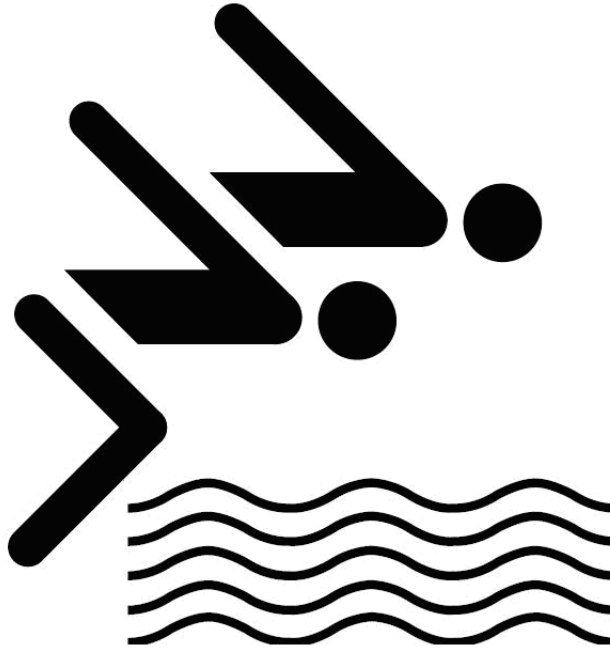


Analyses of Elite Swimming Performances and Their Respective Between-Gender Differences over Time



Seminararbeit von Jasmin Abedieh

Betreut von Manuel Eugster

Im Rahmen des Bachelor Seminars

Statistik im Sport

bei Prof. Dr. Friedrich Leisch, Manuel Eugster, Sebastian Kaiser

am Institut für Statistik

Ludwig-Maximilians-Universität München

Inhaltsverzeichnis

1	Einleitung	1
2	Sportart - Schwimmen	1
3	Analyse von Eliteschwimmern	3
3.1	Problemstellung	3
3.2	1. Studie	3
3.2.1	Methode	4
3.2.2	Ergebnisse	7
3.3	2. Studie	12
3.3.1	Methode	12
3.3.2	Ergebnisse	13
4	Zusammenfassung	14
5	Diskussion	15
6	Ausblick	16
	Anhang A	18
	Anhang B	19
	Literaturverzeichnis	22

1 Einleitung

In vielen Bereichen unserer Gesellschaft wird die Statistik angewendet. Sei es nun zum Beispiel in der Wirtschaft, Medizin, Soziologie oder auch in der Psychologie. In all diesen Bereichen werden Daten erhoben, die mit Hilfe statistischer Methoden ausgewertet werden. Dabei kann die Untersuchung eines speziellen Merkmals, oder aber des Zusammenhangs zweier Variablen im Vordergrund stehen. Des Weiteren können mittels Test- und Schätzverfahren Hypothesen formuliert und auf Signifikanz geprüft werden.

Nun gibt es noch Bereiche in unserer Gesellschaft, in denen statistische Analysen noch am Anfang ihrer Entwicklung stehen. Zu solch einem Gebiet zählt der Sport. Eigentlich wird in fast jeder Sportart eine große Datenmenge erhoben, tatsächlich aber werden statistische Verfahren bisher größtenteils im amerikanischen Raum eingesetzt. Auch im Sport können anhand von Daten beispielsweise Gesetzmäßigkeiten oder sogar Prognosen ermittelt werden. So wird in der folgenden Arbeit die Sportart Schwimmen unter einer statistischen Betrachterperspektive näher untersucht. Als Grundlage dieser Arbeit dient aus dem Journal of Quantitative Analysis in Sports der Artikel *Analyses of Elite Swimming Performances and Their Respective Between-Gender Differences over Time* von Johnson et al. (2009).

2 Sportart - Schwimmen

Schwimmen zählt nicht nur zu einer der beliebtesten Freizeitsportarten, sondern wird auch als eine sportliche Disziplin betrachtet, deren Grundform darin besteht, dass im Wettkampf mehrere Teilnehmer mit dem Ziel gegeneinander antreten, vorgegebene Strecken in möglichst kurzer Zeit zu bewältigen.

Der Schwimmsport lässt sich in 4 verschiedene Grundschwimmstile unterteilen. Brustschwimmen, Kraulschwimmen, Rückenschwimmen und Schmetterlingsschwimmen (Delfinschwimmen). Die Schwimmausbildung beginnt normalerweise mit dem Erlernen des Brustschwimmens. Die Aneignung weiterer Schwimmtechniken folgt meist danach. Der Bruststil zeichnet sich durch einen Wechsel von Armzug und Grätschbeinschlag aus. Der Antrieb wird zusätzlich von der sogenannten Undulationstechnik, die einer Wellenbewegung ähnelt, gefördert.

Der Kraulstil stellt die schnellste Schwimmtechnik dar, da der Antrieb durch die abwechselnde Armzugbewegung und der wechselseitigen Auf- und Abwärtsbewegung des Beinschlags ohne Unterbrechung durchgeführt wird. Der Rückenstil unterscheidet sich kaum vom Kraulstil, außer der Rückenlage im Wasser und dem Abwinkeln der Arme unter Wasser, was einen höheren Krafteinsatz erlaubt, als das bei gerader Führung des Armes im Wasser der Fall wäre.

Der Schmetterlingsstil gilt als zweitschnellste Schwimmart. Diese Technik setzt sich aus einem gleichzeitigen nach vorne Bringen der Arme und eines geschlossenen Beinschlages zusammen. Hierbei entsteht eine Ganzkörper-Wellenbewegung die den Antrieb zusätzlich unterstützt.

Schwimmwettkämpfe haben als Austragungsort im Normalfall Schwimmbecken in Hallen oder seltener, im Freien. Das sogenannte Langstreckenschwimmen bildet eine Ausnahme, denn hier werden die Wettkämpfe in offenen Gewässern ausgetragen.

Zu den verschiedenen Wettkampfbestimmungen, die erfüllt sein müssen, gehören: eine Bahnlänge von 25 oder 50 m, eine Bahnbreite von 2.50 m, eine Wassertiefe von mindestens 1.35 und einer Wassertemperatur von 25 bis 28 °C, sowie Bahnleinen die jeden Schwimmer von seinen Mitstreitern abtrennen und zudem als Wellenbrecher fungieren. Darüber hinaus benötigt man fixierte Startblöcke, die aus einer rutschfesten Oberfläche bestehen.

Zu den zu bewältigenden Strecken gehören: die 50-m-Distanz als Sprintstrecke, 100 m, 200 m, 400 m, 800 m und 1500 m sowie 5000 m und 25000 m als Langstrecke.

Des Weiteren ist zu erwähnen, dass im Schwimmsport Einzelwettkämpfe und Staffelwettkämpfe stattfinden. Allerdings handelt es sich bei den Staffelwettkämpfen ausschließlich um 4er-Staffeln, hierbei müssen die Schwimmer folgende drei Distanzen bewältigen: 4×50 m, 4×100 m und 4×200 m.

In Einzelwettbewerben werden die Disziplinen Brust-, Rücken-, Kraul-, Schmetterling-, sowie Freistil geschwommen. In der Disziplin Freistil kann der Schwimmer einen beliebigen Schwimmstil auswählen, wobei ein Wechsel des Schwimmstils während des Wettkampfes durchaus erlaubt ist. Nach den FINA-Wettkampfbestimmungen (Fédération Internationale de Natation) ist beim Freistil wichtig, dass durchgehend, außer nach dem Start und den Wenden, ein Körperteil des Schwimmers die Wasseroberfläche durchbricht. Die Wettkampfbestimmungen sind auf FINA (2009) veröffentlicht.

Bei Staffelwettkämpfen handelt es sich entweder um eine Freistilstaffel oder um eine Lagenstaffel. Hier treten erfahrungsgemäß vier Schwimmer an, wobei bei der Lagenstaffel eine vorgegebene Reihenfolge der Schwimmstile im Gegensatz zur Freistilstaffel zu absolvieren ist. Die Reihenfolge ist wie folgt festgelegt: Rücken-, Brust, Schmetterling und Freistil. Beim herkömmlichen Lageneinzelwettkampf ist jedoch die Reihenfolge der vier Schwimmtechniken eine andere, nämlich Schmetterling-, Rücken-, Brust- und schließlich Freistil.

Zum Ablauf des Wettkampfes sind folgende Regeln zu beachten: Der Startsprung erfolgt nach einem Startsignal vom Startblock, der sich jeweils am Beckenkopf der jeweiligen Schwimmbahn befindet. Beim Rückenschwimmen bzw. Lagenstaffel handelt es sich um eine Ausnahme, denn hier erfolgt der Start direkt im Wasser.

Der Start wird wiederholt, wenn ein oder mehrere Schwimmer einen Fehlstart machen. Einem Schwimmer wird in der Regel bloß ein Fehlstart erlaubt. Bei einem zweiten Fehlstart erfolgt eine Disqualifikation des Schwimmers, während der Wettkampf für die anderen Schwimmer beim zweiten Start auf jeden Fall gewertet wird. Bei Staffelwettkämpfen ist es dem Folgeschwimmer erlaubt bis zu 0.03 Sekunden vor dem Anschlag des schwimmenden Teamkollegen zu starten, man spricht hier von einem sogenannten fliegendem Wechsel.

Bei einem Wettkampf muss nach jeder geschwommenen Bahn eine regelgerechte Wende vollzogen werden. Die Gesamtzeit wird in dem Moment gestoppt, in dem der Schwimmer die Wand erreicht, wobei beim Brust- und Schmetterlingsstil zu beachten ist, dass der Schwimmer mit beiden Händen die Wand berühren muss.

Erwähnenswert ist noch, dass es im Schwimmsport unterschiedliche Wettkampfrichter gibt: Zeitnehmer, Zielrichter, Wenderichter, Schwimmrichter und Starter, die zusätzlich von einem Auswerter, einem Protokollführer und einem Schiedsrichter unterstützt werden (vgl. DSV, 2009).

Eine detaillierte Beschreibung des Schwimmsports bietet Wikipedia (2009).

3 Analyse von Eliteschwimmern

3.1 Problemstellung

Im Sport ist die Untersuchung von einem Leistungsunterschied zwischen Frauen und Männern von sehr großem Interesse. Schwimmen eignet sich sehr gut für solch einen Vergleich und aus diesem Grund beschäftigt sich die folgende Arbeit mit der Entwicklung der Schwimgeschwindigkeit und dem Vergleich zwischen weiblichen und männlichen Eliteschwimmern über eine längere Zeitspanne. Für die Untersuchung bietet sich die Sportart Schwimmen aus mehreren Gründen hervorragend an: Eliteschwimmer, ob männlich oder weiblich, haben größtenteils das gleiche Trainingsmuster. Dazu kommt, dass beim Schwimmen die Zeitmessung kaum durch Störvariablen beeinflusst wird. Es handelt sich somit um eine bereinigte Messung, in der Zusatzfaktoren wie Wetter, Strategie oder Teameffekte kaum eine Rolle spielen. Hinzukommt, dass Schwimmen zu den Sportarten gehört die am längsten mitdokumentierte Daten aufweisen, was die Analyse durch ihr Datenaufkommen repräsentativer macht.

Zahlreiche Studien haben sich mit diesem Thema bereits auseinandergesetzt. Aus einigen geht hervor, dass der geschlechtliche Leistungsunterschied immer mehr abgenommen hat und zum Teil fast verschwunden ist (vgl. Tatem et al., 2004 und Whipp and Ward, 1992). Andere wiederum kommen zu dem Ergebnis, dass die Weltrekorde der Männer auch in Zukunft immer schneller bleiben werden als die der Frauen (beispielsweise Chevront et al., 2005, Nevill and Whyte, 2005, Seiler et al., 2007 und Sharp, 2004). Zudem hat man bei vorherigen Studien festgestellt, dass Eliteschwimmer im Laufe der Zeit immer bessere Rekorde erzielt haben. Diese Arbeit soll nun untersuchen, inwiefern diese Hypothesen mit unserer Datenauswertung übereinstimmen

3.2 1. Studie

Die Daten der ersten Studie wurden von der Website USA Swimming (USAS, 2009) herangezogen. Die Analyse basiert auf den jährlichen Aufzeichnungen der männlichen und weiblichen

Rekorde von 1962 bis 2007 von den folgenden Schwimmdisziplinen: 100 m in Frei-, Rücken-, Brust- und Schmetterlingsstil, sowie 200 m Freistil. Außerdem gibt es noch Daten von 50 m Freistil ab 1980 bis 2007, denn 1980 wurden erstmals die Rekorde dieser Disziplin mitdokumentiert.

Von einer zahlreichen Vielfalt an Schwimmdisziplinen wurden diese sechs ausgewählt, um einerseits einen Vergleich der Geschwindigkeit eines Schwimmstils für drei unterschiedliche Bahnlängen und andererseits einen Vergleich der Geschwindigkeit bei fester Bahnlänge für mehrere Schwimmstile zu ermöglichen.

Zu erwähnen ist, dass im Falle mehrerer neuer Rekorde innerhalb eines Jahres die Bestzeit genommen wurde und falls in einem Jahr kein neuer Rekorde aufgestellt wurde, so wurde für dieses Jahr der bestehende Rekord verwendet. Man betrachtet also nun den Trend der Rekorde. Für den Datensatz der im Folgenden untersucht wird, werden nicht die absoluten Zeiten verwendet, die ein Schwimmer für die Strecke benötigt, sondern die zugehörige Durchschnittsgeschwindigkeit. Eine schnellere Streckenzeit resultiert also in einer höheren Durchschnittsgeschwindigkeit.

Im Folgenden soll nun die Entwicklung der männlichen und weiblichen Durchschnittsgeschwindigkeit, zunächst einzeln und anschließend im Vergleich zu einander, über den Zeitraum von 1962 bis 2007 analysiert werden.

Hierfür wird nun die abhängige Variable Y „Geschwindigkeit“ auf die unabhängige Variable X „Jahreszahl“ regressiert. Um festzustellen, ob geschlechtsbedingte Unterschiede bezüglich der Durchschnittsgeschwindigkeit existieren, wird eine weitere Regressionsanalyse durchgeführt. Dabei wird als neue abhängige Variable nun die Differenz aus den männlichen und weiblichen Durchschnittsgeschwindigkeiten gewählt.

3.2.1 Methode

Die Regressionsanalyse ist ein statistisches Analyseverfahren mit dem Ziel, Beziehungen zwischen einer abhängigen Variable Y und einer oder mehreren unabhängigen Variablen $X_k, k = 1, \dots, p$, zu modellieren. Zur Beschreibung dieses Zusammenhangs zwischen Y und den X_k können unterschiedliche Funktionen verwendet werden, die sich in ihrer Komplexität unterscheiden. Im einfachsten Fall verwendet man eine lineare Funktion, die sogenannte Regressionsgerade. In diesem Fall nimmt man an, dass das interessierende Merkmal Y gut durch eine Linearkombination aus den sogenannten Prädiktoren X_k erklärt werden kann (lineare Regression). Das Basismodell besitzt also folgende Form (vgl. Fahrmeir et al., 2007 sowie Küchenhoff,

2009):

$$\begin{aligned}y_i &= \beta_0 + \beta_1 x_{1i} \dots + \beta_p x_{pi} + \varepsilon_i; & i = 1, \dots, n \\E(\varepsilon_i) &= 0 \\V(\varepsilon_i) &= \sigma^2 \\ \{\varepsilon_i \mid i = 1, \dots, n\} & \text{stoch. unabhängig} \\ \varepsilon_i & \sim N(0, \sigma^2)\end{aligned}$$

(BM)

- y_i : Zielgröße (Zufallsgröße), abhängige Variable
- x_{1i}, \dots, x_{pi} : **fest**e bekannte Einflussgrößen, unabhängige Variablen
- ε_i : Zufallsfehler mit unbekannter Varianz σ^2
- $\beta_0, \beta_1, \dots, \beta_p$: Regressionsparameter
- n : Anzahl der Beobachtungen

Die Regressionsparameter β_k können relativ einfach geschätzt werden, zum Beispiel mittels Kleinst-Quadrat-Methode (KQ-Methode). Die Güte der Anpassung die mit dem Regressionsmodell erreicht wird, wird durch die Größe R^2 bestimmt, das sogenannte *Bestimmtheitsmaß*.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad R^2 \in [0; 1]$$

mit den Größen $\hat{y}_i := \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_p x_{pi}$ und dem arithmetischen Mittel $\bar{y} := \frac{1}{n} \sum_{i=1}^n y_i$. Dabei bezeichnen die $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ die Schätzungen der Regressionsparameter, die mittels KQ-Methode berechnet werden können. Dieses Maß beschreibt den Anteil der Varianz, die durch das Regressionsmodell erklärt werden kann, d.h. je größer R^2 umso besser passt das gerechnete Modell für die jeweilige Datensituation.

Ob das Modell überhaupt sinnvoll ist bzw. sich eignet, um den Sachverhalt zu beschreiben, kann mit dem *Overall-F-Test* geprüft werden (vgl. Küchenhoff, 2009). Der Test geht von folgendem Hypothesenpaar aus:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad \text{vs.} \quad H_1 : \beta_k \neq 0 \text{ für mindestens ein } k, k = 1, \dots, p \quad (\text{H1})$$

Als Teststatistik verwendet man die Größe F , die einer F -Verteilung mit p und $n - p - 1$ Freiheitsgraden genügt und sich wie folgt berechnet:

$$F = \frac{n - p - 1}{p} \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \sim F_{p, n-p-1}$$

Dabei wird H_0 verworfen, falls gilt:

$$F > F_{1-\alpha, p, n-p-1}$$

wobei $F_{1-\alpha, p, n-p-1}$ das $(1 - \alpha)$ -Quantil der F -Verteilung mit p und $n - p - 1$ Freiheitsgraden bezeichnet.

In unserem Fall wird zunächst eine lineare Einfachregression zwischen Y „Geschwindigkeit“ und der einzigen Einflussgröße X „Jahreszahl“, d.h.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i; \quad i = 1, \dots, n \quad (1)$$

durchgeführt. Dabei kann β_1 als die Rate interpretiert werden, mit der sich die amerikanischen Rekordgeschwindigkeiten konstant erhöhen. Geht man aber davon aus, dass der Anstieg der Rekordgeschwindigkeiten nicht konstant bleibt, sondern sich im Laufe der Zeit verlangsamt, kann durch Verwendung eines Polynoms zweiten Grades als Regressionsfunktion eine bessere Modellanpassung erreicht werden. Die Regressionsgleichung hat hier folgende Form:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i; \quad i = 1, \dots, n \quad (2)$$

Verwendet man ein Polynom dritten Grades als Regressionsfunktion, d.h.

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \varepsilon_i; \quad i = 1, \dots, n \quad (3)$$

so beschreibt das Modell zunächst eine abnehmende Geschwindigkeitsverbesserung, die anschließend wieder zunimmt.

Um aus den Modellen (1) - (3) das am besten passende zu finden, wurde wie folgt vorgegangen. Zunächst wurde für alle Disziplinen und Geschlechter eine standardisierte lineare Regression (1) durchgeführt.

Anschließend wurde das Modell (1) mit dem nächst komplizierten Modell (2) verglichen, indem die Differenz $\Delta R^2 = R_2^2 - R_1^2$ betrachtet wurde, wobei R_i^2 das Bestimmtheitsmaß von Modell (i) bezeichnet. Modell (2) wurde dem Modell (1) vorgezogen, wenn sich das Bestimmtheitsmaß durch das neue Modell (2) um mindestens 0.02 verbessert hat, wenn also gilt $\Delta R^2 \geq 0.02$. Falls Modell (2) eine bessere Modellanpassung als Modell (1) erreicht hat, wurde zusätzlich noch Modell (3) gerechnet. Ob Modell (3) nun dem Modell (2) vorzuziehen ist, wurde wieder auf Basis der neuen Differenz $\Delta R^2 = R_3^2 - R_2^2$ entschieden. Für $\Delta R^2 \geq 0.02$ sollte also Modell (3) anstatt Modell (2) gewählt werden.

In einer weiteren Regressionsanalyse soll nun auf die geschlechtsbedingten Unterschiede bezüglich der Durchschnittsgeschwindigkeit eingegangen werden. Dazu verwenden wir wieder lineare Regressionsmodelle der Form (1) - (3), wobei wir nun als neue abhängige Variable \check{Y} die Differenz

aus den männlichen Durchschnittsgeschwindigkeiten Y^m und den weiblichen Durchschnittsgeschwindigkeiten Y^w verwenden, also $\check{y}_i := y_i^m - y_i^w, i = 1, \dots, n$. Die jeweilige Modellentscheidung basiert analog zu den Modellen (1) - (3) auf den Differenzen ΔR^2 .

3.2.2 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Regressionsanalysen für die verschiedenen Disziplinen und Geschlechter anhand von zwei Tabellen präsentiert. Die erste Tabelle zeigt die Ergebnisse der Entwicklung der Schwimmrekorde von männlichen und weiblichen Eliteschwimmern. Anschließend folgt die zweite Tabelle mit der Entwicklung der Schwimmgeschwindigkeit in anbetracht der geschlechtlichen Differenz.

Entwicklung amerikanischer Schwimmrekorde

Tabelle 1 beinhaltet die standardisierten β -Koeffizienten der jeweiligen Regressionsgleichung, für die die beste Datenanpassung erzielt wurde, sowie die Ergebnisse der Overall-F-Tests, R^2 , ΔR^2 und die p -Werte.

Disziplin	Standardisierte Regressionskoeffizienten			F-Test	R^2	ΔR^2
	β_1	β_2	β_3			
50 m Freistil:						
Männer	2.548	-1.723		F(2, 25) = 198.64*	0.94	0.17
Frauen	0.978			F(1, 26) = 573.83*	0.96	
100 m Freistil:						
Männer	2.417	-1.547		F(2, 25) = 1391.82*	0.99	0.14
Frauen	2.035	-1.123		F(2, 43) = 710.08*	0.97	0.08
200 m Freistil:						
Männer	3.895	-5.950	2.949	F(3, 42) = 1391.82*	0.99	0.03
Frauen	4.392	-6.111	2.589	F(3, 42) = 687.91*	0.98	0.03
100 m Rücken:						
Männer	2.087	-1.179		F(2, 43) = 805.38*	0.97	0.03
Frauen	1.983	-1.061		F(2, 43) = 981.38*	0.98	0.07
100 m Brust:						
Männer	1.810	-0.881		F(2, 43) = 511.78*	0.96	0.05
Frauen	2.036	-1.138		F(2, 43) = 389.59*	0.95	0.06
100 m Schmetterling:						
Männer	4.141	-7.187	4.096	F(3, 42) = 468.87*	0.97	0.07
Frauen	2.465	-1.626		F(2, 43) = 383.85*	0.95	0.16

* $p \leq 0.01$

Tabelle 1: Regressionskoeffizienten, F-Test, R^2 , ΔR^2 und p -Werte für den Verlauf der 12 amerikanischen Schwimmrekorde (Geschlecht · 6 Disziplinen) von 1962 bis 2007.

Folgende Punkte fallen bei der Regressionsanalyse auf: Zunächst weisen die einzelnen Overall-F-Tests darauf hin, dass alle Modelle für die Beschreibung des Sachverhaltes geeignet sind. Die Teststatistiken sind alle auf einem 0.1 Niveau signifikant und sind jeweils größer als ihre Quantile und somit kann die Nullhypothese aus (H1) verworfen werden.

Zudem demonstrieren die Ergebnisse der einzelnen Bestimmtheitsmaße R^2 , dass der Anteil der Streuung sehr gut durch die Modelle erklärt wird.

In Tabelle 1 sind die β -Koeffizienten einer standardisierten Regression angegeben. Das bedeutet, dass sowohl die unabhängigen Variablen X, X^2, X^3 als auch der Response Y in den Modellen (1) - (3) standardisiert wurden und anschließend mit den standardisierten Größen die Regression durchgeführt wurde. Eine Zufallsvariable Z wird standardisiert, indem man ihren Erwartungswert subtrahiert und sie durch ihre Standardabweichung teilt. Für die neue Zufallsvariable

$$\tilde{Z} = \frac{Z - \mu_Z}{\sigma_Z}$$

gilt dann

$$E[\tilde{Z}] = 0, \quad \text{Var}(\tilde{Z}) = 1.$$

Üblicherweise wird vor allem dann standardisiert, wenn die verschiedenen Einflussgrößen sehr unterschiedliche Maßeinheiten bzw. Wertebereiche besitzen, um ihre Effekte besser vergleichbar zu machen. Die standardisierten Regressionskoeffizienten geben dann die Änderung als ein Vielfaches der Standardabweichung der abhängigen Variable Y an, die sich ergibt, wenn eine unabhängige Variable $X_k, k = 1, \dots, p$ um genau eine Standardabweichung σ_{X_k} erhöht wird, während alle anderen Variablen festgehalten werden. Dies wird in der folgenden Rechnung am Beispiel der linearen Einfachregression gezeigt. Bei gegebener Stichprobe werden alle Erwartungswerte durch das arithmetische Mittel und alle Standardabweichungen durch die empirische Standardabweichung ersetzt und man erhält:

$$\begin{aligned} \tilde{y} &= \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x} & (4) \\ \Leftrightarrow \frac{y - \bar{y}}{s_Y} &= \tilde{\beta}_0 + \tilde{\beta}_1 \frac{x - \bar{x}}{s_X} \\ \Leftrightarrow \frac{y}{s_Y} &= \tilde{\beta}_0 + \frac{\bar{y}}{s_Y} - \tilde{\beta}_1 \frac{\bar{x}}{s_X} + \tilde{\beta}_1 \frac{1}{s_X} x \\ \Leftrightarrow y &= \underbrace{\left(\tilde{\beta}_0 + \frac{\bar{y}}{s_Y} - \tilde{\beta}_1 \frac{\bar{x}}{s_X} \right)}_{:=\beta_0} s_Y + \underbrace{\tilde{\beta}_1 \frac{s_Y}{s_X}}_{:=\beta_1} x & (5) \end{aligned}$$

Diese Rechnung zeigt, wie man die standardisierten Regressionskoeffizienten $\tilde{\beta}_k$ aus (4) zu den nicht standardisierten Regressionskoeffizienten β_k aus (5) zurücktransformieren kann. Setzt man in (5) $x = s_X$ ein, dann ändert sich Y um $\tilde{\beta}_1 s_Y$. Der Regressionskoeffizient $\tilde{\beta}_1$ gibt also den Anteil von s_Y an, um den sich Y verändert.

Die folgende Abbildung zeigt die standardisierten Regressionskurven beider Geschlechter für die Disziplinen 50, 100 und 200 m Freistil sowie 100 m Schmetterling aus Tabelle 1. Um die Kurven besser vergleichbar zu machen, wurden sie auf dem Bereich $x \in [-1.5, 1.5]$ und ohne Berücksichtigung des Intercepts β_0 dargestellt. Es fällt auf, dass sich beispielsweise für die Disziplin 200 m Freistil, kaum ein Unterschied zwischen Mann und Frau feststellen lässt. Ande-

reerseits wird der unterschiedliche Kurvenverlauf von Männern und Frauen für die Disziplinen 50 m Freistil und 100 m Schmetterling besonders deutlich. Eine Zwischenform stellen die Kurven für die Disziplin 100 m Freistil dar, dort sind Unterschiede erkennbar, die jedoch nicht gravierend sind. Insgesamt läßt sich also keine eindeutige Entwicklung der Geschwindigkeitsdifferenzen zwischen Mann und Frau erkennen, da beide Kurven sich je nach Disziplin bzw. für die selbe Disziplin je nach Distanz völlig unterschiedlich zu einander verhalten.

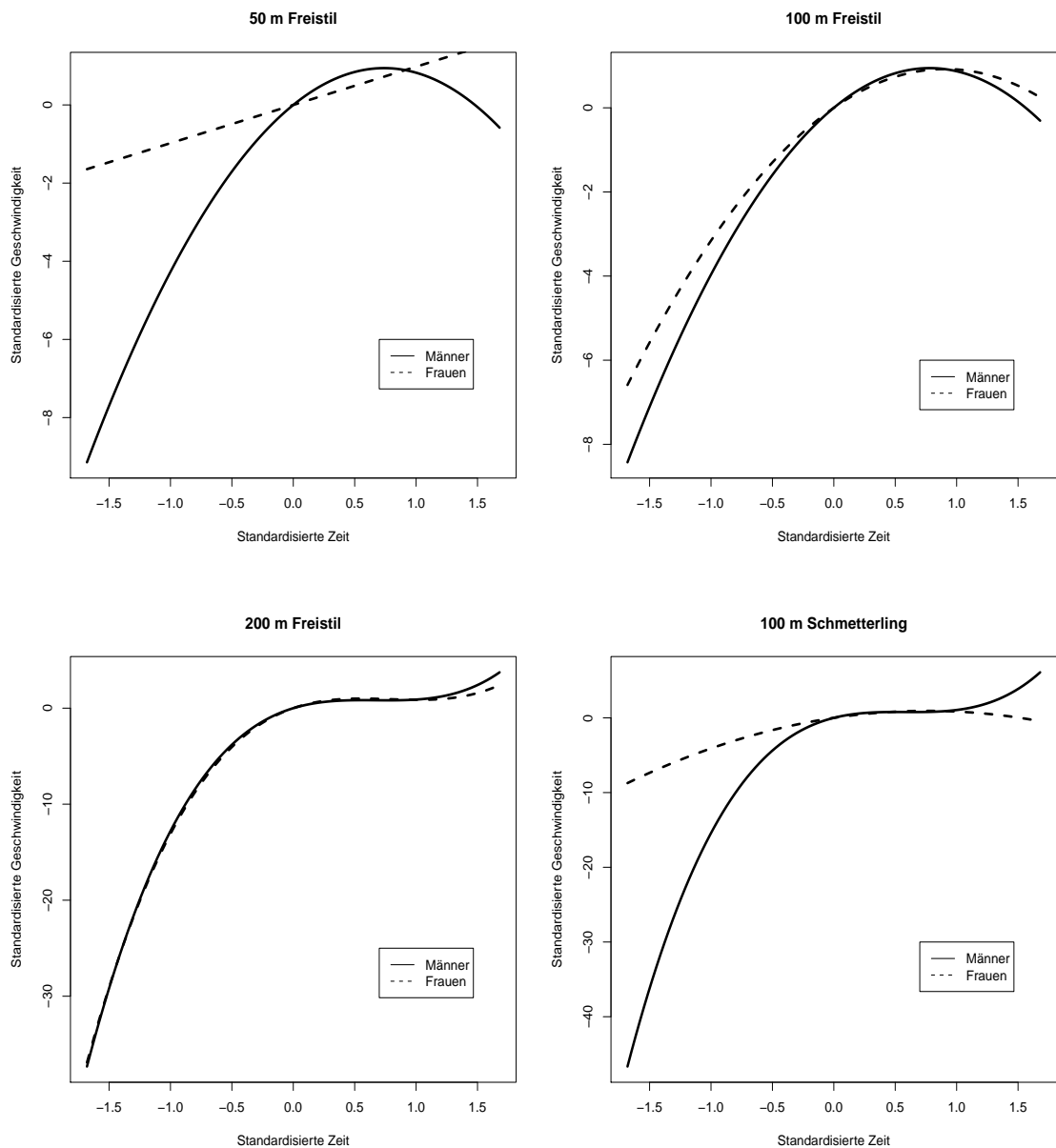


Abbildung 1: Vergleich der standardisierten Regressionsfunktionen beider Geschlechter für die Disziplinen 50, 100 und 200 m Freistil sowie 100 m Schmetterling

Die weiteren Regressionskurven aus Tabelle 1 werden im Anhang A aufgeführt.

Von allen zwölf Untersuchungseinheiten hat sich herausgestellt, dass nur bei der Disziplin 50 m

Freistil Frauen eine lineare Regression geeignet ist, d. h. dass sich die Geschwindigkeit konstant über die Jahre von 1962 bis 2007 erhöht hat.

Alle weiteren Regressionsfunktionen sind zweiten oder dritten Grades, weil die Berechnung des R jeweils einen höheren Wert als 0.02 ergeben hat. Die β -Koeffizienten zweiten Grades haben jeweils ein negatives Vorzeichen, somit handelt es sich hier um nach unten geöffnete Parabeln, die zunächst eine starke Steigung besitzen, jedoch am Ende immer flacher werden. In unserem Fall bedeutet das, dass die Steigung der Geschwindigkeitsentwicklung der Schwimmer im Vergleich zu den ersten Jahren abgenommen hat.

Für die Untersuchungseinheiten 200 m Freistil Männer und Frauen sowie 100 m Schmetterling Männer hat sich herausgestellt, dass eine Regressionsgleichung dritten Grades die beste Anpassung erreicht hat. Dieses Ergebnis demonstriert eine zunächst langsamer werdende Steigung der Geschwindigkeitsentwicklung, auf die später eine Zunahme der Geschwindigkeiten folgt.

Dieser Verlauf kann beispielsweise bei der Disziplin 100 m Schmetterling beobachtet werden. Gemäß Johnson et al. (2009) hat sich der Rekord bis 2000 hier nur langsam verbessert. Allerdings verbesserte sich die Zeit auf 50.40 Sekunden im Jahr 2005, während der Rekord im Jahr 2000 noch bei einer Zeit von 52.76 Sekunden lag.

Geschlechtlicher Vergleich der Entwicklung amerikanischer Schwimmrekorde

So wie in Tabelle 1 beinhaltet Tabelle 2 ebenfalls die standardisierten β -Koeffizienten der jeweiligen Regressionsgleichung, für die die beste Datenanpassung erzielt wurde, sowie die Ergebnisse der einzelnen Overall-F-Tests, R^2 , ΔR^2 und die p -Werte.

Disziplin	Standardisierte Regressionskoeffizienten			F-Test	R^2	ΔR^2	p -Wert
	β_1	β_2	β_3				
50 m Freistil:	4.668	-8.560	3.619	F(3, 24) = 11.69	0.59	0.05	≤ 0.01
100 m Freistil:	0.745	1.832	-2.720	F(3, 42) = 9.49	0.40	0.03	≤ 0.01
200 m Freistil:	-3.069	3.207		F(2, 43) = 34.79	0.62	0.62	≤ 0.01
100 m Rücken:	1.062	-0.947		F(2, 43) = 1.73	0.07	0.05	0.19
100 m Brust:	-1.114	0.894		F(2, 43) = 2.62	0.11	0.05	0.08
100 m Schmetterling:	1.985	-9.351	7.284	F(3, 42) = 26.91	0.66	0.23	≤ 0.01

Tabelle 2: Regressionskoeffizienten, F-Test, R^2 , ΔR^2 und p -Werte für die Differenzen der Geschwindigkeitsrekorde zwischen Männern und Frauen für den Zeitraum 1962 bis 2007.

Im Vergleich zu Tabelle 1 ergaben die F-Tests, dass nicht bei allen Disziplinen ein geeignetes Regressionsmodell gefunden werden konnte. Die Modelle für Brust- und Rückenstil sind unpassend, da sich jeweils beim Testen ein p -Wert größer als 0.05 ergeben hat und man somit nicht mehr signifikant sagen kann, dass sich die Modelle für die Situation eignen.

Außerdem sind alle Werte der Bestimmtheitsmaße vergleichsweise zur ersten Tabelle schlechter und verweisen auf eine geringere Modellanpassung bei allen sechs Disziplinen.

Darüber hinaus wurde beim Vergleich der Modelle durch R nicht ein einziges Mal das lineare

Modell gewählt, d. h. es gibt keine Hinweise darauf, dass sich die geschlechtsspezifischen Unterschiede bezüglich der Geschwindigkeit konstant verbessert haben. Eher im Gegenteil wurden parabolische oder noch dynamischere Modelle für signifikant befunden. Die folgende Abbildung zeigt die Kurven der vier signifikanten Modelle aus Tabelle 2. Diese Modelle ermöglichen anders als im linearen Fall keine eindeutige Interpretation der Entwicklung der geschlechtlichen Leistungsunterschiede. Die Kurven der beiden anderen Disziplinen sind wieder im Anhang A aufgeführt.

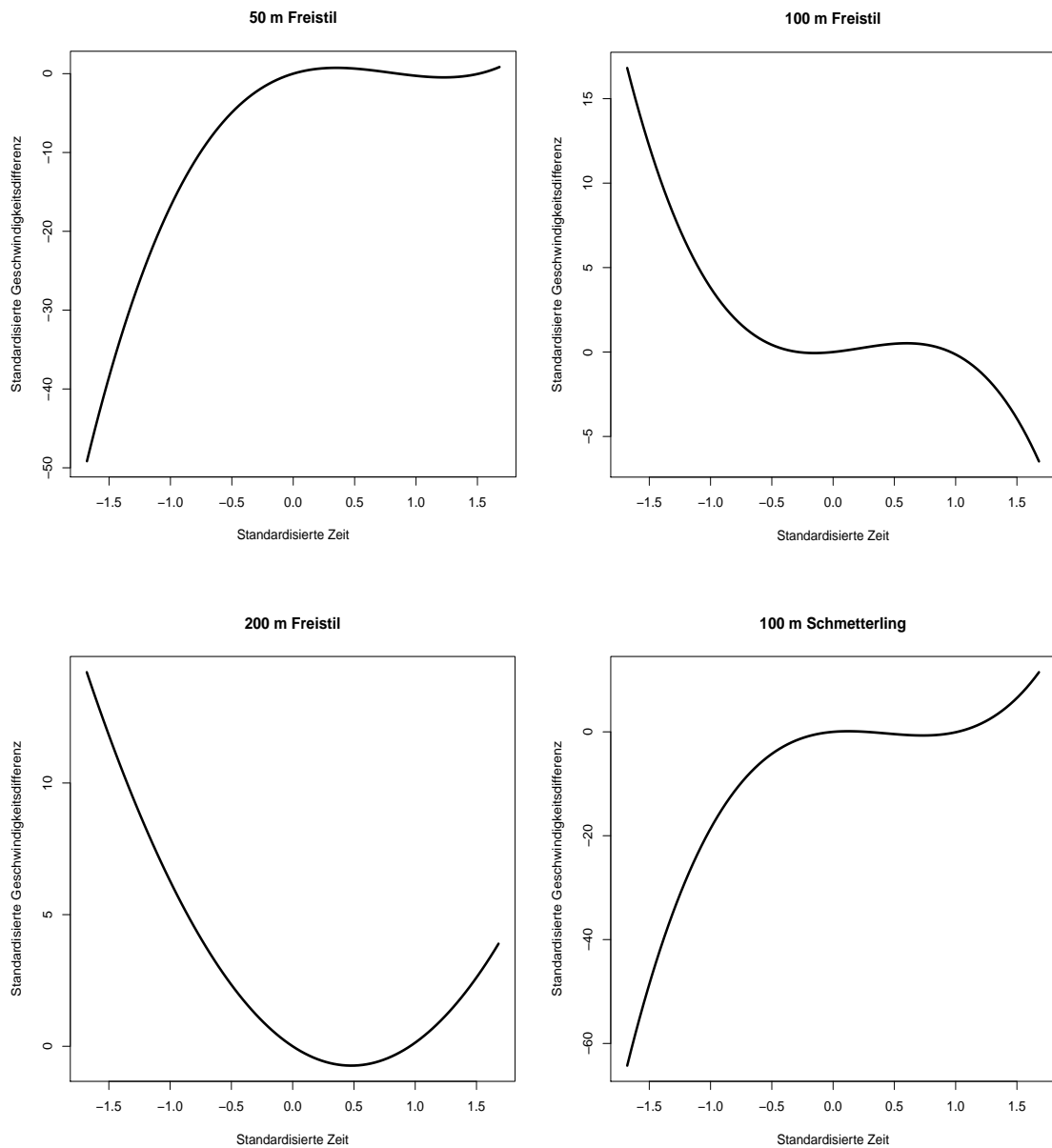


Abbildung 2: Standardisierte Regressionsfunktionen der Geschwindigkeitsdifferenzen von Mann und Frau für die Disziplinen 50, 100 und 200 m Freistil sowie 100 m Schmetterling

3.3 2. Studie

Die Daten der zweiten Studie stammen ebenfalls von der Website USA Swimming (USAS, 2009). Zu den zu untersuchenden Disziplinen gehören wie schon bei der ersten Studie: 100 m Freistil, Rücken-, Brustschwimmen und Schmetterling, sowie 200 m und 50 m Freistil.

Von diesen sechs Disziplinen wurden jährlich die 16 besten männlichen und weiblichen Zeiten aufgezeichnet, wobei für die Analyse nur die Jahre 1962, 1969, 1982, 1992 und 2002 verwendet worden sind. Aufzeichnungen von den Disziplinen Brust-, Rückenschwimmen und Schmetterling haben allerdings erst ab 1980 stattgefunden.

In dem Datensatz wurden Schwimmer von fünf Altersgruppen betrachtet, nämlich 10 Jahre & jünger, 11-12 Jahre, 13-14 Jahre, 15-16 Jahre und 17-18 Jahre. Für alle fünf Altersgruppen wurden für die fünf verschiedenen Jahreszahlen für drei oder sogar alle sechs Disziplinen Bestenlisten mit maximal 16 Geschwindigkeiten erfasst, so dass der Datensatz insgesamt 3614 Einträge erhält.

Nun soll wieder die Entwicklung der männlichen und weiblichen Durchschnittsgeschwindigkeit untersucht werden, zunächst einzeln und anschließend im Vergleich zu einander. Hierfür wird ebenfalls die abhängige Variable Y „Geschwindigkeit“, bzw. für den geschlechtlichen Vergleich die Differenz \check{Y} , auf die unabhängige Variable X „Jahreszahl“ regressiert. Zusätzlich wird die Stärke von geschlechtsspezifischen Effekten analysiert.

3.3.1 Methode

Für die Regressionsanalyse hinsichtlich der allgemeinen Entwicklung der Durchschnittsgeschwindigkeit bzw. hinsichtlich der geschlechtlichen Leistungsunterschiede werden analog zu Studie 1 wieder die Modelle (1) - (3) verwendet. Die jeweilige Modellentscheidung basiert wieder auf den Differenzen ΔR^2 .

Für die Analyse geschlechtsspezifischer Effekte wurden die Effektgrößen mittels *Cohens d* gemäß

$$d = \frac{\bar{y}_m - \bar{y}_w}{s_m}$$

für alle Altersgruppen und Disziplinen berechnet, wobei \bar{y}_m das arithmetische Mittel aus den Geschwindigkeiten einer entsprechenden Kombination aus Alter und Disziplin für männliche Schwimmer bezeichnet, analog \bar{y}_w für die entsprechende weibliche Gruppe. Die Differenz aus den Mittelwerten wird dann noch durch die empirische Standardabweichung s_m der männlichen Gruppe dividiert.

Der zu Cohens d gehörige Test mit den Hypothesen

$$H_0 : \mu_m = \mu_w \quad vs. \quad H_1 : \mu_m \neq \mu_w \quad (\text{H2})$$

dient zur Überprüfung, ob signifikante Mittelwertunterschiede in zwei gleich großen Gruppen vorliegen. Die Testentscheidung erfolgt unter Berücksichtigung des p -Werts.

3.3.2 Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse der Regressionsanalysen für die verschiedenen Disziplinen und Geschlechter erläutert. Anschließend werden die Ergebnisse der Effektgrößen anhand Tabelle 3 veranschaulicht und erklärt.

Entwicklung der Schwimmleistung amerikanischer Elitejugendschwimmer

Die Regressionsanalyse hat für alle 60 Disziplinen und Altersgruppenkombinationen ergeben, dass als Regressionsfunktion Polynome mindestens zweiten Grades ausgewählt werden. Dabei wurde in 52 der 60 Kombinationen der beste Fit mit einem Polynom dritten Grades erzielt, was die Vermutung bestätigt, dass keine konstante Geschwindigkeitsverbesserung oder -verschlechterung über die Zeitspanne 1962-2002 stattgefunden hat.

In den übrigen acht Fällen scheint aufgrund der parabolischen Form der Geschwindigkeitsanstieg mit der Zeit abzunehmen.

Im Vergleich zur Studie 1 fällt auf, dass sich die Modellanpassung verschlechtert hat. Dabei liegen die Werte des Bestimmtheitsmaßes R^2 jetzt nur noch zwischen 0.141 und 0.744. Darüber hinaus ist zu erwähnen, dass dennoch die Overall-F-Tests in 57 Fällen ein signifikantes Ergebnis erzielten und somit der Sachverhalt durch die Modelle sinnvoll beschrieben wird. Die einzelnen Ergebnissen werden in den Tabellen 3 und 4 in Johnson et al. (2009) präsentiert.

Geschlechtlicher Vergleich der Schwimmleistung amerikanischer Elitejugendschwimmer

Bei der Untersuchung der geschlechtsspezifischen Leistungsunterschiede ergab sich nur in drei von 30 Fällen ein lineares Regressionsmodell als am besten geeignet, nämlich in den Disziplinen 10 Jahre & jünger 50 m Freistil, 13-14 Jahre 200 m Freistil und 15-16 Jahre 100 m Rückenschwimmen.

Das bedeutet, dass in diesen drei Fällen die Differenz zwischen weiblichen und männlichen Elitejugendschwimmern konstant ab- oder zugenommen hat. In den übrigen Fällen unterliegt die Entwicklung der geschlechtlichen Geschwindigkeitsunterschiede einem komplizierterem Verlauf und erschwert somit die Interpretation.

Geschlechtsspezifische Effekte

In der folgenden Tabelle sind die Effektgrößen für die zwischengeschlechtlichen Geschwindigkeitsdifferenzen angegeben.

Disziplin	Altersgruppen				
	10 & jünger	11-12	13-14	15-16	17-18
50 m Freistil:	0.63**	2.44**	5.97**	5.33**	8.14**
100 m Freistil:	1.00**	2.00**	5.08**	5.51**	7.36**
200 m Freistil:	0.81**	1.64**	3.60**	6.49**	6.21**
100 m Rücken:	0.39	1.30**	3.29**	4.81**	5.36**
100 m Brust:	0.18	1.14**	2.37**	3.94**	5.01**
100 m Schmetterling:	0.60**	2.08**	3.09**	4.99**	5.22**

Tabelle 3: Differenzen der Effektgrößen und zugehörige p -Werte (* für $p < 0.05$, ** für $p < 0.01$) für den Geschlechtsvergleich bei amerikanischen Elitejugendschwimmern nach Alter, Schwimmdisziplin und Bahnlänge für den Zeitraum 1962 bis 2007.

Für 28 der 30 Kombinationen aus Altersgruppe und Schwimmstil bestehen signifikante Mittelwertunterschiede zwischen den Geschlechtern. Der zugehörige Test zur Hypothese (H2) wies einen p -Wert unter 0.01 auf. Die beiden Ausnahmen bilden die Gruppen 10 Jahre & jünger bei den Disziplinen 100 m Rücken- sowie Brustschwimmen.

Bis auf zwei Ausnahmen sieht man, dass die Effektgrößen mit steigendem Alter der Schwimmer zunehmen.

Darüber hinaus kann innerhalb einer Altersgruppe eine gewisse Heterogenität über die verschiedenen Disziplinen festgestellt werden. Die Effektgröße nimmt in allen Altersgruppen immer beim Brustschwimmen den kleinsten und bei einer der Freistildisziplinen den höchsten Wert an.

4 Zusammenfassung

Ziel dieser Arbeit ist in erster Linie die Untersuchung der Geschwindigkeitsentwicklung von erwachsenen und heranwachsenden Eliteschwimmern sowie der geschlechtsspezifische Unterschied hinsichtlich der Geschwindigkeitsentwicklung. Die Untersuchungsergebnisse werden dann mit den unterschiedlichen bereits bestehenden Hypothesen aus früheren Artikeln bezüglich der Geschwindigkeitsentwicklung im Sport verglichen, nämlich einerseits, dass zukünftig der geschlechtsspezifische Unterschied fast verschwinden wird (vgl. Tatem et al., 2004 und Whipp and Ward, 1992), und andererseits, dass der männliche Rekord immer über dem der Frauen liegen wird (vergleiche hierzu Chevront et al., 2005, Nevill and Whyte, 2005, Seiler et al., 2007 und Sharp, 2004). Die Auswertungen dieser Arbeit sprechen weder für die Hypothese, dass geschlechtsspezifische Leistungsunterschiede in der Zukunft nicht mehr existieren werden, noch für die Ansicht, dass Männer im Vergleich zu Frauen auch in Zukunft immer schneller

bleiben werden, da die Ergebnisse bei verschiedenen Disziplinen und Distanzen stark variieren und somit keine eindeutige Interpretation der Geschwindigkeitsentwicklung zulassen.

Nur bei den Auswertungen der Effektgrößen der einzelnen Stichproben zeigt sich, dass die normierten männlichen Durchschnittsgeschwindigkeiten höher liegen, als diejenigen der Frauen und sich sogar noch mit wachsendem Alter der Schwimmer verstärken, da alle Werte in Tabelle 3 positiv sind und mit dem Alter der Schwimmer steigen. Des Weiteren ergeben die zugehörigen Tests, dass die Hypothese $\mu_m = \mu_w$ in fast allen Fällen signifikant verworfen werden kann und somit die Erwartungswerte der Durchschnittsgeschwindigkeiten von Männern und Frauen unterschiedlich sind, also $\mu_m \neq \mu_w$.

5 Diskussion

Im folgenden Abschnitt werden die von Johnson et al. (2009) verwendeten statistischen Methoden kritisch hinterfragt.

Bei näherer Betrachtung der Auswertungen der ersten Studie ist aufgefallen, dass die durchgeführten Regressionsanalysen bemerkenswert gute Werte bei der Anpassung an die Daten erzielen. Interessant hierbei sind zunächst die Ergebnisse der Bestimmtheitsmaße. Die Werte liegen zwischen 0.94 und 0.99, was bedeutet, dass ein sehr hoher Anteil der Gesamtstreuung der Daten durch die Modelle erklärt wird.

Hier lässt sich nun fragen, ob es den überhaupt realistisch ist solche Werte bei solch einer Analyse zu erlangen. Wie man bei der zweiten Studie sehr gut sieht, sind die Werte des Bestimmtheitsmaßes im Vergleich zur ersten Studie schlechter, obwohl dieselbe Problemstellung, natürlich mittels unterschiedlicher Datensätze, untersucht worden ist. Die Tatsache, dass bei der Untersuchung des gleichen Sachverhaltes nur mit unterschiedlichen Datensätzen so unterschiedliche Ergebnisse erzielt werden, lässt vermuten, dass der Datensatz bzw. die Wahl der Datensätze eine ausschlaggebende Rolle spielt. Nun wurde in beiden Datensätzen die Geschwindigkeitsentwicklung, zum einen von erwachsenen Eliteschwimmern und zum anderen von heranwachsenden Eliteschwimmern, untersucht.

In Studie 1 wurde im Falle mehrerer neuer Rekorde innerhalb eines Jahres die Bestzeit der jeweiligen Eliteschwimmer genommen und falls in einem Jahr kein neuer Rekorde aufgestellt wurde, so wurde für dieses Jahr der bestehende Rekord verwendet. Bestehende Rekorde werden also so lange weiter verwendet, bis ein neuer Rekord aufgestellt wird, so dass die Geschwindigkeitsrekorde Y bei wachsender Jahreszahl X eine monoton steigende Folge bilden. Diese Tatsache ist aber äußerst kritisch zu sehen in Bezug auf die Anwendung des Basisregressionsmodells (BM) aus Kapitel 3.2.1, da die Modellannahme $E[\varepsilon_i] = 0$ verletzt wird. Denn bei normalverteilten, zu einander unabhängigen Störgrößen ε_i , könnte es passieren, dass sich zu einem späteren Zeitpunkt ein stark negatives ε_i realisiert und somit der spätere Geschwindigkeitsrekord unter dem vorherigen liegt. Man sieht auch anhand der Abbildungen 1, 2 und 4,

dass die Regressionsfunktionen in manchen Bereichen durchaus fallend sind, was aber mit der wahren Datenstruktur unvereinbar ist. Auch die Unabhängigkeitsannahme für die y_i scheint nicht gegeben, denn bei bekanntem y_t zum Zeitpunkt t , gilt für einen späteren Zeitpunkt $t + 1$: $y_t \leq y_{t+1}$, somit ist der Wertebereich des späteren y_{t+1} von der Realisation y_t zum früheren Zeitpunkt abhängig.

Ein weiteres Problem besteht darin, dass wenn beispielsweise fünf Jahre lang der Rekord nicht gebrochen wurde, ein möglicher Rückgang gar nicht berücksichtigt werden kann, weil nicht die Bestzeiten der Eliteschwimmer pro Jahr erfasst worden sind, sondern nur überbotene Rekorde. Im Gegensatz hierzu wurden in der zweiten Studie jeweils immer die Bestzeiten der heranwachsenden Eliteschwimmer erfasst, was viel sinnvoller für die Untersuchung der Geschwindigkeitsentwicklung erscheint. Somit wäre ein Datensatz in dem jeder Athlet nur mit einer einzigen Zeit, seiner Bestzeit, vertreten ist, unabhängig davon ob der Schwimmer den Weltrekord unterboten hat oder nicht, geeigneter.

Darüber hinaus sind nach Johnson et al. (2009) möglicherweise noch weitere Einflussfaktoren auf die Schwimmleistung zu berücksichtigen, wie beispielsweise die Auswirkung von Trainer, Herkunft, psychischer Verfassung, sozi-kultureller Umwelt sowie physiologische oder biomechanische Faktoren. Da gerade bei jungen Schwimmern zwischen den Geschlechtern noch keine großen Unterschiede bezüglich Körpergröße und Muskelkraft vorhanden sind, in Tabelle 3 aber eindeutig geschlechtsspezifische Unterschiede vorliegen, müssen weitere Faktoren berücksichtigt werden. Diese wirken auf den Reifeprozess eines jungen Menschen ein und können positiven oder negativen Einfluss auf seine Trainingsfortschritte nehmen und seine Entwicklung bis ins Erwachsenenalter prägen.

Ein weiterer Aspekt, der zur Erklärung der Effektgrößen aus Tabelle 3 beitragen könnte, besteht darin, dass bei jungen Sportlern im Training auch bewusst andere Spielformen, wie zum Beispiel Fußball oder Basketball, geübt werden, die keinen speziellen Trainingseffekt hinsichtlich der eigentlichen Sportart erkennen lassen, aber dennoch allgemeine Fähigkeiten eines Athleten schulen (siehe Berry et al., 2008 und Côté, 1999). Gerade in diesen Ergänzungsspielformen bestehen gemäß Côté et al. (2003) und Côté and Hay (2002) signifikante geschlechtliche Unterschiede bei der Wahl der Sportart, beispielsweise präferieren Jungen Fußball oder Basketball, während Mädchen lieber Volleyball oder Völkerball spielen. Dabei können die Jugendlichen verschiedenste, positive oder negative, persönliche oder zwischenmenschliche Erfahrungen machen, die sich bis ins Erwachsenenalter auswirken können.

6 Ausblick

Abschließend werden weitere Ideen genannt, wie die Sportart Schwimmen bzw. die Entwicklung der Geschwindigkeit im Schwimmsport untersucht werden kann. Zum einen könnte man diesen Sachverhalt mittels der *Extremwerttheorie* untersuchen. In Einmahl and Smeets (2009) wurde

bereits der Extremwertansatz auf die Leichtathletikdisziplin *100-Meter-Lauf* angewandt. Ziel im Kontext dieser Arbeit wäre dann, mittels des Extremwertansatzes anhand von Bestzeitdaten der Schwimmer eine Prognose der maximalen Rekordzeit zu schätzen. Grundlagen der Extremwerttheorie werden ausführlich in de Haan and Ferreira (2006) behandelt.

Zum anderen könnte man weitaus flexiblere Regressionsmodelle einsetzen, um die Geschwindigkeitsentwicklung bzw. die Leistung im Schwimmsport zu analysieren. In dieser Arbeit wurden nur die linearen Regressionsmodelle (1) - (3) verwendet, die alle auf dem Basismodell (BM) aus Kapitel 3.2.1 aufbauen. Die Flexibilität der drei Modelle ist beschränkt, da die Variable X „Jahreszahl“ ausschließlich *linearen* Einfluss besitzt. Erweitert man das Basismodell (BM) auf die allgemeine *additive* Form

$$y_i = f_1(z_{i1}) + \dots + f_q(z_{iq}) + \beta_0 + \beta_1 x_{1i} \dots + \beta_p x_{pi} + \varepsilon_i,$$

so kann für die Kovariablen z_1, \dots, z_q mittels der unspezifizierten glatten Funktionen f_1, \dots, f_q ein nichtlinearer Einfluss modelliert werden. Da für die Funktionen f_1, \dots, f_q keine restriktive funktionale Form vorgeschrieben ist, besitzt dieses nichtparametrische Modell hohe Flexibilität. Für die Schätzung werden die Funktionen $f_j, j = 1, \dots, q$ üblicherweise in Basisfunktionen entwickelt, d.h.

$$f_j(z_j) = \sum_{l=1}^{d_j} \alpha_{jl} B_l(z_j),$$

wobei die Basisfunktionen B_l Polynomsplines, beispielsweise B-Splines repräsentieren. Die Gewichte α_{jl} werden dann über Penalisierungsansätze geschätzt. Das additive Regressionsmodell wird ausführlich in Fahrmeir et al. (2007) behandelt.

Aufgrund der Aktualität neuartiger Schwimmanzüge, wie zum Beispiel nahtlose oder Haifischhaut ähnliche Anzüge, wäre ein Vergleich von den Zeiten der einzelnen Schwimmdisziplinen vor und nach der Verwendung dieser Bekleidung von Interesse, um feststellen zu können, ob die zahlreichen Rekorde der Jahre 2008 und 2009 wirklich auf die Bekleidung zurückzuführen sind. Außerdem lässt sich gut darüber debattieren, inwiefern die neu festgelegten Regeln der FINA, bezüglich der Bekleidung, vertretbar sind (vgl. FocusOnline, 2009). Denn in Anbetracht der Tatsache, dass in den letzten Jahrzehnten neue Techniken, wie zum Beispiel die Kraulwende beim Rückenschwimmen, eingeführt worden sind, die ebenfalls erheblichen Einfluss auf die Geschwindigkeit ausüben, lässt sich nun darüber diskutieren, weshalb neue Anzüge, die die Geschwindigkeit positiv beeinflussen, verboten werden sollten, während neuartige Schwimmtechniken erlaubt sind.

Anhang A

Hier finden sich die restlichen Graphiken der Regressionskurven aus Tabelle 1.

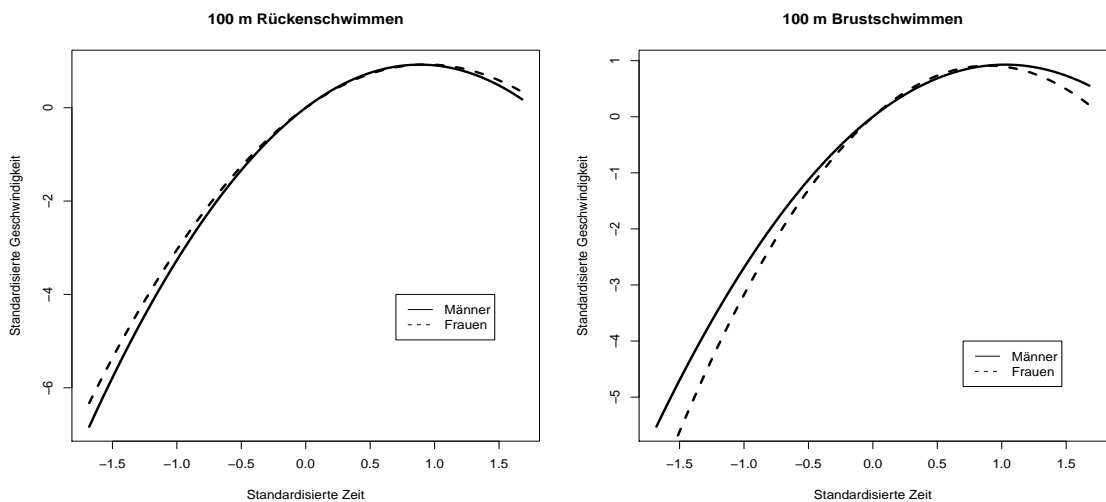


Abbildung 3: Vergleich der standardisierten Regressionsfunktionen beider Geschlechter für die Disziplinen Rückschwimmen (links) und Brustschwimmen (rechts) der Distanz 100 m

Der Vollständigkeit halber sind hier nun noch die Regressionskurven der Geschwindigkeitsdifferenzen zwischen Männern und Frauen aus Tabelle 2 für die Disziplinen Rücken- und Brustschwimmen dargestellt. Man beachte, dass diese Regressionsmodelle keine gute Anpassung erzielten, da ihre p -Werte den Wert 0.05 überschreiten. Johnson et al. (2009) haben die Daten einer graphischen Analyse unterzogen, die ergab, dass in beiden Fällen wohl eine Verletzung der Homoskedastizitätsannahme vorlag.

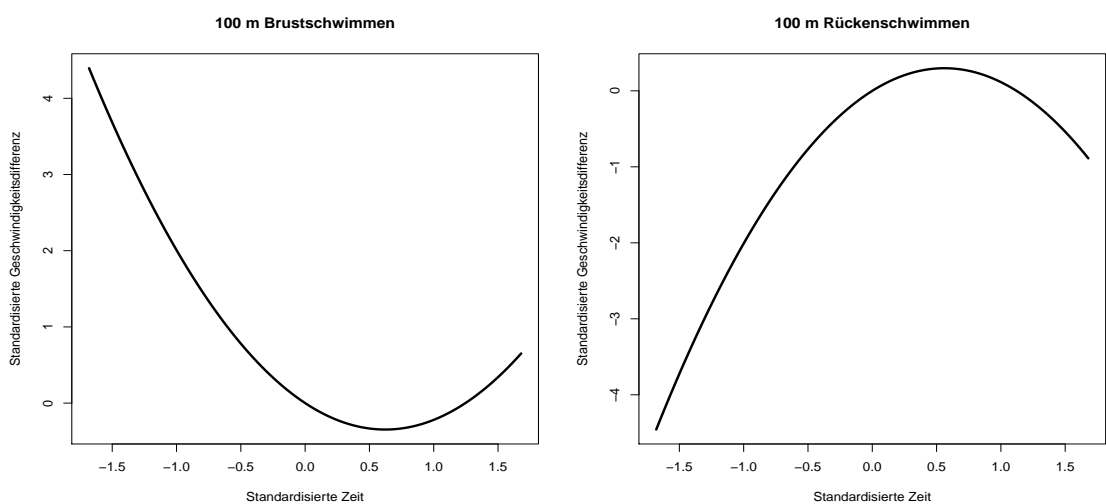


Abbildung 4: Standardisierte Regressionsfunktionen der Geschwindigkeitsdifferenzen von Mann und Frau für die Disziplinen 100 m Brust- (links) und Rückschwimmen (rechts)

Anhang B

Im Folgenden wird der R-Code präsentiert, mit dem die Graphiken aus Kapitel 3 erstellt wurden:

```
#####  
##### Polynomfunktion  
  
f <- function(b1, b2, b3) {  
  function(x) {  
    b1 * x + b2 * x^2 + b3 * x^3  
  }  
}  
  
##### x ist das standardisierte Jahr;  
##### y ist die standardisierte Geschwindigkeit;  
  
#####  
##### 50m Freistil (Tabelle 1):  
  
# Männer:  
  
plot(f(b1=2.548, b2=-1.723, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,  
xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeit", main="50 m Freistil")  
  
# Frauen:  
  
plot(f(b1=0.978, b2=0, b3=0), xlim = c(-1.68, 1.68), lwd="3", lty=2, add = TRUE)  
legend(0.7, -6, c("Männer", "Frauen"), lty=c(1,2))  
  
#####  
##### 100m Freistil (Tabelle 1):  
  
# Männer:  
  
plot(f(b1=2.417, b2=-1.547, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,  
xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeit", main="100 m  
Freistil")  
  
# Frauen:  
  
plot(f(b1=2.035, b2=-1.123, b3=0), xlim = c(-1.68, 1.68), lwd="3", lty=2, add =  
TRUE)  
legend(0.7, -6, c("Männer", "Frauen"), lty=c(1,2))  
  
#####  
##### 200m Freistil (Tabelle 1):  
  
# Männer:  
  
plot(f(b1=3.895, b2=-5.950, b3=2.949), xlim = c(-1.68, 1.68), type = 'l', lwd="3",  
lty=1, xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeit", main="200  
m Freistil")  
  
# Frauen:  
  
plot(f(b1=4.392, b2=-6.111, b3=2.589), xlim = c(-1.68, 1.68), lwd="3", lty=2, add  
= TRUE)  
legend(0.7, 25, c("Männer", "Frauen"), lty=c(1,2))
```

```

#####
##### 100m Rückenschwimmen (Tabelle 1):

# Männer:

plot(f(b1=2.087, b2=-1.179, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,
xlab="Standardisierte Zeit",ylab="Standardisierte Geschwindigkeit", main="100 m
Rückenschwimmen")

# Frauen:

plot(f(b1=1.983, b2=-1.061, b3=0), xlim = c(-1.68, 1.68), lwd="3", lty=2, add =
TRUE)
legend(0.7,-4,c("Männer","Frauen"),lty=c(1,2))

#####
##### 100m Brustschwimmen (Tabelle 1):

# Männer:

plot(f(b1=1.810, b2=-0.881, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,
xlab="Standardisierte Zeit",ylab="Standardisierte Geschwindigkeit", main="100 m
Brustschwimmen")

# Frauen:

plot(f(b1=2.036, b2=-1.138, b3=0), xlim = c(-1.68, 1.68), lwd="3", lty=2, add =
TRUE)
legend(0.7,-4,c("Männer","Frauen"),lty=c(1,2))

#####
##### 100m Schmetterling (Tabelle 1):

# Männer:

plot(f(b1=4.141, b2=-7.187, b3=4.096), xlim = c(-1.68, 1.68), type = 'l', lwd="3",
lty=1, xlab="Standardisierte Zeit",ylab="Standardisierte Geschwindigkeit", main="100
m Schmetterling")

# Frauen:

plot(f(b1=2.465, b2=-1.626, b3=0), xlim = c(-1.68, 1.68), lwd="3", lty=2, add =
TRUE)
legend(0.7,-30,c("Männer","Frauen"),lty=c(1,2))

#####
##### Abbildungen für Regressionskurven der Geschwindigkeitsdifferenzen

#####
##### 50m Freistil (Tabelle 2):

plot(f(b1=4.668, b2=-8.560, b3=3.619), xlim = c(-1.68, 1.68), type = 'l', lwd="3",
lty=1, xlab="Standardisierte Zeit",ylab="Standardisierte Geschwindigkeitsdifferenz",
main="50 m Freistil")

#####
##### 100m Freistil (Tabelle 2):

plot(f(b1=0.745, b2=1.832, b3=-2.720), xlim = c(-1.68, 1.68), type = 'l', lwd="3",
lty=1, xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeitsdifferenz",
main="100 m Freistil")

```

```
#####  
##### 200m Freistil (Tabelle 2):  
  
plot(f(b1=-3.069, b2=3.207, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,  
xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeitsdifferenz",  
main="200 m Freistil")
```

```
#####  
##### 100m Rückenschwimmen (Tabelle 2):  
  
plot(f(b1=1.062, b2=-0.947, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,  
xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeitsdifferenz",  
main="100 m Rückenschwimmen")
```

```
#####  
##### 100m Brustschwimmen (Tabelle 2):  
  
plot(f(b1=-1.114, b2=0.894, b3=0), xlim = c(-1.68, 1.68), type = 'l', lwd="3", lty=1,  
xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeitsdifferenz",  
main="100m Brustschwimmen")
```

```
#####  
##### 100m Schmetterling (Tabelle 2):  
  
plot((f(b1=1.985, b2=-9.351, b3=7.284), xlim = c(-1.68, 1.68), type = 'l', lwd="3",  
lty=1, xlab="Standardisierte Zeit", ylab="Standardisierte Geschwindigkeitsdifferenz",  
main="100m Schmetterling ")
```

Literatur

- Berry, J., B. Abernethy, and J. Côté (2008). The contribution of structured activity and deliberate play to the development of expert perceptual and decision-making skill. *Journal of Sport and Exercise Psychology* 30, 685–708.
- Chevront, S. N., R. J. Moffatt, and K. DeRuisseau (2005). Body composition and gender differences in performance. In J. A. Driskell & I. Wolinsky (Eds.). *Nutritional assessment of athletes*, Boca Raton, FL: CRC Press, 177–200.
- Côté, J. (1999). The influence of the family in the development of talent in sports. *The Sport Psychologist* 13, 395–417.
- Côté, J., J. Baker, and B. Abernethy (2003). From play to practice: A development framework for the acquisition of expertise in team sports. In J. M. Silva and D. Stephens (Eds.). *Expert performance in sports: Advances in research on sport expertise*, Champaign, IL: Human Kinetics, 89–110.
- Côté, J. and J. Hay (2002). Children’s involvement in sports: A developmental perspective. In J. M. Silva and D. Stevens (Eds.). *Psychological foundations of sports 2nd. edition*, Boston: Merrill, 484–502.
- de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: An Introduction*. New York: Springer.
- DSV (2009). Deutscher Schwimm-Verband e.v. <http://www.dsv.de/DesktopDefault.aspx>, Stand: 7.12.2009.
- Einmahl, J. H. J. and S. G. W. R. Smeets (2009). Ultimate 100m world records through extreme-value theory. *CentER Discussion Paper Series*.
- Fahrmeir, L., T. Kneib, and S. Lang (2007). *Regression*. Heidelberg: Springer Verlag.
- FINA (2009). Fédération Internationale de Natation. <http://fina.org/>, Stand: 7.12.2009.
- FocusOnline (2009). Schwimmanzug-Streit: Neue Regeln, alte Probleme. http://www.focus.de/sport/mehrsport/tid-13794/schwimmanzug-streit-neue-regeln-alte-probleme_aid_384386.html, Stand: 7.12.2009.
- Johnson, M. B., W. A. Edmonds, S. Jain, and J. C. Jr. (2009). Analyses of elite swimming performances and their respective between-gender differences over time. *Journal of Quantitative Analysis in Sports* 5, Issue 4, Article 2.
- Küchenhoff, H. (2009). Skript zur Vorlesung Lineare Modelle. http://www.stat.uni-muenchen.de/~helmut/limo09/skript_jm.pdf, Stand: 7.12.2009.
- Nevill, A. M. and G. Whyte (2005). Are there limits to running world records? *Medicine and Science in Sports and Exercise* 37, 1785–1788.
- Seiler, S., J. J. de Koning, and C. Forster (2007). The fall and rise of the gender difference in elite anaerobic performance 1952-2006. *Medicine and Science in Sports and Exercise* 39, 534–540.
- Sharp, N. C. (2004). Mind the gap: Women racers are falling behind. *Nature* 432, 147.
- Tatem, A. J., C. A. Guerra, P. M. Atkinson, and S. I. Hay (2004). Athletics: Momentous sprint at the 2156 Olympics? *Nature* 431, 525.
- USAS (2009). USA Swimming. <http://www.usaswimming.org/USASWeb/DesktopDefault.aspx?TabId=205&Alias=Rainbow&Lang=en>, Stand: 7.12.2009.
- Whipp, B. J. and S. A. Ward (1992). Will women soon outrun men? *Nature* 355, 25.
- Wikipedia (2009). Schwimmsport. <http://de.wikipedia.org/wiki/Schwimmsport>, Stand: 7.12.2009.