer ausgefallen und der ermittelte Effekt der Intervention noch deutlicher wäre. Die Leistungen von Jungen und Mädchen unterschieden sich nicht wesentlich, die Intervention erwies sich bei den Mädchen allerdings als effektiver als bei den Jungen. Ebenfalls zeigte sich die Intervention als wirkungsvoller bei den weniger leistungsfähigen und übergewichtigen Kindern gegenüber leistungsfähigeren und normalgewichtigen Kindern (Einzelheiten s. Krombholz 2005, 2012, 2016).

Eine Verminderung des Körpergewichts bei übergewichtigen und adipösen Kindern konnte allerdings durch die Intervention nicht erreicht werden, was angesichts der geringen Anzahl von solchen Kindern in der Stichprobe, der begrenzten Dauer von etwa zwei Jahren und fehlender Maßnahmen zur Beeinflussung der Ernährung kaum verwunderlich ist (Krombholz 2012). Auch vergleichbare Studien fanden, dass ein erhöhtes Bewegungsangebot in Kindertagesstätten keinen wesentlichen Einfluss auf das Körpergewicht der Kinder hat (Reilly & McDowell (2006).



## 4.5 Fazit

Entgegen den Erwartungen konnte nicht gezeigt werden, dass Kinder, die bis zum Abschluss einer Längsschnittstudie in der Stichprobe verblieben, sich hinsichtlich ihrer Ausgangswerte von den Kindern unterschieden, die innerhalb von zwei Jahren aus der Studie ausschieden. Da sich Fälle mit Missing Data und vollständige Fälle nicht unterschieden, wurde auf das geplante Heranziehen imputationsbasierter Methoden verzichtet. Keine Einflüsse auf die Ergebnisse hatten die Verwendung der Auswertungsmethoden Listenweiser oder Paarweiser Fallausschluss (complete cases).

Ungeachtet dieser Ergebnisse erscheint es empfehlenswert, bei Längsschnittstudien möglichst alle ProbandInnen, die an einem Untersuchungstermin nicht erreicht werden, zu einem späteren Zeitpunkt nachzutesten. Dies wurde in der vorliegenden Studie versucht, stieß jedoch an Grenzen, da das Zeitintervall für die Datenerhebung nicht beliebig verlängert werden kann.

Vor Beginn der statistischen Auswertung von Datensätzen mit Missing Data sollte – wo immer dies möglich ist – geprüft werden, ob sich Fälle mit Missing Data von solchen ohne Missing Data unterscheiden. Zur Prüfung der Art der Ausfälle und zur angemessenen Analyse von Datensätzen mit Missing Data finden sich zahlreiche Hinweise (u.a. Möntmann, Bollinger & Hermann 1983, Little & Rubin 2002, Wirtz 2004, Lüdtke, Robitzsch, Trautwein & Köller 2007, Göthlich 2009, Baltes-Götz 2013).

In jedem Fall sollte der Informationsgehalt eines Datensatzes so vollständig wie möglich ausgeschöpft werden. Bei Verwendung multivariater Verfahren, bei der nur Probanden mit vollständigen Datensätzen berücksichtigt werden, können zusätzlich univariate Verfahren herangezogen werden, bei denen ein Paarweiser Probandenausschluss erfolgen kann. Daneben sollte versucht werden, einen vorliegenden Datensatz sowohl nach Variablengruppen als auch – bei Längsschnittstudien – nach Erhebungszeitpunkten aufzubrechen und Teilstichproben mit jeweiligen „vollständigen Datensätzen“ zu analysieren. Zeigen sich hierbei keine Widersprüche gewinnt die Annahme, dass die Ergebnisse für die ursprünglich vorgesehene Population gelten, an Plausibilität. Bei widersprüchlichen Ergebnissen ist eine Erörterung der Ursachen unerlässlich. Neuere, modellbasierte Verfahren und Multiple Imputationsverfahren versprechen zwar, Probleme bei der Behandlung von Datensätzen zumindest zu verringern, inwieweit dies gelingt, ist jedoch umstritten (Wuthke 2008, Göthlich 2009). Selbst wenn mittels der Verwendung von sophisticated Methoden Probleme bei der Auswertung vermindert werden können, sollten Forscher Alles versuchen, um fehlende Beobachtungen zu vermeiden – auch wenn das nicht immer gelingen dürfte.

## 5 Literatur

- Baltes-Götz, B. (2013). Behandlung fehlender Werte in SPSS und Amos. Trier: ZIMK.  
Verfügbar unter <https://www.uni-trier.de/fileadmin/urt/doku/bfw/bfw.pdf> (aufg. am 26.03.2019).
- Bondy, C., Cohen, R., Eggert, D. & Lüer, G. (1975). Testbatterie für geistig behinderte Kinder TBGB. Weinheim: Beltz.
- Díaz-Ordaz, K., Kenward, M.G, Cohen, A., Coleman, C.L. & Eldridge, S, (2014). Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. *Clinical Trials*. 11 (5), 590-600. DOI: 10.1177/1740774514537136.
- Fiero, M.H., Huang, S., Oren, E. & Bell, M.L. (2016). Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. *Trials*, 17:72.  
DOI 10.1186/s13063-016-1201-z
- Göthlich, S.E. (2009). Zum Umgang mit fehlenden Daten in großzahligen empirischen Erhebungen. In: S. Albers, D. Klapper, U. Konradt, A. Walter & J. Wolf (Hrsg.): *Methodik der empirischen Forschung*. Wiesbaden: Gabler, S. 119-135.
- Haslam, D.W., & James, W. P. (2005) Obesity. *Lancet*, 366, 1197-1209.
- Kiphard, E.J. & Schilling, F. (2007). *Der Körperkoordinationstest für Kinder KTK*. Weinheim: Beltz (2. Aufl.)
- Kirsch A. (1968). Internationaler Standard Fitness Test. *Lehrhilfen für die Leibeserziehung*, 17, 12, 133 – 140
- Krombholz, H. (1985). Fehlen "fehlende Beobachtungen" zufällig? Eine Analyse von "Missing Data" am Beispiel einer Längsschnittuntersuchung. *Zeitschrift für Empirische Pädagogik und Pädagogische Psychologie*, 9, 139-148.
- Krombholz, H. (1988). *Sportliche und kognitive Leistungen im Grundschulalter - Eine Längsschnittuntersuchung*. Frankfurt, Bern, New York, Paris: Lang.
- Krombholz, H. (1998). Theorien, Modelle und Befunde zur motorischen Entwicklung im Kindesalter. *Sportonomics*, 4 (2), 55-76.
- Krombholz, H. (2004a). *Bewegungsförderung im Kindergarten - Ergebnisse eines Modellversuchs*. Teil 1: Ziele, Inhalte und Umsetzung. *Motorik*, 27 (3), 130-137.
- Krombholz, H. (2004b). *Bewegungsförderung im Kindergarten - Ergebnisse eines Modellversuchs*. Teil 2: Ergebnisse. *Motorik*, 27 (4) 166-182.
- Krombholz, H. (2005). *Bewegungsförderung im Kindergarten – Ein Modellversuch*.

Schorndorf, Hofmann.

Krombholz, H. (2011). Testbatterie zur Erfassung motorischer Leistungen im Vorschulalter MoTB 3-7. Beschreibung, Gütekriterien, Normwerte und ausgewählte Ergebnisse. <http://psydok.sulb.uni-saarland.de/volltexte/2011/2711/>.

Krombholz, H. (2012). The Impact of a 20-Month Physical Activity Intervention in Child Care Centers on Motor Performance and Weight in Overweight and Healthy-Weight Preschool Children. *Perceptual and Motor Skills*, 115, 919-932 (DOI: 10.2466/06.10.25.PMS.115.6.919-932).

Krombholz, H. (2016). Die Wirksamkeit motorischer Fördermaßnahmen im Kindergarten - Eine Zusammenfassung von Ergebnissen in Münchener Kindertagesstätten. *Motorik*, 39 (4), 190 – 199 (DOI: 10.2378/motorik2016.art34d).

Kurth, B.-M., & Schaffrath Rosario, A. (2007). Die Verbreitung von Übergewicht und Adipositas bei Kindern und Jugendlichen in Deutschland. Ergebnisse des Kinder und Jugendgesundheits surveys (KiGGS)]. *Gesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 50, 736-743.

Little, R.J.A. & Rubin, D.B. (2002). *Statistical analyses with missing data*. New York: Wiley.

Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58 (2), 103-117.

Möntmann, V.G., Bollinger, G. & Herrmann, A. (1983). Tests auf Zufälligkeit von „Missing Data“. In: H. Wilke et al. (Hrsg.): *Statistik Software in der Sozialforschung*. Berlin: Quorum, S. 87-101.

Raatz, U. & Möhling, K. (1971). *Frankfurter Test für Fünfjährige - Konzentration FTF-K*. Weinheim: Beltz.

Reilly, J. J., & McDowell, Z. C. (2003) Physical activity interventions in the prevention and treatment of pediatric obesity: systematic review and critical appraisal. *Proceedings of the Nutrition Society*, 62, 611-619.

Rubin, D.B. (1976). Inference and Missing Data. *Biometrika*, 63 (3), 581-592.

Schilling, F. (2009). *PTK - LDT Punkttest und Leistungs-Dominanztest für Kinder (5 - 12 Jahre)*. Dortmund: modernes lernen.

Schwarz, R. (2014). Effekte der Bewegungsförderung. Review längsschnittlicher Evaluationsstudien zu Bewegungsinterventionen in der frühen Kindheit. *Motorik*, 37 (2), 52-63.

Weiß, R. & Osterland, J. (1997). *Grundintelligenztest Skala 1 CFT 1*. Göttingen: Hogrefe.

Wirtz, M. (2004). Über das Problem fehlender Werte: Wie der Einfluss fehlender Informationen auf Analyseergebnisse entdeckt und reduziert werden kann. *Die Rehabilitation*, 43 (2): 109-115. DOI: 10.1055/s-2003-814839

World Health Organization (WHO) (2006). WHO European Ministerial Conference on counteracting obesity. Retrieved from <http://www.euro.who.int/en/what-we-do/health-topics/noncommunicable-diseases/obesity/publications/pre-2009/european-charter-on-counteracting-obesity>.

Wuttke, J. (2008). Erhöhter Dokumentationsbedarf bei Imputation fehlender Werte. *Psychologische Rundschau*, 59, 178-179.