

**Entwicklung der motorischen Fitness
von Thüringer und Brandenburger DrittklässlerInnen
nach der Corona-Pandemie (2022 – 2024)**

Reinhold Kliegl¹, Paula Teich¹, Mareike Voigt¹, Mandy Vogel³

Toni Wöhr², Kathleen Golle¹, Florian Bähr²

¹Universität Potsdam, ²Universität Erfurt, ³Universität Leipzig

dvs2025, Münster, 2025-09-17

EMOTIKON Test: Ergebnisse aus Brandenburg und Thüringen (2022–2024)



Ausdauer ▷
6-Minuten-Lauf



Koordination mit
Zeitdruck ▷ Sternlauf



Schnelligkeit ▷
20-m-Sprint



Schnellkraft Beine ▷
Standweitsprung

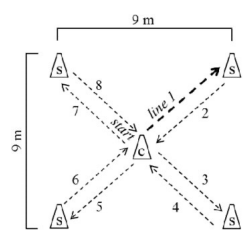


Schnellkraft Arme ▷
Medizinballstoßen



Gleichgewicht ▷
Einbeinstand

△c = center pylone △s = spike pylone



- 1 = forward run
- 2 = backward run
- 3 = sidestep right
- 4 = sidestep left
- 5 = backward run
- 6 = forward run
- 7 = sidestep left
- 8 = sidestep right

Seit 2009 in **Brandenburger Grundschulen (3. Jahrgangsstufe)** flächendeckend jährliche motorische Testung durch Sportlehrkräfte (bisher ≈ 280 000 Kinder). In **Thüringen (BeKiGeKi)** seit 2017 und seit 2022 auch flächendeckend (bisher ≈ 66 000 Kinder)

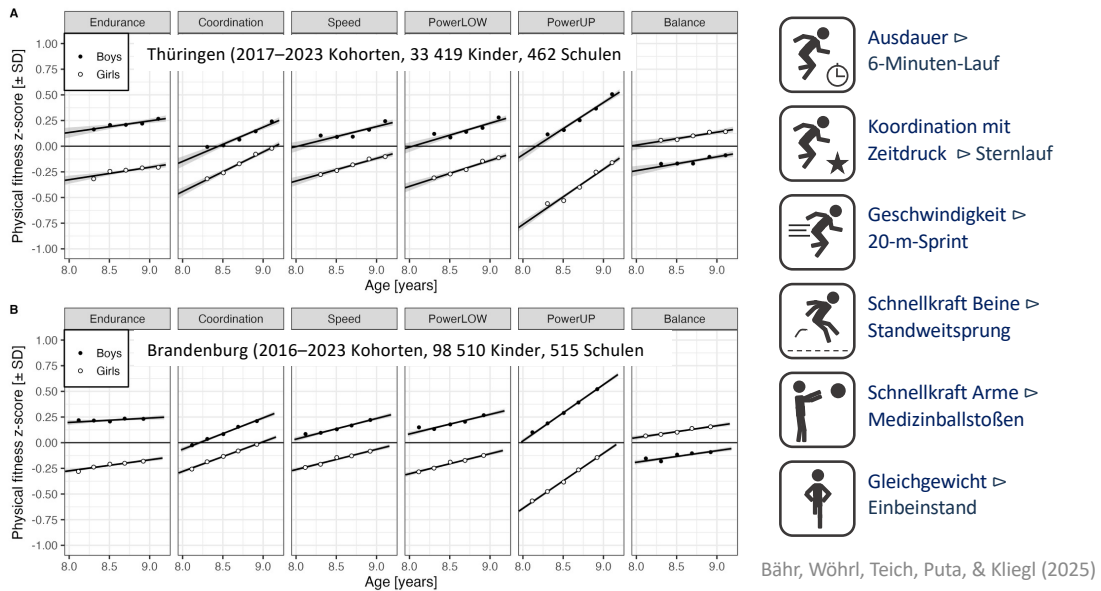
2022 – 2024		Total N	Elternbericht zu Größe und Masse	
Brandenburg	Stichlinge (74.2 %)	46.368	21.977	47.4 %
	Spätlinge	16.100	6.861	43.6%
Thüringen	Stichlinge (82.3 %)	35.197	32.487	92.3 %
	Spätlinge	7.556	6.501	86.0%

<https://www.uni-potsdam.de/de/emotikon/projekt/methodik.html>

2

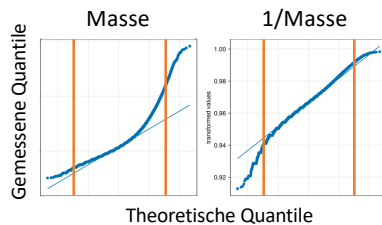
2. In Brandenburg und Thüringen testen die Sportlehrkräfte die Drittklässler in diesen sechs Aufgaben. Im Vortrag beschränke ich mich auf den 6-Minuten-Lauf und das Medizinballstoßen und auf die letzten drei Jahre. Inhaltlich geht es um Unterschiede (1) zwischen Brandenburg und Thüringen, und (2) den Effekt der Anthropometrie auf die motorische Fitness. Hier haben wir zwei große Probleme: Erstens haben wir Größe und Masse nicht gemessen, sondern als freiwilligen Bericht der Eltern. Zweitens haben uns zwar über 90 % der Eltern in Thüringen, aber nur knapp 50 % der Eltern in Brandenburg diese Informationen gegeben. Zunächst als Grundlage ein kurzer Blick auf die Leistung in den sechs Aufgaben nach Geschlecht und Alter im 9. Lebensjahr. (1 Min)

Entwicklungsraten im 9. Lebensjahr und Geschlechtsunterschiede



3. Unsere Metrik für Laufaufgaben ist Geschwindigkeit, nicht die Zeit, d. h. große z-Werte bedeuten hohe Geschwindigkeit und eine bessere Leistung. Das Überraschende ist, dass die Entwicklung für Jungen und Mädchen in jedem Panel statistisch parallel ist. Besonders schön für uns ist auch, dass wir den Brandenburger Nullbefund für diese Interaktion in Thüringen repliziert haben.

Masse $\Rightarrow 30 \text{ [kg]} / \text{Masse [kg]} = \text{„Leichtheitsfaktor“ (Light)}$



Box-Cox-Verteilungsanalyse	λ _Größe	λ _Masse
Brandenburg SEU @ 5.5 Jahre (Schularztuntersuchung)	+1.14	-0.98

- Brandenburger Schuleingangsuntersuchung (SEU; Kooperation mit LAVG Brandenburg): 2009–2024, **342.307** Kinder
- Box-Cox-Analyse konditional auf Geschlecht, Alter und Kohorte
- Normalverteilung für Größe (*height*)
- Reziproke Transformation für Masse normiert am MW der Masse der Zielgruppe: $30 \text{ [kg]} / \text{Masse [kg]} \Rightarrow \text{Leichtkeitsfaktor (light)}$
- Nebeneffekt: $\text{Height [m]} + \text{Light} + \text{Height} \times \text{Light} = \text{Height [m]} + \text{Light} + \text{Height-Light-Product [m]}$
 $\text{HLP} = \text{Height-Light-Produkt [m]} \Leftrightarrow \text{Height-Mass-Ratio [m/kg]} \approx \text{Mass / Height}^2 \text{ [kg/m}^2\text{]} \text{ oder: bye, bye, BMI!}$
- Das **Height-Light-Product** ist die mit Leichtheit gewichtete Größe in Meter

Bähr et al. (2024): <https://www.researchsquare.com/article/rs-3885133/v1>

4

4. Wir haben nicht nur die Laufaufgaben von Zeit zu Speed transformiert, sondern rechnen auch nicht mit Masse, sondern dem Kehrwert von Masse, also der Leichtheit.

Warum? Durch die reziproke Transformation bekommen wir zwei normalverteilte Variablen: Größe und Leichtheit (*light*). Im linearen Modell ist die Interaktion das *Height-Light-Produkt*. Das ist im Normalfall das gleiche wie das Größe-Gewichts-Verhältnis. Normiert man zusätzlich die Leichtheit bei 30 kg, dem Mittelwert der Stichtagskinder, ist das *Height-Light-Produkt* eine mit Leichtheit gewichtete Größe. Die Schätzungen beruhen auf 340K Kindern aus der Brandenburger Schuleingangsuntersuchung, kurz SEU. Das ist die Population, also maximal repräsentativ.

Probleme mit Repräsentativität – Non-Response- und Response-Bias, Stichtagskinder, KiGGS-Referenzdaten

1. Kritische Relevanz von Körperkonstitution, aber im EMOTIKON- und BeKiGeKi-Projekt
keine direkte Messung von Größe und Masse, sondern freiwillige Elternberichte
 - a. **Non-Response-Bias:** Einwilligungsbereitschaft (Rücklaufquoten BB: 47 % - TH: 92 %)
 - b. **Response-Bias:** Über- oder Unterschätzung der Angabe
2. **Stichtagskinder** sind motorisch fitter als zurückgestellte Kinder (Fühner et al., 2022; Teich et al., 2023, 2024) => eingeschränkte Repräsentativität bzgl. Altersstufe (ca. 20-25% sind in 2. Klasse)
3. **KiGGS-Referenzdaten** (2006) für die Berechnung von alters- und geschlechtskorrigierten z-Werten für Anthropometrie (Größe, Gewicht, BMI, etc.) Kohorteneffekte, vor allem eine Unterschätzung der Größe

5

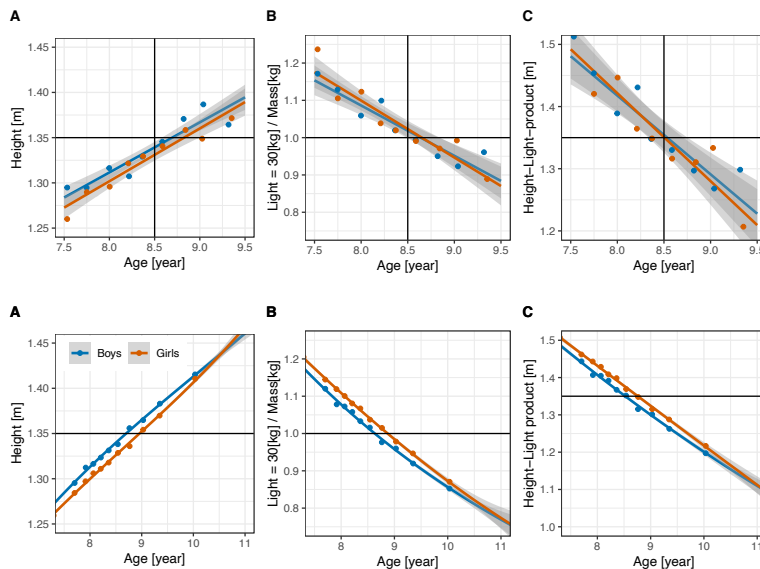
5. Das Symposium hat die Repräsentativität großer Stichproben als Rahmenthema. Wir haben große Stichproben, aber auch vier Probleme mit der Repräsentativität.

(1) und (2) Da wir mit Elternberichten von Größe und Masse arbeiten, müssen wir zwischen dem Non-Response-Bias und dem Response-Bias unterscheiden. Der Non-Response-Bias bezieht sich darauf, ob die Eltern überhaupt antworten oder eben nicht. Und falls es eine Antwort gibt, dann ist der Response-Bias die Ungenauigkeit der Antwort. Der Unterschied in der Rücklaufquote von 47 % in Brandenburg und 92 % in Thüringen öffnet uns hier Möglichkeiten für eine Dissoziation.

(3) Stichtagskinder sind nicht repräsentativ für ihre Altersstufe, sie sind motorisch fitter als zurückgestellte Kinder, wenn man den chronologischen Altersunterschied berücksichtigt. Da wir nur Daten aus der 3. Jahrgangsstufe haben, fehlen uns ca. 20% - 25% der gleichaltrigen Kinder in der 2. Jahrgangsstufe, die sich nicht zufällig, sondern systematisch von denen in der 3. Klasse unterscheiden.

(4) Und das letzte Problem ist, dass die KiGGS-Normen, mit denen wir bisher für Alters- und Geschlechtseffekte kontrolliert haben, für die aktuellen Kohorten keine gute Baseline mehr sind. Die Kinder sind jetzt größer als vor 20 Jahren.

Zwei neue Referenzdaten (AUT FIT-Projekt: N=666; OOE-Project: N=37.301)



AUT FIT: [Jarnig et al. \(2022\)](#)

- Zufällige Auswahl von 12 von 39 Volksschulen aus den Land-/ Stadtbezirken von Klagenfurt,
- Testung: 2019
- **Messungen: u.a. Größe, Masse, 6-Minuten-Lauf und Medizinballstoßen**

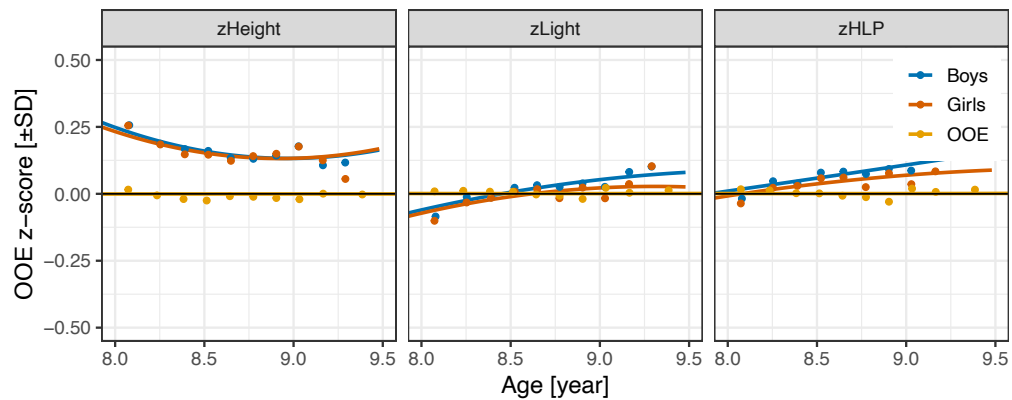
OOE: [Drenowatz et al. \(2021\)](#)

- Freiwillige Teilnahme von Volksschulen aus allen Bezirken Oberösterreichs
- Testung: 2016–18; 2022–24
- **Messungen: u.a. Größe, Masse, 6-Minuten-Lauf und Medizinballstoßen**

6

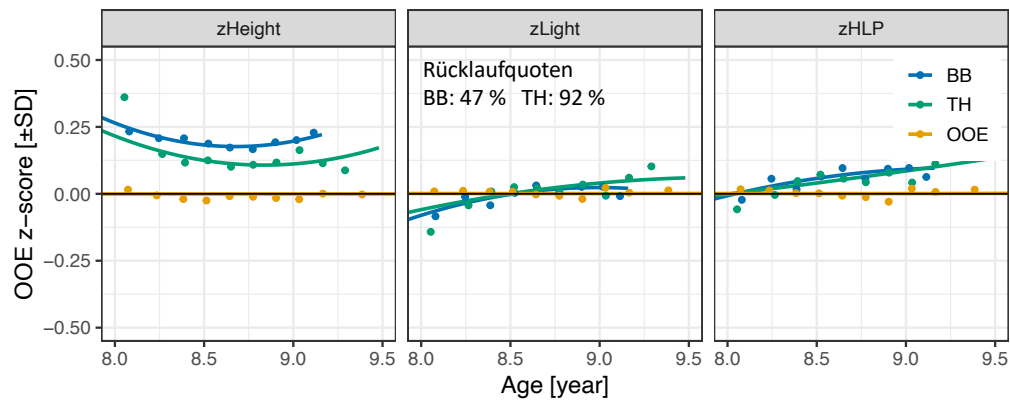
6. Zwei Kollegen aus Österreich, Gerald Jarnig aus Graz und Clemens Drenowatz aus Linz, haben spontan ihre Daten zur Verfügung gestellt, mit denen wir eigene Normierungen implementieren konnten. Die Abbildungen zeigen, dass die Beziehung zwischen *Height*, *Light* und dem *Height-Light-Product* für 7-10 jährige Jungen und Mädchen linear ist. Der Geschlechtseffekt zeigt sich im großen Datensatz besser. Allerdings ist hier aber die Repräsentativität der Daten im Vergleich zum AUT FIT-Projekt eingeschränkt. Das ist auch ein interessanter Tradeoff für die Diskussion. (2025-09-22: Korrektur der y-Achsenbeschriftung zu Übereinstimmung mit Folie 4.)

Alter x Geschlecht relativ zu OOE-Referenzdaten (N=37.301)



7. Wir haben also für die beiden Datensätze alters- und geschlechtsbereinigte z-Werte berechnet und dann die Parameter für die Korrektur von Alters- und Geschlechtseffekten in unseren Daten verwendet. Ich zeige Ihnen die Ergebnisse für die Normierung mit den Kindern aus Oberösterreich. Die OOE-Kinder liegen genau auf der Nulllinie für *Height*, *Light* und *HLP*. Das muss so sein, weil wir die Effekte ja für die Normierung rausgerechnet haben. Die Brandenburger und Thüringer Jungen und Mädchen liegen aber für *Height* deutlich über der Nulllinie. Die Größenangaben der Eltern sind also vermutlich etwas zu optimistisch. Beim *Light* sehen wir leichte Unterschätzung für junge und Überschätzung für ältere Kinder. Das hatte ich nicht erwartet, ist aber plausibel, meine ich

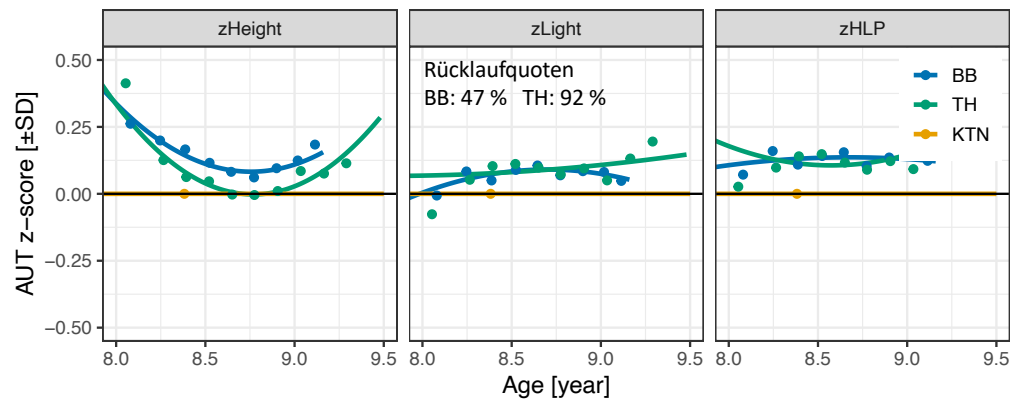
Alter x Bundesland relativ zu OOE-Referenzdaten (N=37.301)



8

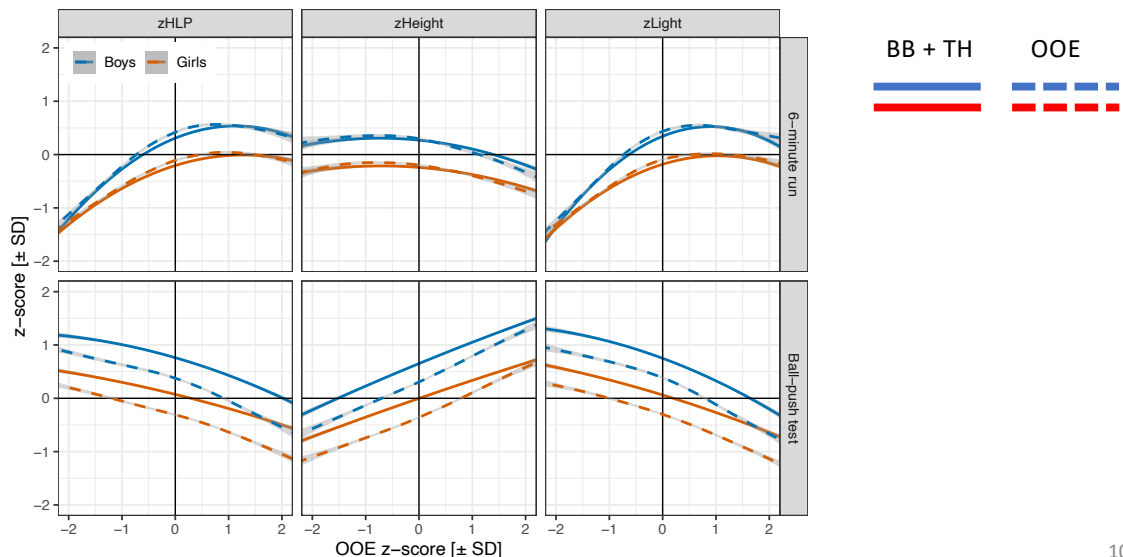
8. Können wir mit unseren Daten Response-Bias und Non-Response-Bias unterscheiden?
 Brandenburg liegt für *Height* deutlich über Thüringen. Das ist die Verzerrung durch den Non-Response-Bias. Die Differenz zwischen Thüringen und der Nulllinie ist unsere Schätzung für den Response Bias. Für das *Height-Light-Product* sehen wir nur einen Response-Bias, der mit dem Alter der Kinder größer wird. Wir bekommen ähnliche Ergebnisse, wenn wir die Daten aus Kärnten als Referenzdaten nehmen. Das Besondere an diesen beiden Datensätzen ist, dass auch der 6-Minuten-Lauf und Medizinballstoßen gemessen wurden. Darum können wir die Fitness in diesen beiden Aufgaben mit Bezug auf direkt gemessene und berichtete anthropometrische Indikatoren vergleichen.

Alter x Bundesland relativ zu AUT-Referenzdaten (N=666)



9.. Ein ähnliches Ergebnis für Kärnten.

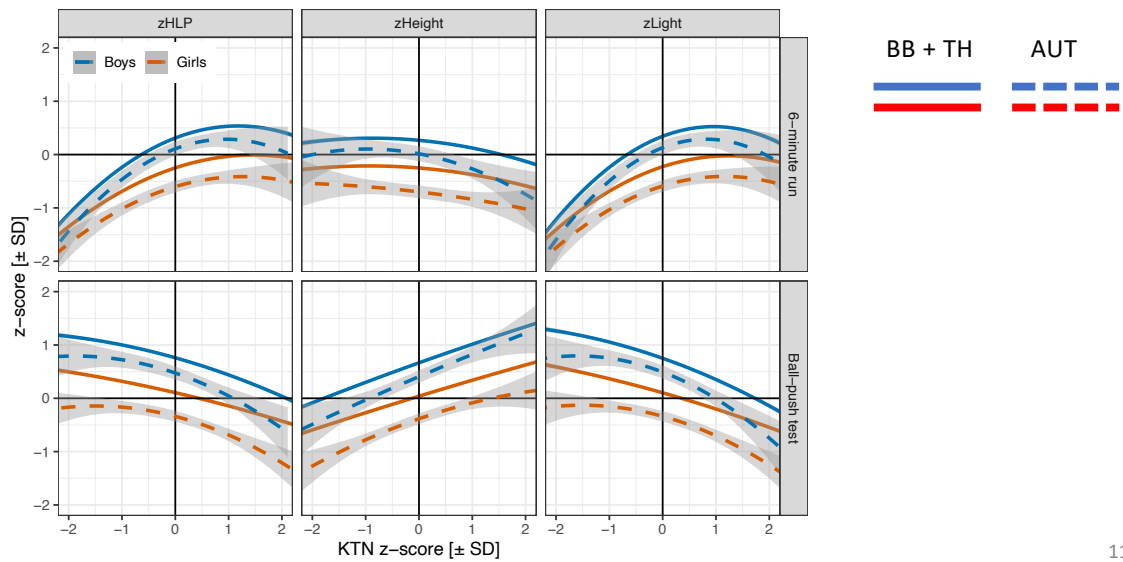
Interaktionen von Geschlecht und berichtete / gemessene Anthropometrie



10

10. Wir brauchen die Anthropometrie ja vor allem, weil sie grundlegend für die Fitness der Kinder ist. Gibt es Unterschiede in den Fitnessaufgaben bzgl. berichteter und direkter Messung der Anthropometrie? Die Ergebnisse für den Sechs-Minuten-Lauf sind in der ersten und die für den Medizinball in der zweiten Reihe. Die gestrichelten Linien sind die Ergebnisse aus Oberösterreich. Ein großes *Height-Light-Product* und *Light* sind vorteilhaft für die Laufristanz -- aber nur bis zu einem bestimmten Wert. Für das Medizinballstoßen sehen wir ein qualitativ anderes Leistungsprofil, nämlich das umgekehrte Muster für *Height-Light-Product* und *Light*. Außerdem hat *Height* einen stärker positiven Effekt. Hier ist also von Vorteil eher schwer und groß zu sein. Für das Laufen haben wir keine Evidenz für Unterschiede zwischen gemessener und berichteter Anthropometrie. Das sieht auf den ersten Blick schön aus, aber eigentlich hatten wir auf Grund der Überschätzung von Größe, einen Vorteil für Brandenburger und Thüringer Kinder erwartet. Beim Medizinballstoßen zeigt sich ein Bias in der erwarteten Richtung. Gründe für Differenzen gibt es viele, z.B. Details in der Implementation der Laufaufgabe, aber es kann ja auch Unterschiede in Fitness der Kinder dieser Stichproben geben, evt. auch aufgabenspezifisch.

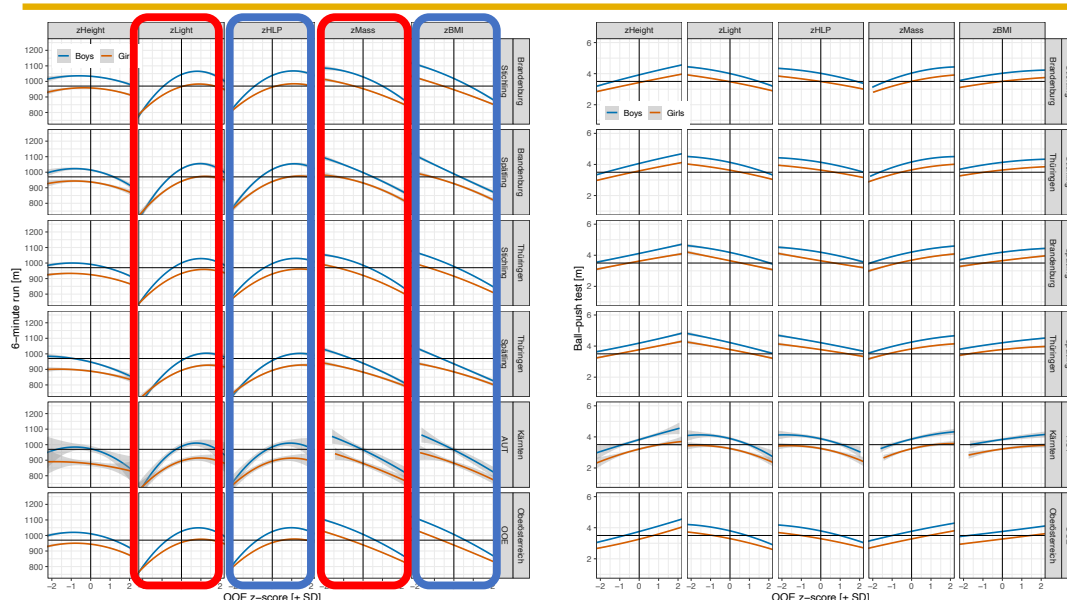
Interaktionen von Geschlecht und berichtete / gemessene Anthropometrie



11

11. Und für den anderen, repräsentativeren, aber kleineren Datensatz aus Kärnten sehen wir das erwartete Muster für beide Aufgaben.

Überblick über alle Ergebnisse des 6-Minuten-Laufs und Medizinballstoßens



12

11. Hier nochmal ein Überblick über alle Anthropometrie x Geschlecht x Aufgaben x Datensatz Befunde. Ich nehme an, Sie stimmen mir zu, dass die Befunde in jeder Spalte, also für die sechs Datensätze *qualitativ* übereinstimmen. Das gilt auch für den eher kleinen Datensatz aus Kärnten. Die vorher gezeigten signifikanten *quantitativen* Verzerrungen für Elternberichte von Größe und Masse sind theoretisch wichtig und spannend, aber vermutlich nicht von großer praktischer Relevanz. Ein zweiter Befund betrifft den 6-Minuten-Lauf. Für *Light* und *Height-Light-Product* ist größere Krümmung für Jungen als Mädchen besser zu sehen als für Masse und BMI. Das ist die Macht der Transformation!

Synopse

- Sehr große **qualitative** Übereinstimmung der Befunde zur motorischen Fitness im 6-Min-Lauf und Medizinballstoßen für Elternberichte und direkte Messungen von Größe und Masse
- Referenzdaten aus Oberösterreich und Kärnten als Alternative zu KiGGS-Referenzdaten; Tradeoff von Repräsentativität und Größe der Stichprobe
- Potenzielle Konfundierung mit Fitness-Differenzen zwischen OOE- und BB/TH-Kinder; ideal wäre Messung und Elternbericht von Anthropometrie und Fitness in einer Studie
- Viele Befunde im Detail, zum Beispiel:
 1. Dissoziation von Non-Response- und Response-Bias für Elternbericht von Masse und Größe
 2. Elternberichte überschätzen vor allem die Größe ihres Kindes; Masse tendenziell für jüngere Kinder überschätzt und für ältere Kinder überschätzt
 3. TH-BB-Unterschied (noch) nicht vollständig durch Non-Response-Bias erklärt; sozialstrukturelle und regionale Effekte?

Mehr zu Non-Response-Bias und sozialstrukturellen regionalen Effekten in Mareike Voigts Vortrag






Vielen Dank!








EMOTIKON-Podcast

Besonders großes Dankeschön: Nina Ulatowski und Daria Kuhn (LAVG im MSGIV Brandenburg)
Gerald Jarnig (Universität Graz) – Clemens Drenowatz (PH OÖ Linz)

- Brandenburg – EMOTIKON:** <https://www.uni-potsdam.de/en/emotikon/>
- Thüringen – BeKi-GeKi:** <https://bekigeki.github.io>
- Daten, R-Scripts und Julia-Scripts sind in OSF-Repositorien verfügbar (Links in Artikeln)**

Teich, P., Arntz, F., Wöhlrl, T., Bähr, F., Golle, K., & Kliegl, R. (2025). Association of school social status with COVID-19 pandemic-related changes and post-pandemic rebounds of children's physical fitness. *Sports Medicine - Open* 11:41. doi.org/10.1186/s40798-025-00838-5

Bähr, F., Wöhlrl, T., Teich, P., Puta, C., & Kliegl, R. (2025). Impact of age, sex, body constitution, and the COVID-19 pandemic on the physical fitness of 38,084 German primary school children. *Scientific Reports*, 15, 11300. doi.org/10.1038/s41598-025-95461-5

Teich, P., Golle, K., & Kliegl, R. (2024). Association between time of assessment within a school year and physical fitness of primary school children. *Scientific Reports* 14, 11500. doi.org/10.1038/s41598-024-61038-x

Bähr, F., Wöhlrl, T., Teich, P., Puta, C., & Kliegl, R. (2024). Impact of height-to-mass ratio on physical fitness of German third-grade children. *Research Square* <https://www.researchsquare.com/article/rs-3885133/v1>

Teich, P., Fühner, T., Bähr, F., Puta, C., Granacher, U., & Kliegl, R. (2023). Covid pandemic effects on the physical fitness of primary school children: Results of the German EMOTIKON project. *Sports Medicine - Open* 9:1. doi.org/10.1186/s40798-023-00624-1

Teich, P., Fühner, T., Granacher, U., & Kliegl, R. (2023). Physical fitness of primary school children differs depending on their timing of school enrollment. *Scientific Reports*, 13, 8788. doi.org/10.1038/s41598-023-35727-y

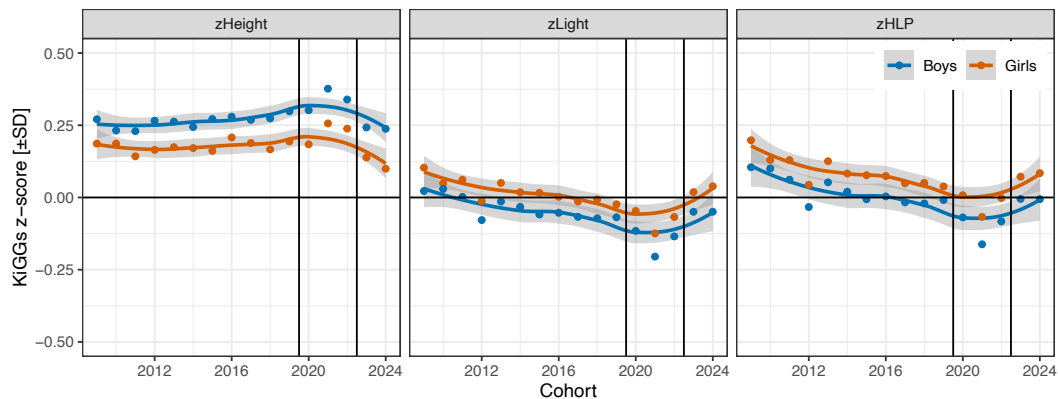
Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2022). Effect of timing of school enrollment on physical fitness in third graders. *Scientific Reports*, 12, 7801. doi.org/10.1038/s41598-022-11710-x

Fühner, T., Granacher, U., Golle, K., & Kliegl, R. (2021). Age and sex effects in physical fitness components of 108,295 third graders including 515 primary schools and 9 cohorts. *Scientific Reports*, 11, 17566. doi.org/10.1038/s41598-021-97000-4

13. Besuchen Sie uns auf den Webseiten. Wir tun unser Bestes, dass die Projekte weiterlaufen. Wir hoffen, dass unsere Partner dabeibleiben und neue dazukommen. Hallo, Österreich! Vielen Dank!

KiGGS-Referenzen für alters- und geschlechtskorrigierte Anthropometrie der Kohorten aus der Brandenburger Schuleingangsuntersuchung (2007-2024)

R Package childsd: <https://git.sc.uni-leipzig.de/my221hepi/childsd> (Vogel, 2024)

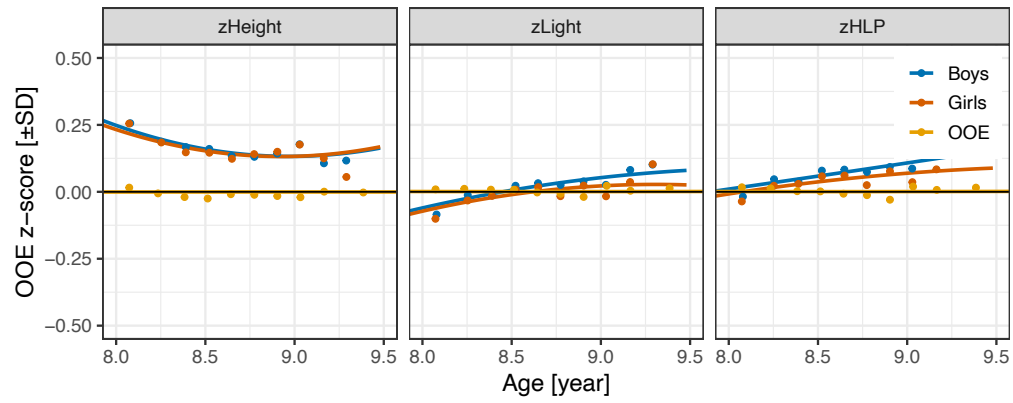


1. **Brandenburger Schuleingangsuntersuchung 2007–2024 (N ≈ 340.000 Kinder; Kooperation mit dem LAVG)**
2. Größe immer deutlich über erwarteten KiGGS-Referenzwerten von Null für Jungen und Mädchen
3. Abweichung bei Größe größer für Jungen als Mädchen
4. Dazu: Covid-Effekt in 2021 und 2022 SEU-Kohorten, d.h. 2023 und 2024 in 3. Jahrgangsstufe

15

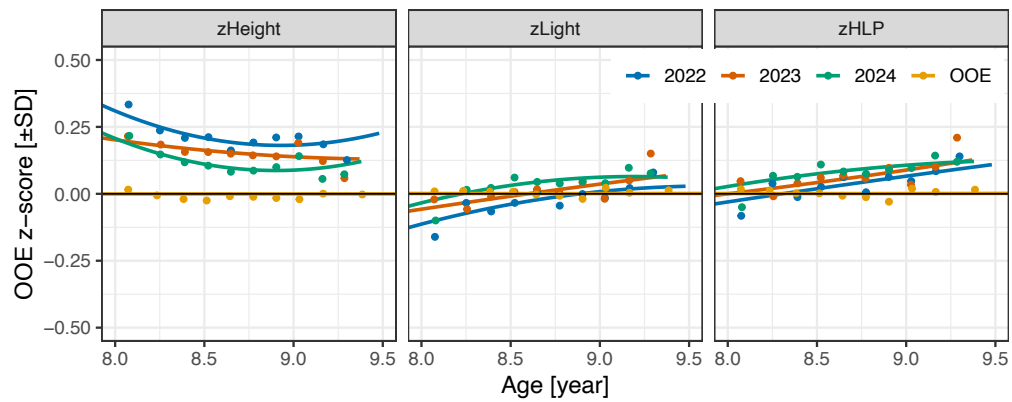
6. In dieser Abbildung habe ich für die Kohorten der SEU-Daten die alters- und geschlechtskorrigierten z-Werte mit Bezug auf die KiGGS-Referenzdaten aus Mandy Vogel's childsd-package gezogen. Mandy Vogel hat uns auch die Normen für Light und HLP berechnet. Man gibt Alter und Geschlecht ein und bekommt für die anthropometrischen Indizes z-Werte, die angeben, wie viele Standardabweichungen das Kind vom Mittelwert Null der Referenzpopulation entfernt ist. Die SEU-Daten sind maximal repräsentativ und haben keinen Bias, liegen aber vor allem für Größe deutlich über der erwarteten Nulllinie – und die Jungen noch mehr als die Mädchen. D.h. wir haben keine klare Baseline für die Elternberichte von Größe und Masse. Wir könnten allerdings die z-Scores der SEU-Daten in der Abbildung als Baseline für die berichteten Anthropometrie in EMOTIKON und BekiGeKi nehmen.

Alter x Sex Bias in BB und TH relativ zu OOE-Referenzdaten (N=37.301)



Keine Evidenz für Geschlechtsunterschiede bei Größe, aber Masse unterschätzt für ältere Jungen

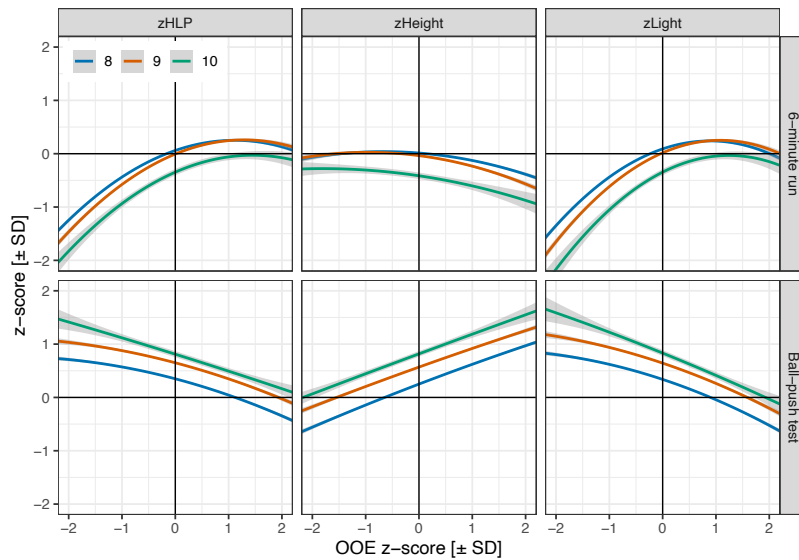
Kohorten-Bias relativ zu OOE-Referenzdaten (N=37.301)



17

Die Überschätzung der Größe reduziert sich über die letzten drei Jahre; für junge Kinder reduziert die Unterschätzung der Leichtigkeit (also Überschätzung von Masse)

Effekte des Alters für Stichlinge (8 + 9) und Einschulungsstatus (10)



1. 8 + 9 sind junge und ältere Stichlinge; 10 sind Spätlinge
2. Evidenz für niedrigere Werte in der Ausdauer und bessere Schnellkraft Arme für Stichlinge