

**STUDIE ZUR EFFEKTIVITÄT  
DES AUDIVA TRAININGS FÜR KINDER**

**BERICHT ÜBER DIE ERGEBNISSE DER STATISTISCHEN  
BERECHNUNGEN MIT EXTREMGRUPPEN**

**ANDREA HINZ, DIPL. PSYCH., AUGUST 2008**

# INHALTSVERZEICHNIS

<b>1 Einleitung .....</b>	<b>3</b>
<b>2 Überblick über die „Landau-Studie“ .....</b>	<b>4</b>
2.1. Messinstrumente, Vorgehen und Gruppenbildung.....	4
2.1.1 Verwendete Messinstrumente .....	4
2.1.2 Vorgehen und zeitlicher Ablauf.....	5
2.1.3 Begründung der Gruppenbildung .....	6
2.1.4 Zusammensetzung der Gruppen .....	7
2.1.5 Beschreibung des AUDIVA Trainings.....	8
2.2 Überblick über die Ergebnisse der „Landau-Studie“ .....	10
2.2.1 Probleme bei der Durchführung.....	10
2.2.2 Ergebnisse .....	11
<b>3 Berechnungen mit Extremgruppen.....</b>	<b>13</b>
3.1 Herleitung und Einordnung der zusätzlichen Berechnungen .....	13
3.2 Testgütekriterien .....	15
3.3 Ausgeschlossene Tests bzw. Fragen.....	15
<b>4 Ergebnisse .....</b>	<b>17</b>
4.1 Ergebnisse des AUDIVA-Aufmerksamkeitstesters.....	17
4.2 Ergebnisse des Audiva-Hörscreeners .....	20
4.2.1 Vergleich von Extremgruppe und Kontrollgruppe .....	21
4.2.2 Die Gruppe der „Hochsensiblen“ .....	26
4.3 Ergebnisse der Audiva-Test-CD .....	29
4.3.1 Ergebnisse des Lautunterscheidungstests Stufe 2.....	29
4.3.2 Ergebnisse des Lautunterscheidungstests Stufe 3.....	32
4.3.4 Ergebnisse des dichotischen Hörtests.....	35
4.4 Ergebnisse des Fragebogens für Lehrer/innen .....	38
4.4.1 Vorgehen .....	38
4.4.2 Skalen .....	38
4.4.3 Ergebnisse .....	42
4.5 Ergebnisse des Abschlussfragebogens für Kinder.....	49
<b>5 Diskussion .....</b>	<b>50</b>
5.1 Limitierungen der Studie und Empfehlungen für das weitere Vorgehen .....	50
5.2 Fazit .....	51

# **1 EINLEITUNG**

In Zusammenarbeit mit der Universität Koblenz-Landau wurde eine Studie an Grundschulkindern durchgeführt, um die Effektivität des AUDIVA Trainings (Musiktherapie Phase A) zu überprüfen. Die Ergebnisse wurden von Prof. Günther und Prof. Petillon im März 2008 im Forschungsbericht "Erprobung und Evaluation des Hörtrainings: AUDIVA in dritten Schulklassen" vorgelegt.

Da einer der Professoren die Universität im Winter 2008 verließ, konnten keine weiteren Wünsche seitens AUDIVA bezüglich der Studienanalysen erfüllt werden. Daher wurde die Verfasserin dieses Berichtes, die freiberuflich im Rahmen des AUDIVA Projekts für die Universität tätig war (u.a. mit Datenmanagement und -auswertung) und die statistischen Berechnungen für den ursprünglichen Bericht anhand der im Projekt erhobenen Daten geliefert hatte, im April 2008 von AUDIVA beauftragt, weitere Analysen mit diesen Daten durchzuführen.

Im Folgenden wird zunächst ein Überblick über die „Landau-Studie“, die Messinstrumente, gebildete Gruppen, Forschungsdesign und Ablauf sowie die Ergebnisse gegeben, da diese die Grundlage für das weitere Vorgehen bilden. Danach wird auf die zusätzlichen Extremgruppenberechnungen eingegangen und deren Ergebnisse werden vorgestellt und diskutiert.

## **2 ÜBERBLICK ÜBER DIE „LANDAU-STUDIE“**

Um die Effektivität des AUDIVA Trainings zu überprüfen, wurde das Training in einem Zeitraum von drei Monaten in verschiedenen Schulen durchgeführt. Die teilnehmenden Kinder wurden von den Lehrerinnen ausgewählt. Die Lehrerinnen waren angehalten, diejenigen Kinder zu bestimmen, die von ihnen als allgemein förderbedürftig eingestuft wurden.

### **2.1. MESSINSTRUMENTE, VORGEHEN UND GRUPPENBILDUNG**

#### **2.1.1 VERWENDETE MESSINSTRUMENTE**

Zur Erfassung positiver Effekte des AUDIVA Trainings wurden in der „Landau-Studie“ verschiedene Tests bzw. Fragebögen eingesetzt:

1. Lehrerfragebogen – wurde von den Lehrerinnen der Kinder ausgefüllt.
2. Elternfragebogen – wurde von den Eltern der Kinder ausgefüllt.
3. Audiva-Testbogen (Studie Landau) für Kinder – in diesen wurden von den Studentinnen die Antworten der Kinder auf die Aufgaben der Audiva-Test-CD eingetragen. Dabei wurden nicht alle Tests der Audiva-Test-CD verwendet, sondern für die Studie besonders geeignete Subtests ausgewählt (Nachsprechen von Nonsenssilben, drei verschiedene Lautunterscheidungstests, dichotischer Hörtest, Geschichte nacherzählen).
4. Audiva-Aufmerksamkeitstester für die Kinder.
5. Audiva-Hörscreener für die Kinder.
6. Abschlussfragebogen für die Kinder.

Die entsprechenden Dokumente zu den verwendeten Testmaterialien sind im Internet kostenfrei verfügbar. Geben Sie "forschung.audiva.de" ein, und klicken Sie auf "Landauer Studie".

### 2.1.2 VORGEHEN UND ZEITLICHER ABLAUF

Zwei Gruppen von Studentinnen führten mit Schülerinnen und Schülern das AUDIVA Training durch. Das Training der ersten Gruppe fand im Herbst 2006 statt, je nach organisatorischen Gegebenheiten in den Schulen entweder von September 2006 bis November 2006 oder von Oktober 2006 bis Dezember 2006.

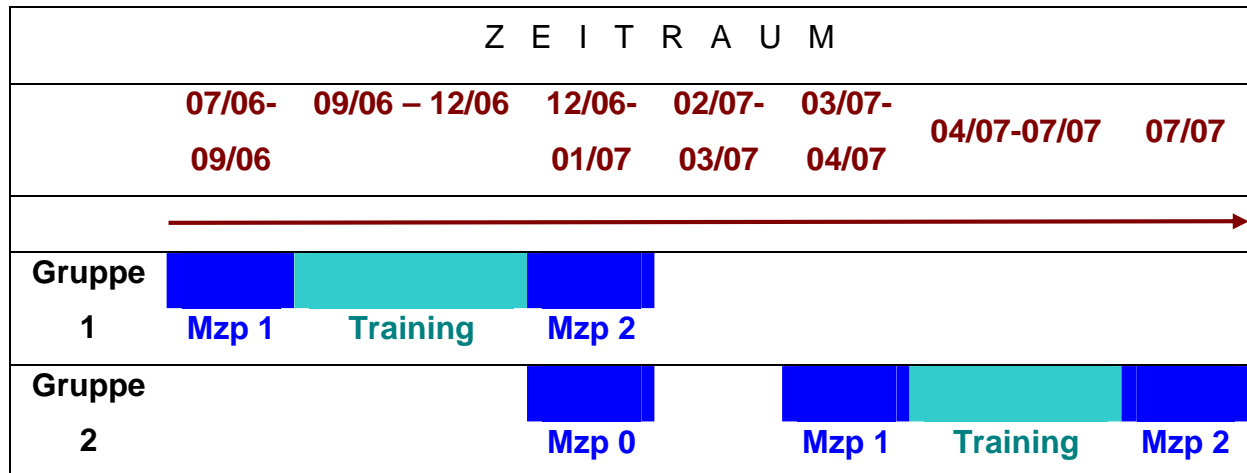
Die Befragung mit dem Lehrerfragebogen und dem Elternfragebogen und die Testung der Kinder mit der Audiva-Test-CD, dem Audiva-Aufmerksamkeitstester und dem Audiva-Hörscreener wurde jeweils vor und nach dem Training vorgenommen: Vor dem Training (Messzeitpunkt 1 = Mzp 1) je nach den schulischen Gegebenheiten entweder noch vor den Sommerferien 2006 (d.h. Juli 2006) oder im August bzw. September 2006, nach dem Training (Messzeitpunkt 2 = Mzp 2) im Dezember 2006 bzw. im Januar 2007.

In der zweiten Gruppe wurde das Training ebenfalls für einen Zeitraum von drei Monaten durchgeführt und zwar von ca. Mitte April 2007 bis ca. Mitte Juli 2007. Auch diese Gruppe führte vor und nach dem Training Befragungen mit dem Lehrer- und Elternfragebogen sowie Testungen der Kinder mit der Audiva-Test-CD, dem Audiva-Aufmerksamkeitstester und dem Audiva-Hörscreener durch. Nach dem Training erhielten die Kinder zusätzlich einen Abschlussfragebogen. Die Befragung und Testung vor dem Training (Messzeitpunkt 1) wurde je nach den organisatorischen Gegebenheiten im März 2007 oder im April 2007 vorgenommen, die Befragung und Testung nach dem Training (Messzeitpunkt 2) im Juli 2007.

Zusätzlich zu den beiden Befragungszeitpunkten vor und nach dem Training wurde in der zweiten Gruppe eine vorausgehende Befragung (Messzeitpunkt 0 = Mzp 0) der Eltern und Lehrer sowie eine Testung der Kinder mit den genannten Instrumenten durchgeführt, und zwar im Dezember 2006 bzw. im Januar 2007.

Insgesamt wurden also die Eltern, Lehrer und Kinder der zweiten Gruppe dreimal getestet bzw. befragt. (Zum genauen zeitlichen Ablauf siehe untenstehende Abbildung.)

## Zeitablauf der Testungen und Training für beide Gruppen



### 2.1.3 BEGRÜNDUNG DER GRUPPENBILDUNG

Der Grund für die Vortestung bei der dritten Gruppe bestand darin, eine sogenannte Kontrollgruppe zu bekommen.

Das Vorliegen einer Kontrollgruppe hat folgenden Grund: Testungen nur vor und nach einem Training (in einer Trainingsgruppe) reichen nicht aus, um den Effekt eines Trainings nachzuweisen. Mögliche Veränderungen können nicht eindeutig auf das Training zurückgeführt werden. Denkbar sind bei einfachen Vorher-Nachher-Testungen auch Veränderungen aufgrund von Testübungseffekten oder einfach aufgrund der natürlichen Entwicklung der Kinder in einem bestimmten Zeitraum.. Hat man aber eine Vergleichsgruppe, die in einem Abstand von drei Monaten ebenfalls zwei Mal getestet wird, ohne dass ein Training stattfindet, können diese beiden Messungen mit denen der Gruppe mit Training in Bezug gesetzt werden. Hierzu wird ein statistisches Verfahren (zweifaktorielle Varianzanalyse mit Messwiederholung) angewendet, das die Verläufe vergleicht. Ein Trainingseffekt, auch Treatmenteffekt genannt, liegt dann vor, wenn die Verbesserungen in der Trainingsgruppe mathematisch nachgewiesen (d.h. statistisch signifikant) größer sind als in der Vergleichsgruppe ohne Training.

Üblicherweise wird hierzu eine Kontrollgruppe verwendet, die der Trainingsgruppe bezüglich der wichtigen Kriterien entspricht und zu denselben Zeitpunkten wie die

Trainingsgruppe getestet wird, aber kein Training erhält. Aus ethischen und organisatorischen Gründen war dies in dieser Untersuchung nicht möglich. Daher wurde in der zweiten Gruppe ein vorausgehender Messzeitpunkt eingeführt, d.h. die zweite Gruppe hat drei Monate auf ihr Training gewartet, weshalb sie für diesen Zeitraum auch als Wartegruppe bezeichnet wird. So können die Veränderungen vom vorausgehenden Messzeitpunkt (Messzeitpunkt 0) bis zum Messzeitpunkt 1 in der Wartegruppe mit den Veränderungen zwischen Messzeitpunkt 1 und Messzeitpunkt 2 der ersten und zweiten Gruppe verglichen und Rückschlüsse auf die Effektivität des Trainings gezogen werden.

#### 2.1.4 ZUSAMMENSETZUNG DER GRUPPEN

Die erste Gruppe besteht aus 60 Kindern aus 8 Schulen, die zweite Gruppe aus 41 Kindern aus 5 Schulen. Von den 41 Kindern der zweiten Gruppe nahmen alle an den Tests der Messzeitpunkte 0 und 1 teil; 38 Kinder nahmen an den Tests der Messzeitpunkte 1 und 2 teil, weshalb sich die Wartegruppe (Gruppe 2: Messzeitpunkte 0 und 1) und die Trainingsgruppe 2 (Gruppe 2: Messzeitpunkte 1 und 2) etwas voneinander unterscheiden.

Für die Auswertungen werden die Messzeitpunkte 1 und 2 für die beiden Trainingsgruppen (Gruppe 1 und 2) beibehalten. Die Wartegruppe wird von Trainingsgruppe 2 getrennt und als eigenständige Gruppe behandelt. Die Messzeitpunkte 0 und 1 der Wartegruppe werden, der Übersichtlichkeit in den Tabellen mit den mathematischen Auswertungen halber, umbenannt, ebenfalls in den ersten Messzeitpunkt der Wartegruppe (Messzeitpunkt 1) und den zweiten Messzeitpunkt der Wartegruppe (Messzeitpunkt 2). Auf diese Weise können von jeder Gruppe (Trainingsgruppe 1 und 2 sowie Wartegruppe) jeweils ein erster und ein zweiter Messzeitpunkt (mit jeweils ca. drei Monaten Abstand) in Bezug auf mögliche Veränderungen verglichen werden. Das in der „Landau-Studie“ verwendete statistische Verfahren hierzu ist die zweifaktorielle Varianzanalyse mit Messwiederholung.

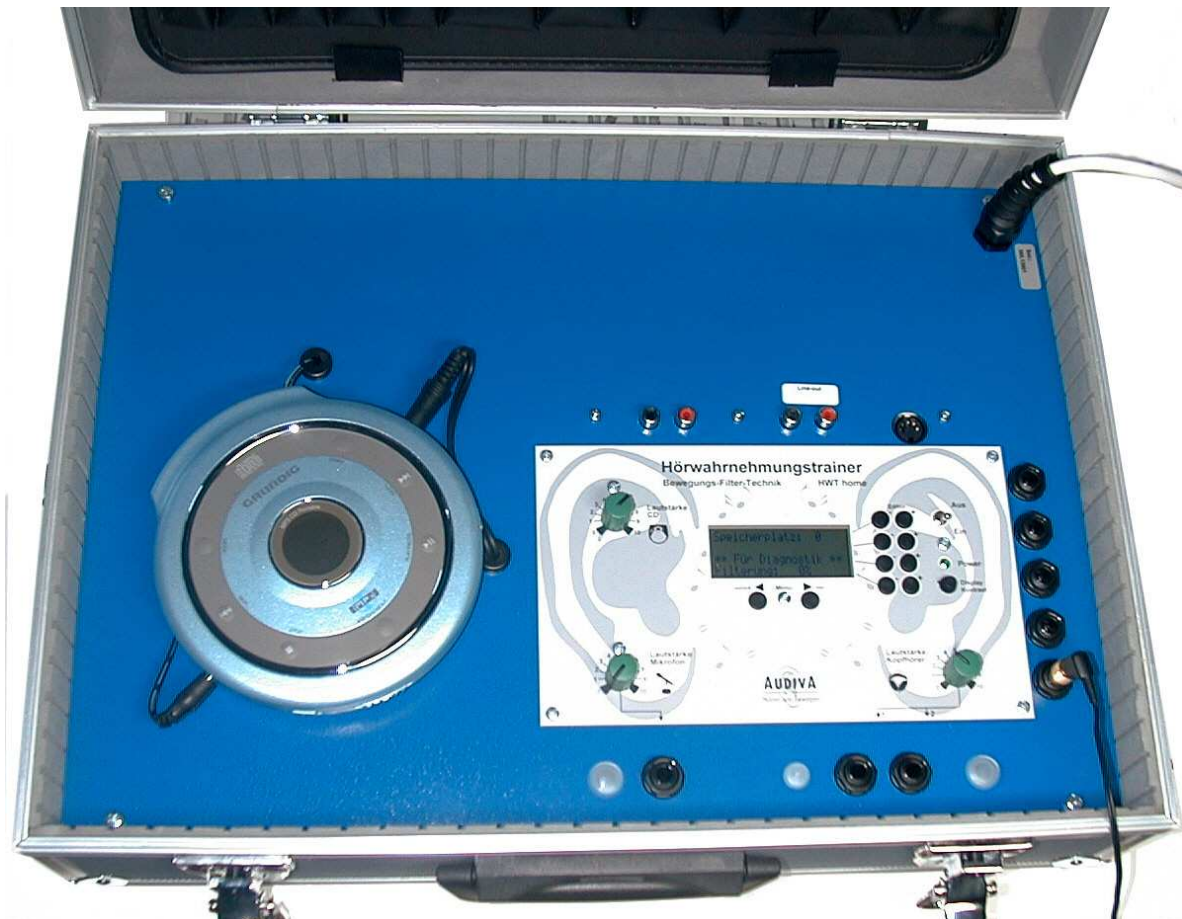
Von einem Treatmenteffekt dabei gesprochen, wenn die positiven Veränderungen in den Trainingsgruppen statistisch signifikant größer sind als in der Wartegruppe.

### 2.1.5 BESCHREIBUNG DES AUDIVA TRAININGS

Das im Rahmen der „Landau-Studie“ an den Schulen durchgeführte AUDIVA Training ist Bestandteil des von AUDIVA entwickeltem Hörwahrnehmungstrainings (kurz: HWT), bei dem über ein HWT-Gerät Hochtonttraining und Lateraltraining (verknüpft werden. Die Hintergründe diese Klangveränderung sind auf der Webseite [www.audiva.de](http://www.audiva.de) nachzulesen oder können dort schriftlich angefordert werden. Als „Trainingsmusik“ wurden die AUDIVA CD´s 1...7 verwendet, die instrumentale Titel von von Mozart (und etwas von Bach und Vivaldi) enthielten. Durch das HWT-home Gerät und einem schriftlichen Trainingsplan wurde den Studenten vorgegeben, die Steigerung der Level 4...6 (steigende Klangveränderung) am HWT home Gerät während der Trainingswochen einzustellen und zu überwachen (siehe Trainingsplan Studie Landau). Eine CD wurde eine Woche lang vorgespielt (wobei ein Wechsel der CD im Rahmen der sieben CDs tolerierbar war). Die Kinder hörten die Musik über beigestellte Kopfhörer vom Typ QP 160, die die spezifischen Klangveränderungen deutlich wiedergaben.

Das Musikhören wurde mit einer handlichen Geräte - Version durchgeführt. Um das Auf- und Abbauen übersichtlich zu gestalten, war für die „Studie Landau“ die Apparatur in einem Koffer untergebracht. Diese Kofferlösung wurde zusätzlich mit einem Kopfhörerverstärker ausgestattet, der es erlaubte bis zu 7 Kopfhörer direkt einzustecken (rechts 5 und unten links weitere 2 Kopfhörer). Weitere Kopfhörer konnten über 2 Fach Verteiler angeschlossen werden. Siehe Bild:





In der Regel wurde ein Gerät verwendet, in einigen Klassen auch zwei Geräte. Während des Musikhörens wurde in den Klassen „Stillarbeit“ durchgeführt, die Kinder konnten Aufgaben machen, malen o.ä. In den Klassen herrschte während dieser Zeit also Ruhe.

Nach einer Einweisung durch die Testleiterinnen (Studentinnen, die im Rahmen dieses Projektes ihre Examensarbeit schrieben), die auch die alle Testungen mit den Kindern vornahmen und die Fragebögen verteilten und einsammelten, bauten die Lehrerinnen das AUDIVA Training in den Unterricht ein. Die Studentinnen besuchten ihre Schule meist einen Tag pro Woche und machten an diesem Tag auch das AUDIVA Training mit den Kindern. Gleichzeitig sammelten Sie Erfahrungen im Austausch mit den Lehrern und bestätigten und korrigierten die Handlungen der LehrerInnen.

An allen anderen Tagen, stellten die LehrerInnen die Technik auf und führten das Training selbstständig durch

## 2.2 ÜBERBLICK ÜBER DIE ERGEBNISSE DER „LANDAU-STUDIE“

### 2.2.1 PROBLEME BEI DER DURCHFÜHRUNG

Bei der Durchführung im Setting Schule ergaben sich einige Probleme:

1. Nicht alle Lehrerinnen haben so mitgearbeitet, wie erwünscht; an einigen Schulen fiel z.B. das Hörtraining aus, wenn die Studentinnen nicht da waren oder wurde sehr unregelmäßig durchgeführt. Auch wurden die Fragebögen nicht immer gewissenhaft ausgefüllt.
2. Die gewünschte Trainingshäufigkeit konnte oft auch deshalb nicht eingehalten werden, weil Projektstage, Projektwochen, Klassenarbeiten usw. dazwischenkamen.
3. Das Setting Schule ist problematisch, da, wenn die Kinder nicht in einem extra Raum zum Musikhören sitzen, Störungen durch Klassenkameraden und den normalen Unterrichtsverlauf häufig sind, was den Effekt zusätzlich einschränken kann.

Insgesamt war das Setting z.T. nicht geeignet und auch die Zeit zu kurz, mit zu vielen Unterbrechungen.

Die Elternfragebögen waren schwer auswertbar. Aufgrund der Rückmeldungen der Studentinnen aber auch anhand der Ergebnisse (sehr schwankend und uneinheitlich) wurde klar, dass erstens nicht immer derselbe Elternteil den Fragebogen ausfüllte und zweitens ein Teil der Eltern sprachliche Probleme hatte (Migrationshintergrund). Außerdem war die Motivation der Eltern häufig nicht besonders hoch. Der Vollständigkeit halber wurden aber auch für die Elternfragebögen Berechnungen durchgeführt, die jedoch zu uneinheitlichen Ergebnissen führten, weshalb dieser Punkt in den weiteren Berechnungen (Extremgruppe) unterblieb.

## 2.1.2 ERGEBNISSE

### *Rückmeldungen der Studentinnen und der Kinder*

Etwa die Hälfte der Studentinnen gibt an, gegen Ende des Trainings durchaus erste Erfolge gesehen zu haben. So waren die Kinder anfänglich während des Trainings unruhig und abgelenkt, was sich mit der Zeit legte; fast alle Kinder waren gegen Ende ruhig und konzentriert. Am Schluss, so wurde von vielen Studentinnen berichtet, waren viele Kinder enttäuscht, dass sie nicht weiter Musik hören durften. Es gab fast kein Kind, dem es nicht gefallen hätte und das nicht gerne weitergemacht hätte. Fast alle Kinder geben dementsprechend (im Kinderfragebogen) an, dass das Musikhören ihnen sehr gut gefallen hat und sie es sehr mochten, tlw. sogar entspannend usw. fanden.

### *Trainingseffekte*

Keine eindeutigen Hinweise auf Trainingseffekte in der „Landau-Studie“ fanden sich in den Elternfragebögen (vermutlich aus den o.g. Gründen) und dem Audiva-Aufmerksamkeitstest und –Hörscreener. Zu letzteren ist anzumerken, dass die Testbedingungen nicht immer optimal waren und die Kinder im Rahmen des Settings Schule teilweise auch das Gefühl hatten, im Sinne eines Leistungsverhaltens Tasten drücken zu müssen, selbst wenn sie nichts hören konnten.

Bei den Ergebnissen des Audiva-Testbogens sind Treatmenteffekte vor allem im Bereich „Nonsenssilben nachsprechen“ erkennbar. Nach dem Training konnten die Kinder signifikant mehr Silben korrekt nachsprechen als vor dem Training.

Bei den Lautunterscheidungstests und dem dichotischen Hörtest finden sich neben einem signifikanten Effekt beim dichotischen Hörtest deutliche Tendenzen in Richtung einer Verbesserung aufgrund des Trainings. Vermutlich werden diese Tendenzen aufgrund der bereits hohen Ausgangswerte (sehr viele richtige Antworten) nicht signifikant, da ausreichend große Veränderungen bei hohen Ausgangswerten nicht mehr möglich sind.

Die meisten Treatmenteffekte ergeben sich aus den Ergebnissen der Lehrerfragebögen. Danach sind die Kinder nach Ende des Trainings im Verlauf des Schulunterrichts weniger müde, unkonzentriert und ablenkbar und zeigen einen geringeren Leistungsabfall im Verlauf des Vormittags. Sie verwechseln Buchstaben beim Lesen weniger oft, haben weniger Probleme beim Sprechen mancher Laute und werden von Anderen beim Sprechen besser verstanden. Zudem fragen die Kinder bei Diktaten weniger häufig nach, lassen bei Diktaten weniger Wörter oder Wortteile aus und lernen besser auch nach mündlicher Textvorgabe. Sie finden Geräusche seltener zu laut, halten sich seltener die Ohren zu, überhören Geräusche weniger oft und erkennen häufiger, aus welcher Richtung Geräusche kommen.

Die Ergebnisse weisen die Effektivität des AUDIVA-Trainings in einigen Teilbereichen nach. Aufgrund der genannten teilweise ausgeprägten Einschränkungen und der teilweise bereits sehr hohen Ausgangsleistungen der getesteten Kinder ist dies als ein positives Resultat zu werden. Betrachtet man die Bedingungen, unter denen die Ergebnisse zustande kamen, ist anzunehmen, dass sie in einem anderen Setting und unter anderen Bedingungen besser ausgefallen wären. Würde das Training z.B. regelmäßig und ohne Ablenkungen, Störungen und Unterbrechungen von außen mit dem einzelnen Kind zuhause durchgeführt, könnte es sicher weitaus deutlichere positive Veränderungen bewirken.

## **3 BERECHNUNGEN MIT EXTREMGRUPPEN**

### **3.1 HERLEITUNG UND EINORDNUNG DER ZUSÄTZLICHEN BERECHNUNGEN**

Ausgangsdaten für die folgenden Berechnungen sind die ursprünglichen Ergebnisse des "Forschungsbericht: Erprobung und Evaluation des Hörtraining: Audiva in dritten Grundschulklassen, Landau 2008". Bei den Auswertungen dieser Studie hatte sich, wie bereits erwähnt, herausgestellt, dass sich in der Treatmentgruppe viele Kinder mit unproblematischen Leistungen befanden, bei denen das AUDIVA Training keine Verbesserung bewirken konnte, da es von der Konzeption her erst ab geringeren bzw. problematischen (Teil-) Leistungen überhaupt einen Effekt zeigen kann und was durch die Testungen dokumentierbar wäre. Dementsprechend ergaben sich bei den statistischen Auswertungen über alle Kinder nur geringe Effekte des AUDIVA Trainings. Die Bildung einer Extremgruppe und die Durchführung weiterer Berechnungen mit dieser Gruppe schien daher angezeigt. Der Extremgruppe wurden all diejenigen Kinder der ursprünglichen Treatmentgruppe (Gruppe 1 und 2: ursprüngliche Messzeitpunkte 1 und 2) zugeordnet, deren Bewertungen im Fragebogen für Lehrer/innen bzw. deren Testleistungen in einem Bereich lagen, der als nicht ausreichend oder sogar problematisch gelten kann. Da bereits bei einer Einzelperson die Leistungen in Teilbereichen (Subtests) stark schwanken können, wurde die Extremgruppe nicht anhand der Gesamtleistung, sondern jeweils pro Teilbereich gebildet, um so spezifische Veränderungen besser darstellen zu können. Daher liegen pro Subtest leicht unterschiedliche Extremgruppen vor.

Für die Treatmentgruppe war die Bildung der jeweiligen Extremgruppe problemlos möglich. Für ein korrektes Vorgehen hätte diese Extremgruppe nun mit einer entsprechend zusammengesetzten Kontrollgruppe verglichen werden müssen. Die Anzahl der Kinder mit problematischen Leistungen bzw. Bewertungen für die einzelnen (Sub-)Tests bzw. Fragebogenskalen war in der vorhandenen Kontrollgruppe allerdings so gering, dass sinnvolle Berechnungen nicht durchgeführt werden konnten. Die Extremgruppe der Treatmentgruppe wurde daher mit der ursprünglichen Kontrollgruppe verglichen, da ein solcher Vergleich dem völligen Verzicht auf eine Kontrollgruppe vorzuziehen ist. Auch wenn die beiden Gruppen

verschiedene Populationen widerspiegeln, kann davon ausgegangen werden, dass sich eine zeitliche Weiterentwicklung der Leistungen der Kinder sowie Testübungseffekte durch einen (signifikanten) Anstieg der Werte in der Kontrollgruppe zeigen müssten. Sollte dies der Fall sein, könnte eine vergleichbare signifikante Verbesserung in der Extremgruppe nicht auf einen Effekt des AUDIVA Trainings zurückgeführt werden, sondern genauso gut Resultat anderer Einflüsse oder Entwicklungen sein. Sollten sich in der Kontrollgruppe keine signifikanten Leistungsverbesserungen finden, in der Extremgruppe hingegen schon, kann dies als Hinweis auf einen Effekt des AUDIVA Trainings gewertet werden.

Das üblicherweise für den direkten statistischen Vergleich des Verlaufs zweier Gruppen über die Zeit hinweg anzuwendende Verfahren, die zweifaktorielle Varianzanalyse (mit Messwiederholung), konnte bei den vorliegenden Daten nicht angewendet werden, da die Grundvoraussetzungen für parametrische Verfahren (Normalverteilung, Varianzhomogenität, z.T. auch Häufigkeit und Intervallskalenniveau) nicht erfüllt waren. Eine nonparametrische Variante steht für diese Art von Vergleich nicht zur Verfügung. Daher wurde in der Extrem- und Kontrollgruppe für alle vorgestellten Berechnungen von Unterschieden zwischen den beiden Messzeitpunkten auf einen einfachen nonparametrischen Mittelwertsvergleich (Wilcoxon-Test) zurückgegriffen, und die Verläufe der beiden Gruppen werden einander im Folgenden deskriptiv gegenüber gestellt.

Von den Berechnungen wurden die Ergebnisse einiger Kinder im Aufmerksamkeitstester und Hörscreener (4-6, je nach Test und Messzeitpunkt) ausgeschlossen, da diese bei den Tests entweder die Anweisungen nicht richtig umsetzen konnten oder nicht mitarbeiten wollten, was dazu führte, dass sie entweder immer oder nie eine Taste drückten, unabhängig von den vorgegebenen Reizen bzw. Aufgaben.

Zusätzlich muss angemerkt werden, dass bei einem großen Teil der Kinder von Seiten der Eltern eine Angabe des Geburtsdatums aus Datenschutzgründen nicht gewünscht wurde, so dass zum Alter keine genauen Zahlen vorliegen. Da aber alle Testungen in der dritten bzw. in einigen Fällen in der zweiten Klasse durchgeführt wurden, kann eine Altersspanne von 7-9 Jahren für die teilnehmenden Kinder angenommen werden.

### **3.2 TESTGÜTEKRITERIEN**

Bei allen eingesetzten Tests (Aufmerksamkeitstest, Hörscreener, Audiva-Test-CD) ist wegen der gleich bleibenden objektiven Testung mit technischen Geräten, stets gleicher Einführung und Anweisungen sowie der maschinellen Ergebnisregistrierung beim Aufmerksamkeitstester und Hörscreener und dem vorgegebenen Auswertungsbogen bei der Audiva-Test-CD eine hohe Durchführungsobjektivität gegeben. Aufgrund der schriftlichen Form des Fragebogens für Lehrer/innen und der schriftlich beigefügten Erläuterung und Bearbeitungshinweise kann auch bei diesem Instrument von einer guten Durchführungsobjektivität ausgegangen werden. Aus den vordefinierten Antwortformaten des Fragebogens für Lehrer/innen und der bei den objektiven Tests immer nach festen Regeln vorgenommenen Umrechnung der Zahlen sowie der Anwendung wissenschaftlicher statistischer Verfahren bei der Auswertung aller Daten resultiert eine ebenfalls als hoch zu bewertende Auswertungs- und Interpretationsobjektivität.

Die Tests und der Fragebogen wurden unter Einbezug bisheriger Erfahrungen aus der Praxis und gleichzeitig eng theoriegeleitet konstruiert, um eine ausreichende Inhalts- bzw. Konstruktvalidität herzustellen. Aus organisatorischen Gründen war im Rahmen der Studie eine externe Validierung (Kriteriumsvalidität) nicht möglich. Da die einzelnen Items des Aufmerksamkeitstests und des Hörscreeners stets von gleicher Art sind, ist die Reliabilität dieser Tests, berechnet als Cronbachs Alpha, hoch (siehe Punkt 4.1 und 4.2). Bei den Subtests der Audiva-CD liegt die Reliabilität der Subskalen im niedrigen bis mittleren Bereich (Punkt 4.3). Die Reliabilität der Summenskalen des Fragebogens für Lehrer/innen kann für alle Skalen als gut bezeichnet werden (siehe Punkt 4.4).

### **3.3 AUSGESCHLOSSENE TESTS BZW. FRAGEN**

In der folgenden Auswertung wurde, im Gegensatz zur „Landau-Studie“ der Fragebogen für Eltern aus den in Punkt 2.2.1 genannten Gründen nicht mit einbezogen. Bei den Aufgaben der Audiva-Test-CD wurde das Nachsprechen von Nonsenssilben (Mottier) herausgenommen, da die Durchführenden das

Abbruchkriterium unterschiedlich gehandhabt hatten, was die Ergebnisse möglicherweise verfälscht haben könnte. Das Nacherzählen der Geschichte wurde wegen unterschiedlicher Handhabung der Fehlernotierung aus denselben Gründen in der Auswertung nicht berücksichtigt.



## **4 ERGEBNISSE**

### **4.1 ERGEBNISSE DES AUDIVA-AUFMERKSAMKEITSTESTERS**

Beim Aufmerksamkeitsstester wurden die Berechnungen basierend auf der Anzahl der Fehler durchgeführt. Zwei Arten von Fehlern waren möglich: die Taste nicht zu drücken, obwohl ein hierfür relevanter Reiz vorgegeben worden war, den Reiz also zu ignorieren oder die Taste zu drücken, ohne dass ein relevanter Reiz vorgegeben war, also ohne Ziel zu reagieren. Da beim Aufmerksamkeitsstester 100 Silben einmal ohne Ablenkung und einmal mit (akustischer und visueller) Ablenkung vorgegeben wurden, ergaben sich daraus die in Tabelle 1 dargestellten vier „Fehlerarten“. Pro Fehlerart und Kind wurde eine Summenskala (Addierung aller Fehler derselben Art) gebildet.

Die Extremgruppe setzt sich für jede Fehlerart aus denjenigen Kindern zusammen, bei denen sich mehr als vier Fehler fanden.

Die über Cronbachs Alpha berechnete Reliabilität des Aufmerksamkeitsstesters liegt für beide Versionen (ohne und mit Ablenkung) bei ,936.

Mit dem Gruppenmittelwert der Fehlersummen (= Durchschnitt für alle Kinder) wurde pro Fehlerart für die Extremgruppe und die Kontrollgruppe der statistische Test zum Vergleich der Mittelwerte zu den beiden Messzeitpunkten gerechnet (Wilcoxon-Test). Die Testergebnisse finden sich ebenfalls in Tabelle 1. Die Abbildungen 1 und 2 veranschaulichen die Summenmittelwerte.

Sowohl beim Test ohne Ablenkung als auch beim Test mit Ablenkung hat sich in der Extremgruppe die durchschnittliche Anzahl der ignorierten relevanten Reize vom ersten zum zweiten Messzeitpunkt statistisch sehr bzw. hoch signifikant reduziert.

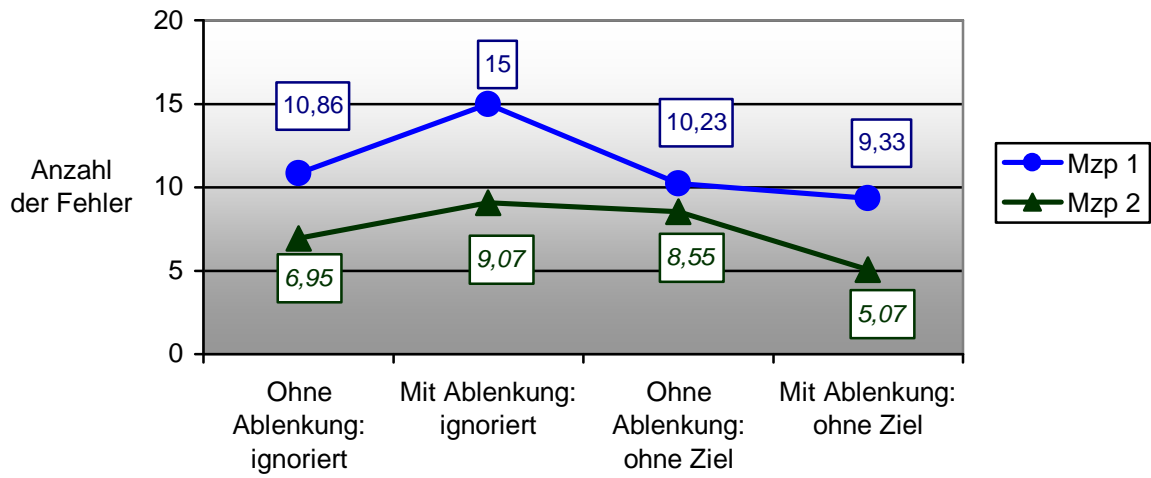
Das Drücken der Taste, ohne dass ein relevanter Reiz vorgegeben wurde, wurde zwar auch seltener, jedoch nicht ausreichend für eine statistische Signifikanz. In der Kontrollgruppe sind keinerlei signifikante Veränderungen zwischen den beiden Messzeitpunkten erkennbar.

Diese Ergebnisse können trotz der methodischen Einschränkungen als Hinweis auf die Effektivität des AUDIVA Trainings für die Extremgruppe gelten.

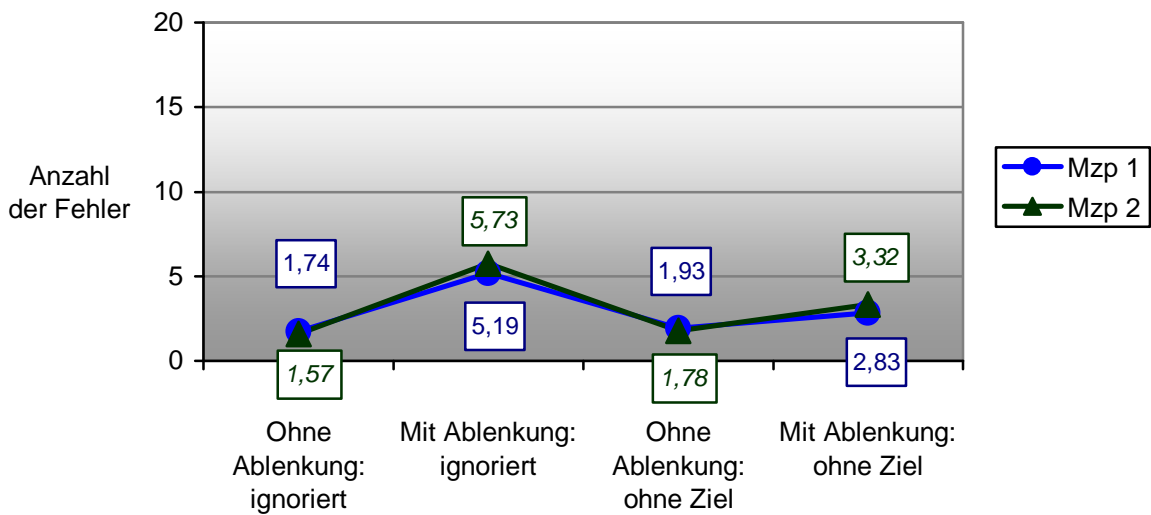
**Tabelle 1: Ergebnisse des Audiva-Aufmerksamkeitstesters ohne und mit Ablenkung für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten.**

Fehlerart	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
Test ohne Ablenkung: Relevanten Reiz ignoriert.	EG	MW	10,83	6,95	-2,59	,010
		SD	6,45	8,32		
		N	23	21		
	KG	MW	1,74	1,57	-,71	,480
		SD	3,36	1,89		
		N	42	37		
Test ohne Ablenkung: Reaktion ohne Ziel.	EG	MW	10,23	8,55	-1,07	,287
		SD	5,80	8,05		
		N	13	11		
	KG	MW	1,93	1,78	-,63	,529
		SD	2,68	2,16		
		N	42	37		
Test mit Ablenkung: Relevanten Reiz ignoriert.	EG	MW	15,00	9,07	3,36	,001
		SD	9,34	8,91		
		N	49	41		
	KG	MW	5,19	5,73	-,73	,464
		SD	4,87	8,50		
		N	42	37		
Test mit Ablenkung: Reaktion ohne Ziel.	EG	MW	9,33	5,07	-1,94	.052
		SD	7,79	4,20		
		N	18	15		
	KG	MW	2,83	3,32	-,26	,792
		SD	3,54	6,50		
		N	42	37		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)  
 Summenskala: bester Wert = 0 (kein Fehler), schlechtester Wert = 25 (maximal mögliche Fehler bei relevanten Reiz ignoriert), bzw. 75 (maximal mögliche Fehler bei Reaktion ohne Ziel)



**Abbildung 1: Audiva-Aufmerksamkeitstester Extremgruppe.**



**Abbildung 2: Audiva-Aufmerksamkeitstester Kontrollgruppe.**

## 4.2 ERGEBNISSE DES AUDIVA-HÖRSCREENERS

Beim Hörscreener wurden getrennt für das rechte und linke Ohr Sinustöne in 5 verschiedenen Frequenzen vorgegeben. Nach Einweisung des Kindes wurde der automatische Testablauf gestartet, bei dem der Versuchsleiter nicht mehr eingriff. Begonnen wurde jede Frequenz bei 0dB (durchschnittliche erwartete Hörschwelle), dann in +5dB Schritten bis +30dB gesteigert. Erfolgte kein Tastendruck, begann die nächste Frequenz entsprechend. Wurde die Taste gedrückt, verringerte sich die Lautstärke solange (bis minimal -12dB), bis kein Tastendruck mehr erfolgte (ein Tastendruck bei -12dB führte zum Wechsel auf die nächste Frequenz). Dann erhöhte sich die Lautstärke erneut und erwartete die 2. Antwort, nach der zur nächsten Frequenz gewechselt wurde.

Folgende dB Schritte standen zur Verfügung: -12, -8, -4, 0, 5, 10, 15, 20, 25, 30dB. Damit bei den Berechnungen auch diejenigen Kinder mit einbezogen werden konnten, die die Taste überhaupt nicht gedrückt hatten, erhielt das „Nichtdrücken“ den höchsten (fiktiven) Wert (35 dB).

Auch für den Hörscreener wurde Cronbachs Alpha berechnet: Einbezogen wurden alle fünf Frequenzen beider Ohren. Dabei ergab sich ein Wert von ,897.

Aus den Ergebnissen des Hörscreeners ergeben sich pro Ohr und Frequenz zusätzlich zur Kontrollgruppe zwei weitere Gruppen:

Die Extremgruppe, der all diejenigen Kinder zugeordnet wurden, die zum ersten Messzeitpunkt den entsprechenden Ton erst bei 15 dB oder einer höheren dB-Zahl bzw. gar nicht gehört hatten sowie die Gruppe der Hochsensiblen, die zum ersten Messzeitpunkt den Ton noch bei einer der sehr niedrigen dB-Zahlen (-8 oder -12 dB) gehört hatten. Gemäß dem Ziel des AUDIVA Trainings sollten die dB-Zahlen nach dem Training in der Extremgruppe sinken, in der Gruppe der Hochsensiblen hingegen ansteigen.

Wie bei den anderen Tests auch, wurde die Extremgruppe der Kontrollgruppe gegenüber gestellt, wobei die Hochsensiblen nicht in die Berechnungen mit einbezogen wurden. Für die getrennt ausgewertete Gruppe der Hochsensiblen war die Bildung einer Kontrollgruppe nicht möglich, da die bereits in der Extremgruppe sehr geringe Anzahl der Hochsensiblen in der Kontrollgruppe deutlich zu niedrig war

(i.d.R. 1-3 Personen) und die Kontrollgruppe per se sich zu einem Vergleich nicht eignet.

Zu den Ergebnissen ist zusätzlich anzumerken, dass eine Verbesserung entweder ein Hinweis auf besseres Hören oder ein Hinweis auf bessere „Hör“-Aufmerksamkeit sein kann.

#### 4.2.1 VERGLEICH VON EXTREMGRUPPE UND KONTROLLGRUPPE

Bei allen fünf Frequenzen pro Ohr, insgesamt zehn, findet sich in der Extremgruppe eine signifikante, sehr signifikante oder hoch signifikante Verbesserung zwischen dem ersten und zweiten Messzeitpunkt (alle Ergebnisse siehe Tabelle 2 und 3, Mittelwertsverläufe Abbildung 3 – 6). In der Kontrollgruppe lässt sich eine signifikante Verbesserung bei zwei von zehn Frequenzen erkennen.

Ein Testübungseffekt kann daher nicht ausgeschlossen werden. Die deutliche Verbesserung in der Extremgruppe weist jedoch auf einen Effekt des AUDIVA Trainings hin.

**Tabelle 2: Ergebnisse des Audiva-Hörscreeners für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten – Teil 1.**

Frequenz	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
1000 Hertz links	EG	MW	27,31	9,50	-2,53	,012
		SD	10,13	10,12		
		N	13	10		
	KG	MW	10,97	11,15	-,20	,840
		SD	12,06	13,88		
		N	36	26		
1000 Hertz rechts	EG	MW	25,88	11,67	-2,66	,008
		SD	9,23	11,13		
		N	17	15		
	KG	MW	10,95	8,39	-1,20	,230
		SD	11,42	10,63		
		N	37	28		
4000 Hertz links	EG	MW	22,21	13,57	-3,58	,000
		SD	8,72	8,48		
		N	34	28		
	KG	MW	12,32	11,21	-1,22	,223
		SD	11,52	10,76		
		N	41	33		
4000 Hertz rechts	EG	MW	22,66	11,33	-3,67	,000
		SD	8,71	10,98		
		N	32	30		
	KG	MW	12,21	9,83	-2,00	,046
		SD	10,54	9,68		
		N	43	29		
3000 Hertz links	EG	MW	22,73	11,79	-3,88	,000
		SD	8,21	8,52		
		N	33	28		
	KG	MW	13,69	12,64	-,51	,608
		SD	10,77	11,68		
		N	42	36		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Mittelwert in dB, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)

**Tabelle 3: Ergebnisse des Audiva-Hörscreeners für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten – Teil 2.**

Frequenz	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
3000 Hertz rechts	EG	MW	21,91	13,10	-3,17	,002
		SD	8,44	10,30		
		N	34	29		
	KG	MW	14,40	12,17	-1,41	,159
		SD	10,94	10,88		
		N	42	30		
2000 Hertz links	EG	MW	20,29	13,28	-2,65	,008
		SD	7,38	9,38		
		N	34	29		
	KG	MW	13,93	9,14	-2,11	,035
		SD	10,91	8,70		
		N	42	35		
2000 Hertz rechts	EG	MW	21,33	13,08	-3,12	,002
		SD	7,65	10,96		
		N	30	26		
	KG	MW	11,71	10,86	-,37	,711
		SD	11,38	10,32		
		N	41	35		
6000 Hertz links	EG	MW	20,43	12,32	-2,67	,008
		SD	7,31	9,76		
		N	35	28		
	KG	MW	13,21	10,00	-1,09	,277
		SD	10,58	9,01		
		N	42	33		
6000 Hertz rechts	EG	MW	21,96	10,29	-2,64	,008
		SD	8,49	10,82		
		N	23	17		
	KG	MW	12,44	9,31	-1,38	,169
		SD	11,17	10,67		
		N	39	29		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Mittelwert in dB, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)

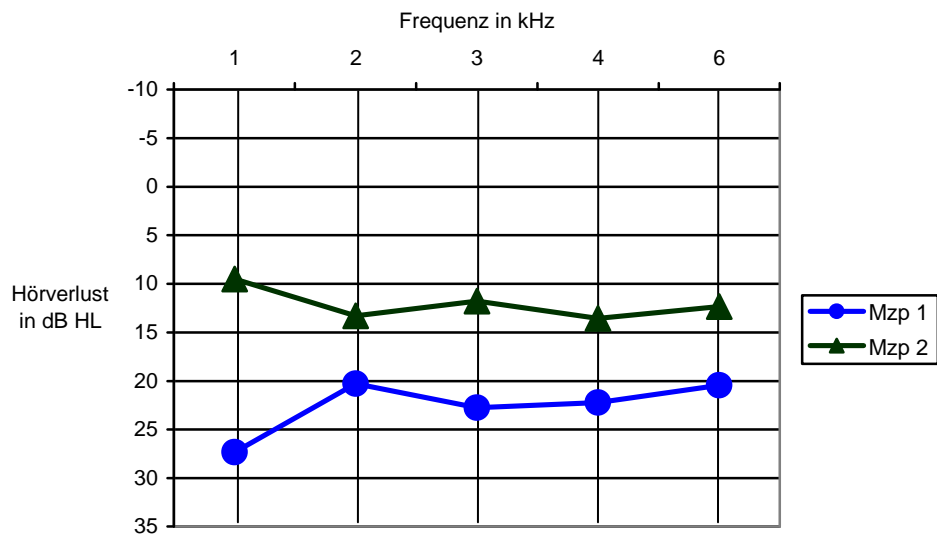


Abbildung 3: Hörscreener: Extremgruppe linkes Ohr.

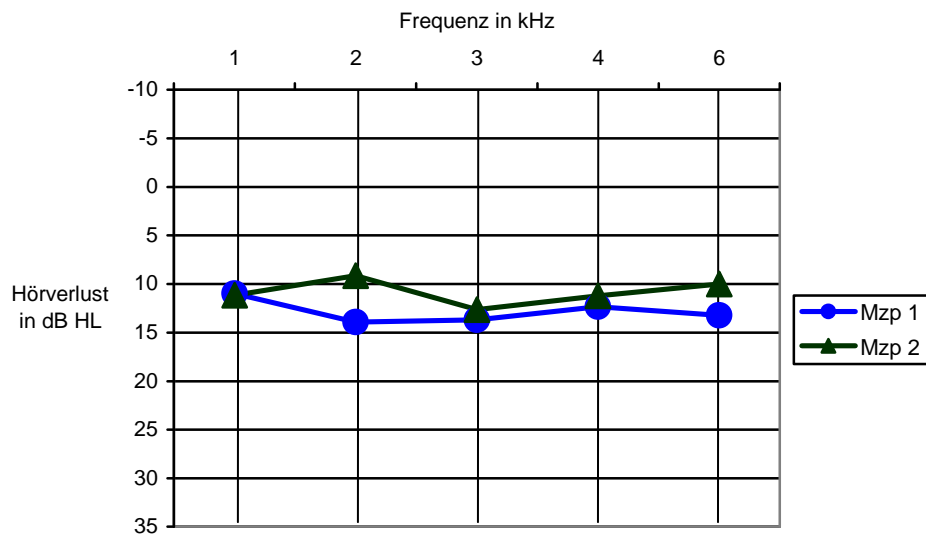


Abbildung 4: Hörscreener Kontrollgruppe linkes Ohr.



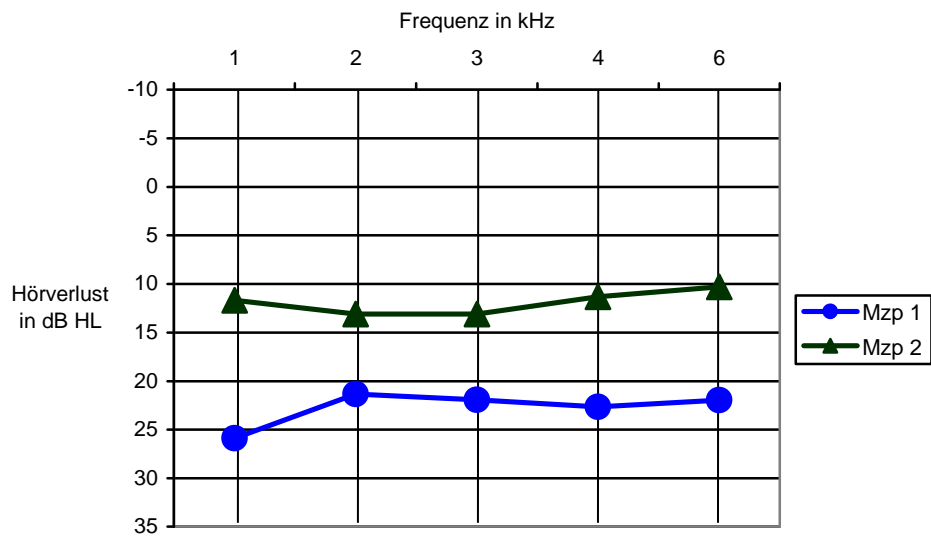


Abbildung 5: Hörscreener: Extremgruppe rechtes Ohr.

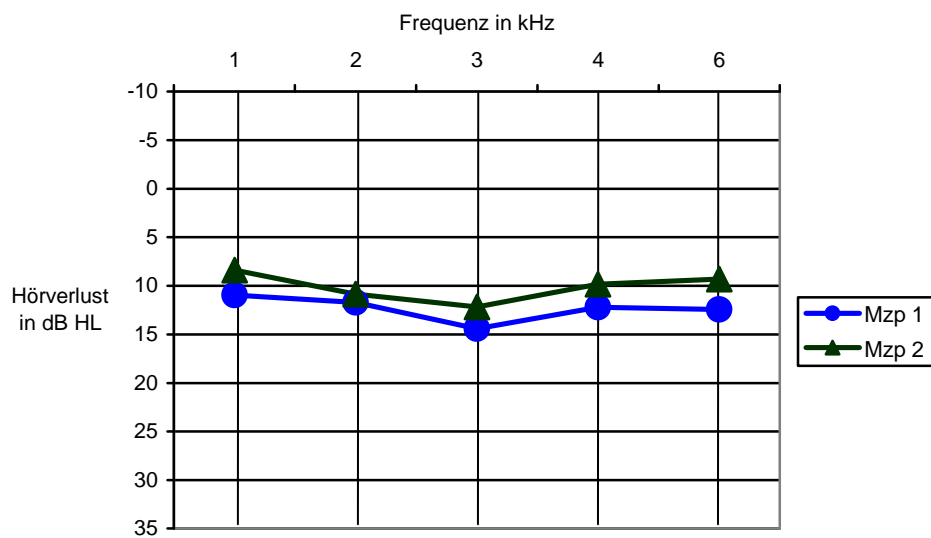


Abbildung 6: Hörscreener: Kontrollgruppe rechtes Ohr.

#### 4.2.2 DIE GRUPPE DER „HOCHSENSIBLEN“

In der Gruppe der Hochsensiblen lässt sich bei all denjenigen Frequenzen, in denen die Anzahl der Kinder hoch genug für einen Mittelwertsvergleich der beiden Messzeitpunkte war, eine signifikante Verbesserung dahingehend erkennen, dass die durchschnittliche dB-Zahl ansteigt (alle Ergebnisse siehe Tabelle 4, Mittelwertsverläufe siehe Abbildung 7 und 8). Auch wenn die Anzahl der Kinder hier teilweise sehr gering ist, könnte dies doch, im Vergleich mit der Extremgruppe, der Vermutung bezüglich des Effekts des AUDIVA Trainings entsprechen und ein Hinweis auf dessen „ausgleichende“ Wirkung (Annäherung an 0 dB) sein.

**Tabelle 4: Ergebnisse des Audiva-Hörscreeners für die Gruppe der Hochsensiblen zu den beiden Messzeitpunkten.**

Frequenz		Mzp 1	Mzp 2	Zeitvergleich	
				Z	p
1000 Hertz links	MW	-9,33	11,20	-	-
	SD	-	-		
	N	6	4		
1000 Hertz rechts	MW	-9,20	-3,89	-2,21	,027
	SD	1,93	4,70		
	N	10	9		
4000 Hertz links	MW	-10,00	5,57	-2,20	,028
	SD	2,14	12,70		
	N	8	7		
4000 Hertz rechts	MW	-10,00	12,00	-	-
	SD	-	-		
	N	4	3		
3000 Hertz links	MW	-10,00	5,00	-	-
	SD	-	-		
	N	2	1		
3000 Hertz rechts	MW	-9,71	9,33	-2,21	,027
	SD	2,14	13,74		
	N	7	6		
2000 Hertz links	MW	-10,67	5,00	-	-
	SD	-	-		
	N	3	2		
2000 Hertz rechts	MW	-10,00	-1,00	-	-
	SD	-	-		
	N	4	3		
6000 Hertz links	MW	-9,33	2,00	-	-
	SD	-	-		
	N	3	3		
6000 Hertz rechts	MW	-11,00	4,00	-2,38	,017
	SD	1,85	8,52		
	N	8	7		

MW = Mittelwert in dB, SD = Standardabweichung,  
 N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)

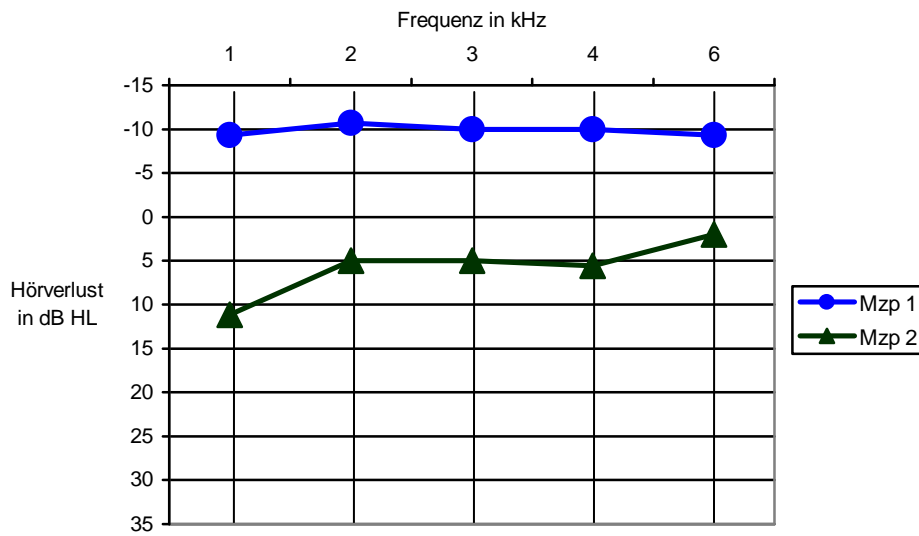


Abbildung 7: Hörscreener: Hochsensible linkes Ohr.

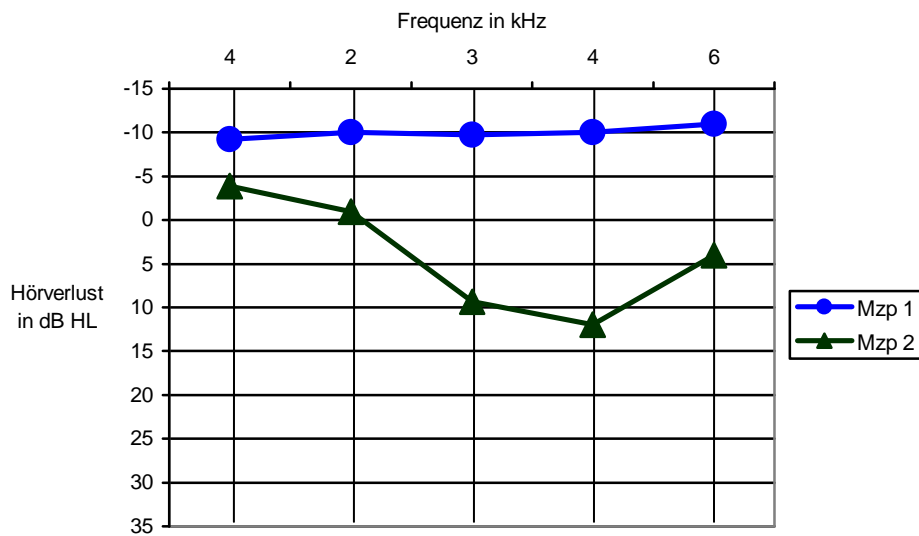


Abbildung 8: Hörscreener: Hochsensible rechtes Ohr.

### **4.3 ERGEBNISSE DER AUDIVA-TEST-CD**

#### **4.3.1 ERGEBNISSE DES LAUTUNTERSCHIEDUNGSTESTS STUFE 2**

Der Lautunterscheidungstest der Stufe 2 wurde einmal ohne und einmal mit Störgeräuschen durchgeführt. Pro Ohr wurden jeweils 16 Silben präsentiert. Die richtigen Antworten der Kinder wurden addiert. Insgesamt ergaben sich daraus pro Kind vier Summenskalen mit einem maximalen Wert von 16. Die Extremgruppen der vier Skalen setzen sich aus Kindern zusammen, die zum ersten Messzeitpunkt jeweils weniger als 12 richtige Antworten gaben.

Cronbachs Alpha für die Skala „Stufe 2 ohne Störgeräusch links“ beträgt ,632, für die Skala „Stufe 2 ohne Störgeräusch rechts“ ,636 Die Skala „Stufe 2 mit Störgeräusch links“ hat ein Cronbachs Alpha von ,642 die Skala „Stufe 2 mit Störgeräusch rechts“ ein Cronbachs Alpha von ,544.

In der Extrem- und in der Kontrollgruppe wurde pro Skala der Gruppenmittelwert zum ersten Messzeitpunkt mit dem Gruppenmittelwert zum zweiten Messzeitpunkt verglichen (Wilcoxon-Test). Alle Ergebnisse finden sich in Tabelle 5, die Verläufe der Mittelwerte in den Abbildungen 9 und 10. In der Extremgruppe erhöhte sich die durchschnittliche Anzahl richtiger Antworten bei allen vier Summenskalen statistisch signifikant bis hoch signifikant, in der Kontrollgruppe sind keine signifikanten Veränderungen erkennbar. Die deutlichen Verbesserungen in der Extremgruppe im Kontrast zu den gleichbleibenden Ergebnissen der Kontrollgruppe sprechen für einen Effekt des AUDIVA Trainings in der Extremgruppe.

**Tabelle 5: Ergebnisse des Lautunterscheidungstests Stufe 2 für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten – Teil 1.**

Skala	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
Stufe 2 ohne Störgeräusch links	EG	MW	9,50	12,33	-2,79	,005
		SD	1,37	2,02		
		N	16	15		
	KG	MW	13,47	13,65	-,47	,637
		SD	1,90	1,90		
		N	38	37		
Stufe 2 ohne Störgeräusch rechts	EG	MW	9,61	11,57	-2,38	,017
		SD	1,83	3,87		
		N	23	21		
	KG	MW	13,46	13,44	-,23	,818
		SD	2,25	1,93		
		N	39	38		
Stufe 2 mit Störgeräusch links	EG	MW	9,16	11,00	-3,19	,001
		SD	1,69	3,24		
		N	37	35		
	KG	MW	12,11	12,22	,00	1,000
		SD	2,55	2,74		
		N	37	37		
Stufe 2 mit Störgeräusch rechts	EG	MW	9,07	10,77	-3,85	,000
		SD	1,94	2,56		
		N	60	52		
	KG	MW	10,62	11,03	-,57	,566
		SD	2,66	2,34		
		N	39	37		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)  
Summenskala: 0 (keine Antwort richtig) – 16 (alle Antworten richtig)

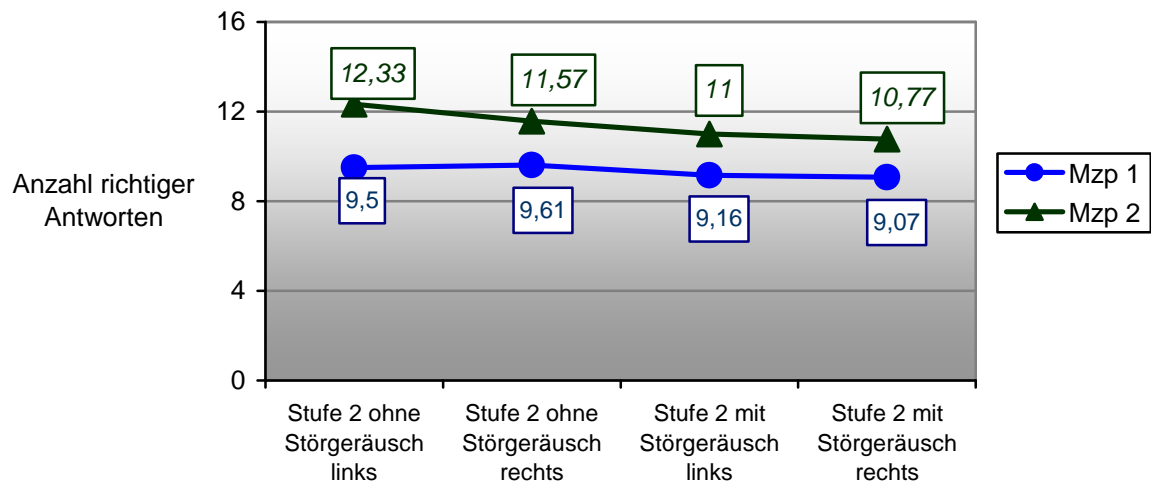


Abbildung 9: Lautunterscheidungstest Stufe 2: Extremgruppe.

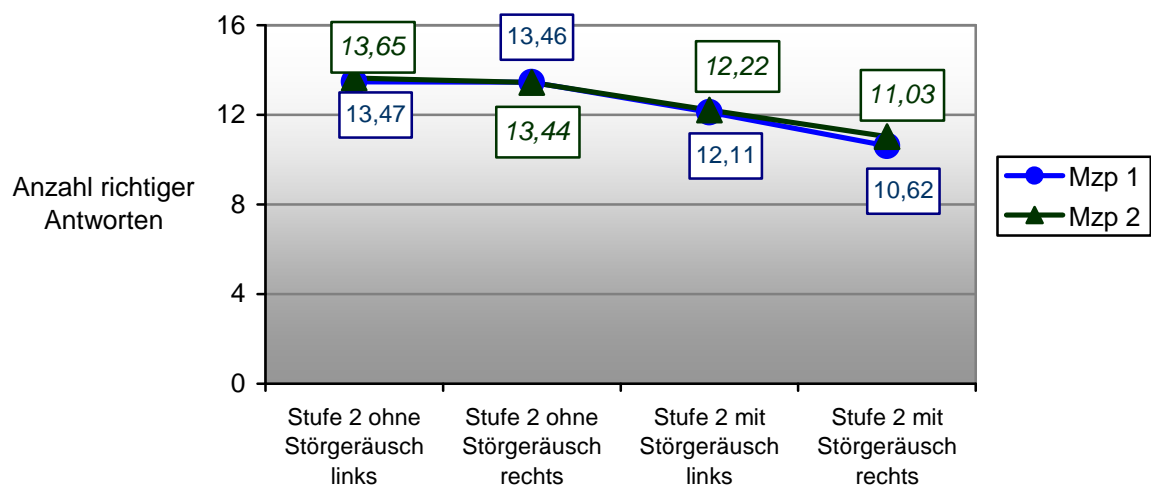


Abbildung 10: Lautunterscheidungstest Stufe 2: Kontrollgruppe.

### 4.3.2 ERGEBNISSE DES LAUTUNTERSCHIEDUNGSTESTS STUFE 3

Im Lautunterscheidungstest der Stufe 3 mit Störgeräusch wurden pro Ohr 16 Silben, immer 2 nacheinander als Pärchen auf das linke/rechte Ohr vorgegeben, so dass achtmal zwei Silben nachgesprochen wurden. Die richtigen Antworten wurden addiert. Pro Kind ergaben sich daraus vier Summenskalen mit einem Maximum von acht richtigen Antworten. Diese Unterteilung (einzelne Skala für jede Silbe) wurde vorgenommen, damit mögliche Effekte der Reihenfolge, in der die Silben dargeboten werden, nicht in einem Gesamtmittelwert „untergehen“. Der Extremgruppe wurden diejenigen Kinder zugeordnet, von denen zum ersten Messzeitpunkt weniger als 6 richtige Antworten vorlagen.

Für die Skala „Stufe 3 links (1)“ ergab sich ein Cronbachs Alpha von ,347, für die Skala „Stufe 3 rechts (1)“ ein Cronbachs Alpha von ,536. Cronbachs Alpha für die Skala „Stufe 3 links (2)“ beträgt ,397, für die Skala Stufe 3 rechts (2) ,359. Trotz der niedrigen Werte wurde eine Summenskala gebildet, da eine Zusammenfassung wegen der Zugehörigkeit zur selben Aufgabe möglich und zudem die Auswertung von Einzelsilben wenig sinnvoll ist.

Für jede Gruppe wurde pro Summenskala der Gruppenmittelwert des ersten Messzeitpunktes mit dem des zweiten Messzeitpunktes verglichen (Wilcoxon-Test). Die durchschnittliche Anzahl richtiger Antworten nahm in der Extremgruppe für alle vier Skalen sehr signifikant bzw. hoch signifikant zu. In der Kontrollgruppe zeigt sich bei einer Skala (Stufe 3 links – 2) eine sehr signifikante Zunahme richtiger Antworten; bei den restlichen drei Skalen ergab sich keine signifikante Veränderung (alle Ergebnisse siehe Tabelle 6, Mittelwertsverläufe siehe Abbildung 11 und 12).

Diese Ergebnisse liefern ebenfalls einen Hinweis auf die Effektivität des AUDIVA Trainings. Testübungseffekte können allerdings wegen der signifikanten Zunahme bei einer Skala der Kontrollgruppe nicht völlig ausgeschlossen werden.



**Tabelle 6: Ergebnisse des Lautunterscheidungstests Stufe 3 mit Störgeräusch für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten.**

Skala	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
Stufe 3 links (1)	EG	MW	4,30	5,53	-4,15	,000
		SD	,91	1,71		
		N	47	43		
	KG	MW	5,87	5,84	-,15	,882
		SD	1,32	1,42		
		N	38	37		
Stufe 3 rechts (1)	EG	MW	3,95	5,25	-3,65	,000
		SD	1,24	1,93		
		N	44	40		
	KG	MW	5,56	5,79	-1,49	,137
		SD	1,52	1,28		
		N	39	38		
Stufe 3 links (2)	EG	MW	4,09	5,04	-3,30	,001
		SD	1,01	1,76		
		N	53	48		
	KG	MW	4,68	5,50	-2,58	,010
		SD	1,76	1,42		
		N	38	36		
Stufe 3 rechts (2)	EG	MW	4,22	5,17	-3,10	,002
		SD	,95	1,78		
		N	45	42		
	KG	MW	5,67	5,63	-,18	,860
		SD	1,46	1,55		
		N	39	38		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)  
Summenskala: 0 (keine Antwort richtig) – 8 (alle Antworten richtig)

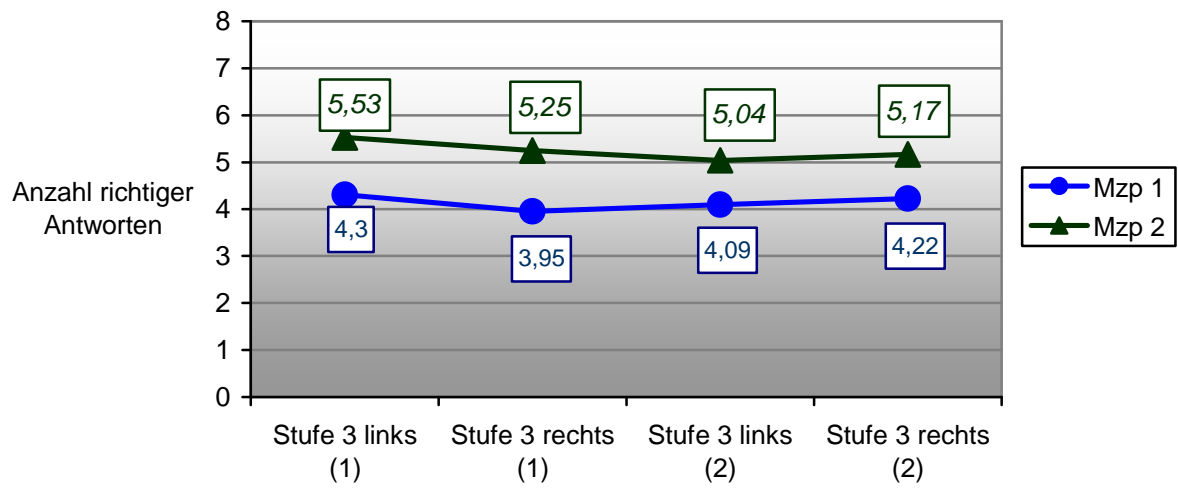


Abbildung 11: Lautunterscheidungstest Stufe 3 mit Störgeräusch: Extremgruppe.

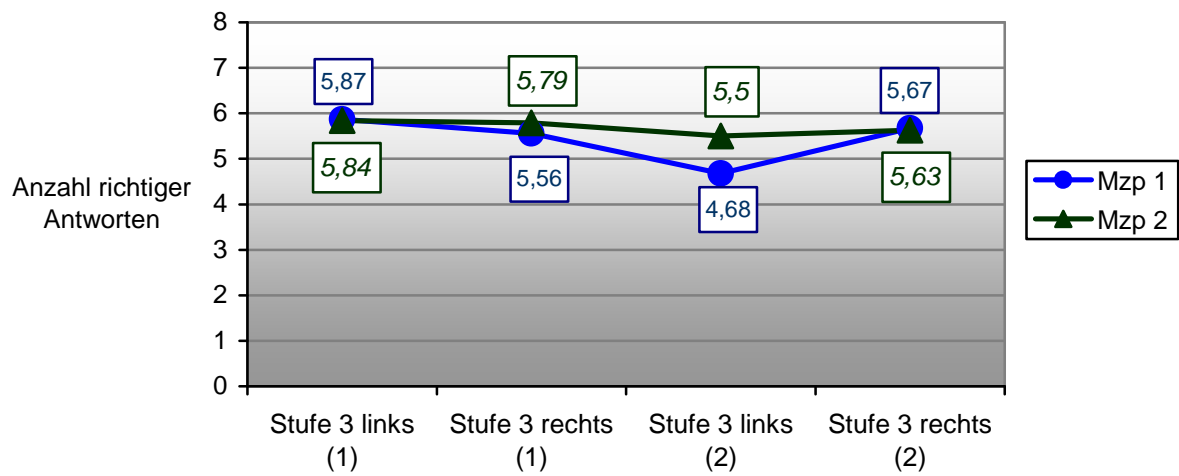


Abbildung 12: Lautunterscheidungstest Stufe 3 mit Störgeräusch: Kontrollgruppe.

#### 4.3.4 ERGEBNISSE DES DICHOTISCHEN HÖRTESTS

Beim dichotischen Hörtest hörten die Kinder gleichzeitig ein aus zwei sinnvollen Wortteilen (Einzelwörter) zusammengesetztes zweisilbiges Wort auf dem linken Ohr und ein anderes, ebenfalls aus zwei sinnvollen Wortteile bestehendes Wort auf dem rechten Ohr. Dies wurde 25 Mal mit verschiedenen Wörtern durchgeführt. Die Anzahl richtiger Antworten pro Ohr (links/rechts) und Wortteil (Wortteil 1/ Wortteil 2) wurde für jedes Kind addiert, woraus sich vier Summenskalen mit einem Maximalwert von jeweils 25 richtigen Antworten ergaben. Auch hier wurden, analog zum Vorgehen beim Lautunterscheidungstest der Stufe 3 mit Störgeräusch, unabhängige Skalen für jeden Wortteil gebildet, um mögliche Effekte der Darbietungs-Reihenfolge zu berücksichtigen. Kinder mit weniger als 16 richtigen Antworten (pro Ohrseite und Wortteil) zum ersten Messzeitpunkt bilden die Extremgruppe der jeweiligen Summenskala.

Für die Summenskala „1. Wortteil links“ wurde ein Cronbachs Alpha von ,799 berechnet, für die Skala „2. Wortteil links“ ein Cronbachs Alpha von ,763, für die Skala „1. Wortteil rechts“ ergab sich ein Wert von ,821 und für die Skala „2. Wortteil rechts“ ein Wert von ,840.

Die Gruppenmittelwerte der vier Skalen für den ersten Messzeitpunkt wurden für Extrem- und Kontrollgruppe mit den Gruppenmittelwerten für den zweiten Messzeitpunkt verglichen. Alle Ergebnisse finden sich in Tabelle 7, die Mittelwertsverläufe in den Abbildungen 13 und 14.

Für die Extremgruppe zeigen die Testergebnisse bei allen vier Summenskalen eine statistisch mindestens sehr signifikante Zunahme der durchschnittlichen Anzahl richtiger Antworten, für die Kontrollgruppe fand sich bei einer Skala (1. Wortteil links) eine signifikante Zunahme, bei den drei übrigen Skalen nicht.

Auch hier kann, analog zu Punkt 4.2, ein Testübungseffekt nicht gänzlich ausgeschlossen werden. Die Ergebnisse sprechen aber dennoch für eine Wirkung des AUDIVA Trainings in der Extremgruppe.

**Tabelle 7: Ergebnisse des dichotischen Hörtests für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten.**

Skala	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
1. Wortteil links	EG	MW	10,42	16,44	-5,38	,000
		SD	3,55	4,87		
		N	45	39		
	KG	MW	17,00	18,54	-2,15	,032
		SD	4,24	4,622		
		N	38	37		
2. Wortteil links	EG	MW	10,45	16,83	-4,99	,000
		SD	3,15	4,72		
		N	40	35		
	KG	MW	17,68	18,78	-1,19	,233
		SD	3,96	5,25		
		N	38	36		
1. Wortteil rechts	EG	MW	11,67	17,59	-3,21	,001
		SD	3,38	6,86		
		N	18	17		
	KG	MW	19,82	20,54	-1,11	,267
		SD	4,24	4,09		
		N	39	37		
2. Wortteil rechts	EG	MW	11,87	17,71	-2,58	,010
		SD	3,27	7,03		
		N	15	14		
	KG	MW	20,33	20,68	-,92	,360
		SD	4,26	4,33		
		N	39	37		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)  
Summenskala: 0 (keine Antwort richtig) – 25 (alle Antworten richtig)

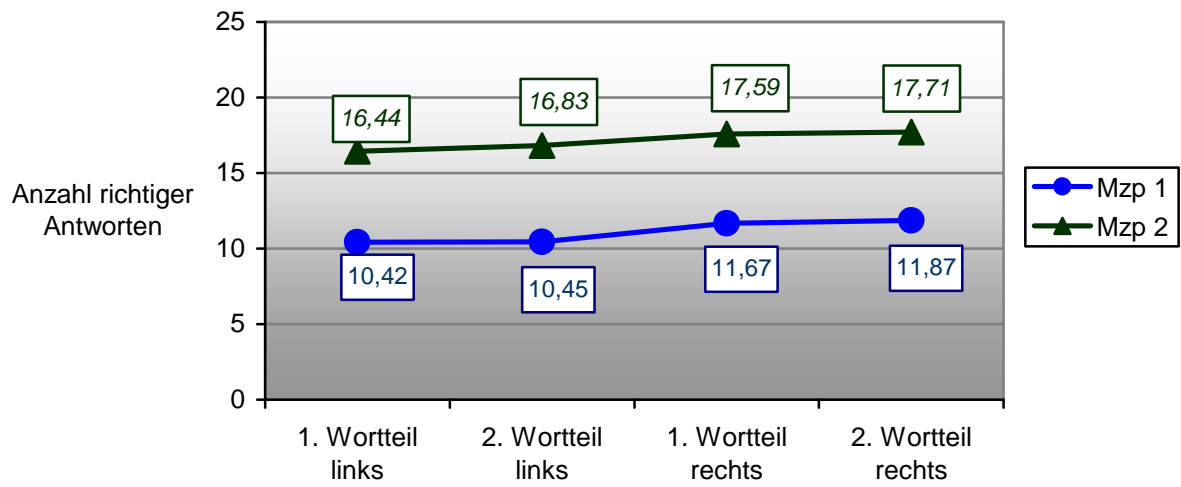


Abbildung 13: Dichotischer Hörtest: Extremgruppe.

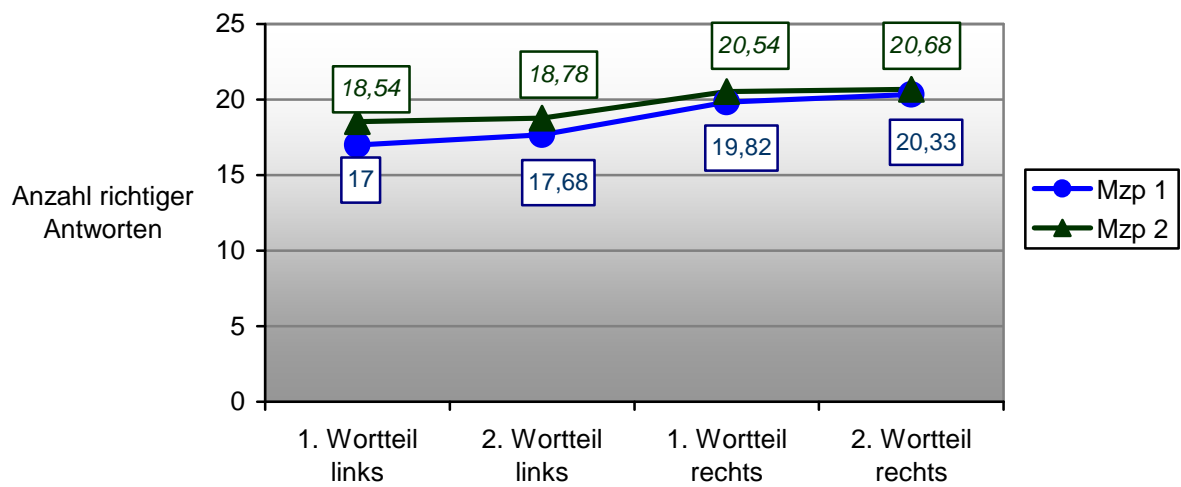


Abbildung 14: Dichotischer Hörtest: Kontrollgruppe.

## 4.4 ERGEBNISSE DES FRAGEBOGENS FÜR LEHRER/INNEN

### 4.4.1 VORGEHEN

Der eingesetzte vierseitige Fragebogen für Lehrer/innen enthielt Einzelitems zum Verhalten eines Kindes für sieben verschiedenen, potenziell problematischen Bereiche, in denen ein möglicher Einfluss des AUDIVA Trainings vermutet wird. Vorgegeben wurde ein vierstufiges Antwortformat (nie, manchmal, oft, immer). Die Einzelitems jedes Bereichs wurden zu einer Summenskala zusammengefasst, mit jeweils von der Anzahl der Items abhängigem Maximalwert. Die Reliabilität der Skalen wurde nach dem Vorliegen aller Daten mittels Cronbachs Alpha überprüft. Nur drei Items mussten entfernt werden.

Im Folgenden werden zunächst die endgültigen Skalen vorgestellt, danach die Ergebnisse der Berechnungen.

### 4.4.2 SKALEN

Für die zu Skalen zusammengefassten Einzelitems des Lehrerfragebogens gilt folgendes Antwortformat:

Immer (4) – Oft (3) – Manchmal (2) – Nie (1).

Das positive Item (Skala 4, Item 3) wurde umkodiert: Immer (1) – Oft (2) – Manchmal (3) – Nie (4).

Dementsprechend bedeuten hohe Werte Probleme in dem jeweiligen Bereich. Für die Skalen ist das Minimum (der beste Wert) jeweils 1 (nie) x die Anzahl der Items, das Maximum (der schlechteste Wert) jeweils 4 (immer) x die Anzahl der Items.

### **Skala 1: Ausdauer – Cronbachs Alpha = ,853**

1. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend motorisch unruhig.
2. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend unkonzentriert.
3. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend ablenkbar.
4. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend müde.
5. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend überfordert.
6. Das Kind zeigt einen deutlichen schulischen Leistungsabfall im Verlauf des Vormittags.

Minimum: 6 (bester Wert), Maximum: 4 x 6 = 24 (schlechtester Wert)

### **Skala 2: Lesen/Schreiben/Aussprache – Cronbachs Alpha: = ,909**

1. Das Kind verwechselt Buchstaben beim Lesen (z.B. b und p, d und t, g und k usw.).
2. Das Kind verwechselt Buchstaben beim Schreiben (z.B. b und p, d und t, g und k usw.).
3. Das Kind hat Probleme beim Sprechen mancher Laute (z.B. s und sch, m und n usw.).
4. Das Kind macht Fehler beim Schreiben nach Gehör.
5. Das Kind verwechselt klangähnliche Wörter.
6. Das Kind lässt bei Diktaten Wörter oder Wortteile aus.
7. Das Kind fragt bei Diktaten nach.

Minimum: 7 (bester Wert), Maximum: 4 x 7 = 28 (schlechtester Wert)

**Skala 3: Allgemeine Aufmerksamkeit/Vergesslichkeit – Cronbachs Alpha = ,832**

1. Wie oft vergisst das Kind etwas, das Sie ihm sagen?
2. Wenn Sie dem Kind etwas sagen, wie oft fragt es bei Ihnen noch mal nach?
3. Wie oft vergisst das Kind etwas, das seine Mitschüler/innen ihm sagen?
4. Wenn seine Mitschüler/innen dem Kind etwas sagen, wie oft fragt es bei diesen noch mal nach?
5. Wie oft scheint das Kind etwas, das Sie oder seine Mitschüler/innen ihm sagen, nicht mitzubekommen?

Minimum: 5 (bester Wert), Maximum:  $4 \times 5 = 20$  (schlechtester Wert)

**Skala 4: Auditive Aufmerksamkeit/Konzentration – Cronbachs Alpha = ,802**

1. Das Kind ist im Kopfrechnen schwächer als im schriftlichen Rechnen.
2. Das Kind lernt Sachen nach einer schriftlichen Vorlage (Text) besser als nach mündlicher Vorlage.
3. Das Kind kann gehörte Texte vollständig nacherzählen (umkodiert für Skala).
4. Das Kind verliert das Interesse, wenn lange nur gesprochen wird (z.B. beim Vorlesen einer Geschichte oder bei längerem Erklären).
5. Das Kind kommt in der Kleingruppe oder in Einzelsituationen besser mit als im großen Klassenverband.
6. Das Kind orientiert sich bei Arbeitsaufträgen an seinen Mitschüler/innen.
7. Das Kind gibt Antworten, die sich nicht auf die gestellte Frage beziehen.
8. Das Kind versteht längere mündliche Anweisungen schlechter als kurze.

Minimum: 8 (bester Wert), Maximum:  $4 \times 8 = 32$  (schlechtester Wert)



**Skala 5: Auditive Sensibilität – Cronbachs Alpha = ,871**

1. Das Kind ist leicht durch Geräusche ablenkbar.
2. Das Kind kann nicht gleichzeitig auf mehr als eine/n Sprecher/in achten.
3. Das Kind kommt bei Gruppengesprächen/Gruppenarbeit nicht mit.
4. Das Kind kommt nicht mit, wenn durcheinander gesprochen wird.
5. Das Kind findet Geräusche schnell zu laut.
6. Das Kind bekommt leise Gesprochenes nicht mit.
7. Das Kind scheint Vieles nicht richtig zu hören.
8. Das Kind hält sich die Ohren zu.
9. Das Kind kann nicht richtig zuhören.
10. Das Kind versteht Informationen falsch.
11. Das Kind erkennt nicht, aus welcher Richtung Sprache und/oder Geräusche kommen.

Minimum: 11 (bester Wert), Maximum:  $4 \times 11 = 44$  (schlechtester Wert)

**Skala 6: Rückzug – Cronbachs Alpha = ,870**

1. Das Kind wird in der Schule im Verlauf des Vormittags zunehmend in sich gekehrt.
2. Das Kind zieht sich zurück.

Minimum: 2 (bester Wert), Maximum:  $4 \times 2 = 8$  (schlechtester Wert)

**Skala 7: Sprache – Cronbachs Alpha = ,721**

1. Das Kind spricht in kurzen Sätzen
2. Das Kind spricht grammatikalisch falsch.

Minimum: 2 (bester Wert), Maximum:  $4 \times 2 = 8$  (schlechtester Wert)

#### 4.4.3 ERGEBNISSE

Bei Skala 1 wurden Kinder mit Summenwerten von 15 oder höher (erster Messzeitpunkt) zur Extremgruppe gezählt, bei Skala 2 mit Summenwerten von 17 oder höher, bei Skala 3 mit Summenwerten von 12 oder höher, bei Skala 4 mit Summenwerten von 20 oder höher, bei Skala 5 mit Summenwerten von 28 oder höher und bei Skala 6 und 7 mit Summenwerten von 5 oder höher.

Beim Vergleich der Gruppenmittelwerte von Messzeitpunkt eins und zwei mittels Wilcoxon-Test ergaben sich in der Extremgruppe für alle sieben Skalen des Fragebogens sehr bzw. hoch signifikante Verbesserungen des Verhaltens. In der Kontrollgruppe ist bei Skala 3 eine sehr signifikante Verbesserung, bei Skala 4 eine hoch signifikante Verschlechterung erkennbar. Bei den übrigen Skalen der Kontrollgruppe zeigte der Wilcoxon-Test keine statistisch bedeutsamen Veränderungen der Gruppenmittelwerte auf. (Alle Ergebnisse siehe Tabelle 8 und 9, Mittelwertsverläufe siehe Abbildung 15 – 21.)

Bei diesem Fragebogen muss berücksichtigt werden, dass die Motivation der Lehrerinnen zum Ausfüllen des Fragebogens zum zweiten und dritten Messzeitpunkt teilweise deutlich nachgelassen hat. Dies hat möglicherweise die Ergebnisse der Fragebögen mit einer ohnehin schon „subjektiven“ Einschätzung der Kinder durch die Lehrerinnen beeinflusst.

Dennoch gehen die Ergebnisse in dieselbe Richtung wie die Ergebnisse der in den vorherigen Unterkapiteln beschriebenen Tests und weisen somit ebenfalls auf eine Effektivität des AUDIVA Trainings hin.

**Tabelle 8: Ergebnisse des Lehrerfragebogens für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten – Teil 1.**

Skala	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
Skala 1: Ausdauer	EG	MW	17,58	15,42	-5,32	,000
		SD	2,38	2,98		
		N	66	66		
	KG	MW	15,85	15,67	-,65	,517
		SD	3,72	3,05		
		N	40	38		
Skala 2: Lesen, Schreiben, Aussprache	EG	MW	20,68	16,14	-4,98	,000
		SD	3,04	4,06		
		N	37	37		
	KG	MW	16,53	17,28	-1,41	,158
		SD	4,42	4,06		
		N	40	38		
Skala 3: Allgemeine Aufmerksam- keit/ Vergesslichkeit	EG	MW	13,24	11,55	-3,79	,000
		SD	1,15	2,44		
		N	33	33		
	KG	MW	11,03	9,97	-3,13	,002
		SD	2,31	2,78		
		N	40	38		
Skala 4: Auditive Aufmerksam- keit/ Konzentration	EG	MW	22,93	21,56	-3,13	,002
		SD	2,13	3,55		
		N	51	50		
	KG	MW	16,95	19,46	-5,09	,000
		SD	3,06	3,00		
		N	40	38		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)

Summenskala: Skala 1: 6 (bester Wert) – 24 (schlechtester Wert), Skala 2: 7 (bester Wert) – 28 (schlechtester Wert), Skala 3: 5 (bester Wert) – 20 (schlechtester Wert), Skala 4: 8 (bester Wert) – 32 (schlechtester Wert)

**Tabelle 9: Ergebnisse des Lehrerfragebogens für die Extrem- und Kontrollgruppe zu den beiden Messzeitpunkten – Teil 2.**

Skala	Gruppe		Mzp 1	Mzp 2	Zeitvergleich	
					Z	p
Skala 5: Auditive Sensibilität	EG	MW	30,40	26,64	-3,12	,002
		SD	2,08	4,72		
		N	25	25		
	KG	MW	24,53	23,95	-1,82	,069
		SD	5,45	4,28		
		N	40	38		
Skala 6: Rückzug	EG	MW	5,83	5,11	-3,39	,001
		SD	,85	1,24		
		N	36	36		
	KG	MW	4,53	4,54	-,05	,962
		SD	1,55	1,35		
		N	40	38		
Skala 7: Sprache	EG	MW	6,09	5,06	-3,42	,001
		SD	1,12	1,66		
		N	32	32		
	KG	MW	4,88	4,70	-1,49	,137
		SD	1,68	1,44		
		N	40	38		

EG = Extremgruppe, KG = Kontrollgruppe, MW = Summenmittelwert, SD = Standardabweichung, N = Häufigkeit, Z = Prüfgröße des Wilcoxon-Tests, p = Signifikanz (5% Niveau)  
 Summenskala: Skala 5: 11 (bester Wert) – 44 (schlechtester Wert), Skala 6 und 7: 2 (bester Wert) – 8 (schlechtester Wert)

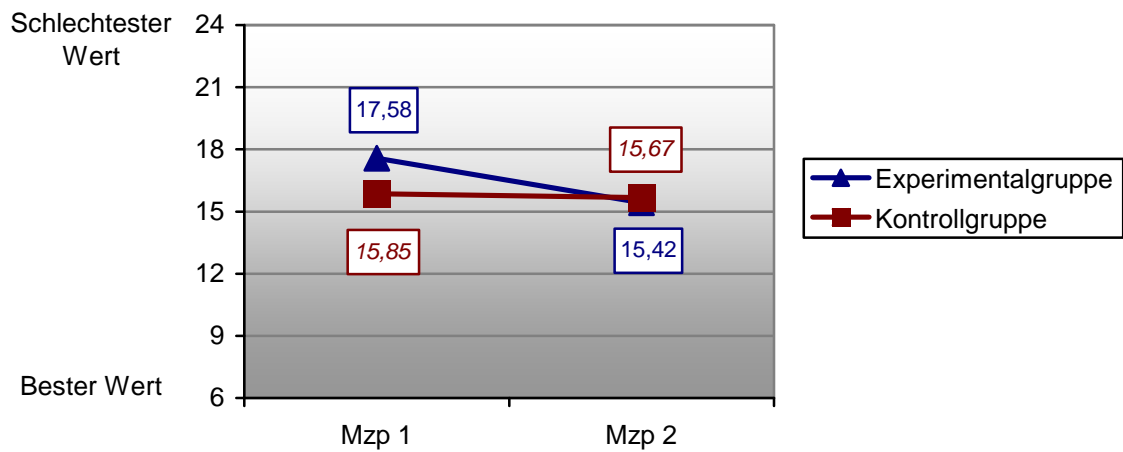


Abbildung 15: Fragebogen für Lehrer/innen Skala 1: Ausdauer – für Extrem- und Kontrollgruppe.

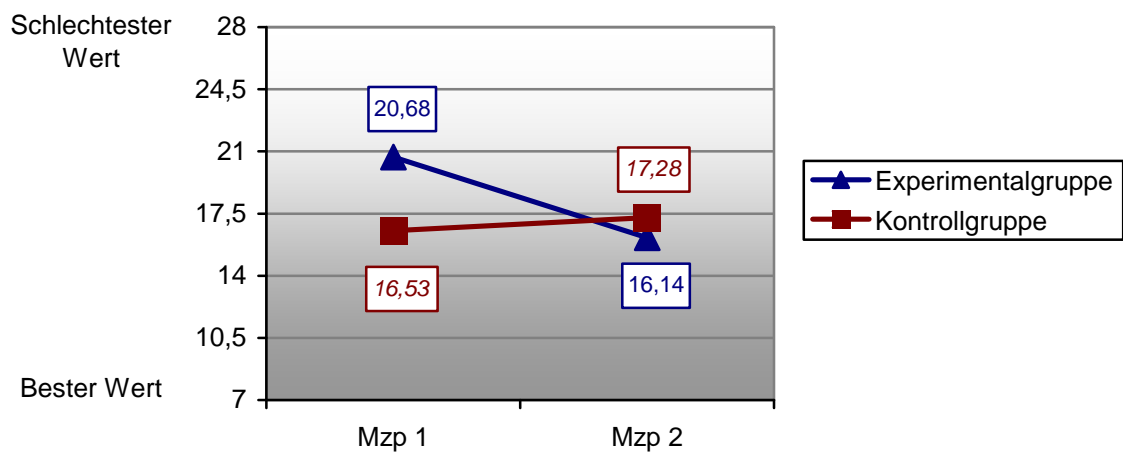
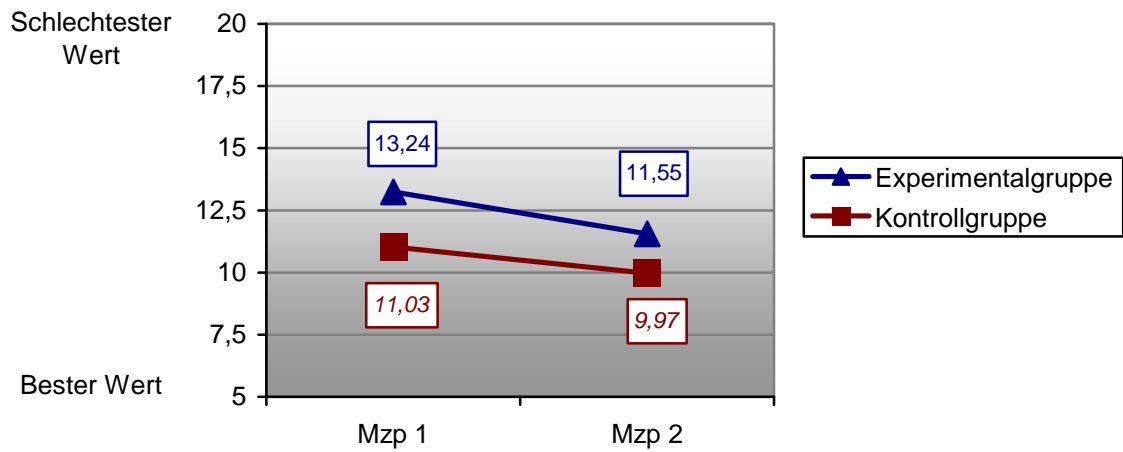
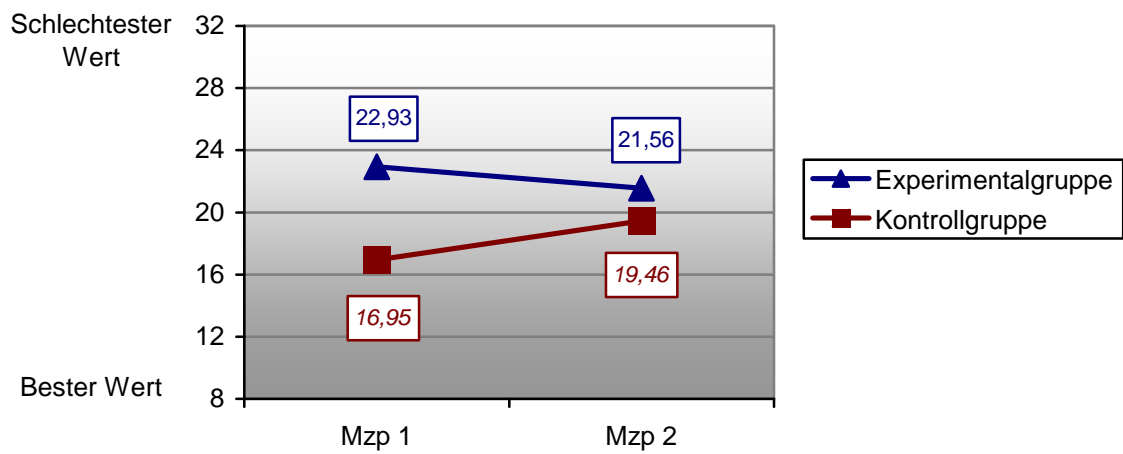


Abbildung 16: Fragebogen für Lehrer/innen Skala 2: Lesen, Schreiben, Aussprache – für Extrem- und Kontrollgruppe.



**Abbildung 17: Fragebogen für Lehrer/innen Skala 3: Allgemeine Aufmerksamkeit/Vergesslichkeit – Extrem- und Kontrollgruppe.**



**Abbildung 18: Fragebogen für Lehrer/innen Skala 4: Auditive Aufmerksamkeit/Konzentration – Extrem- und Kontrollgruppe.**

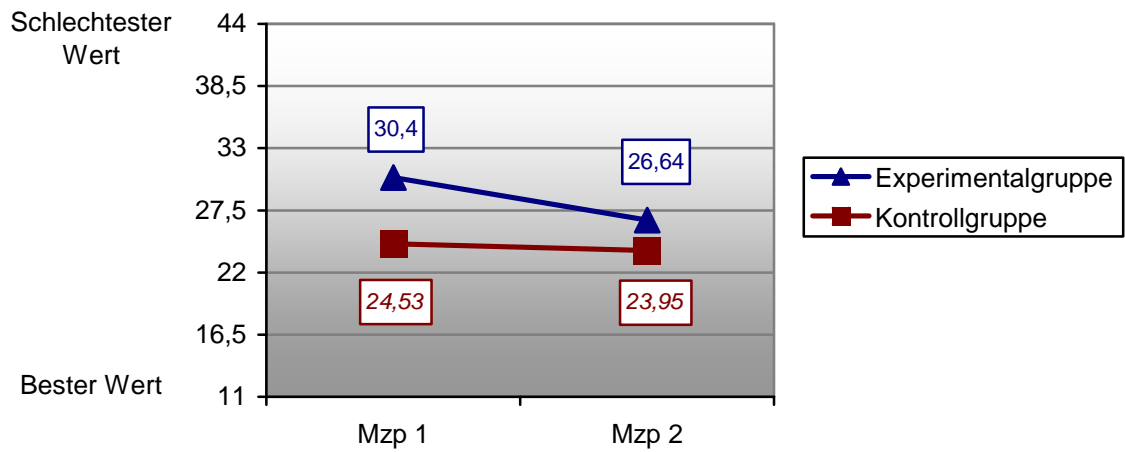


Abbildung 19: Fragebogen für Lehrer/innen Skala 5: Auditive Sensibilität – Extrem- und Kontrollgruppe.

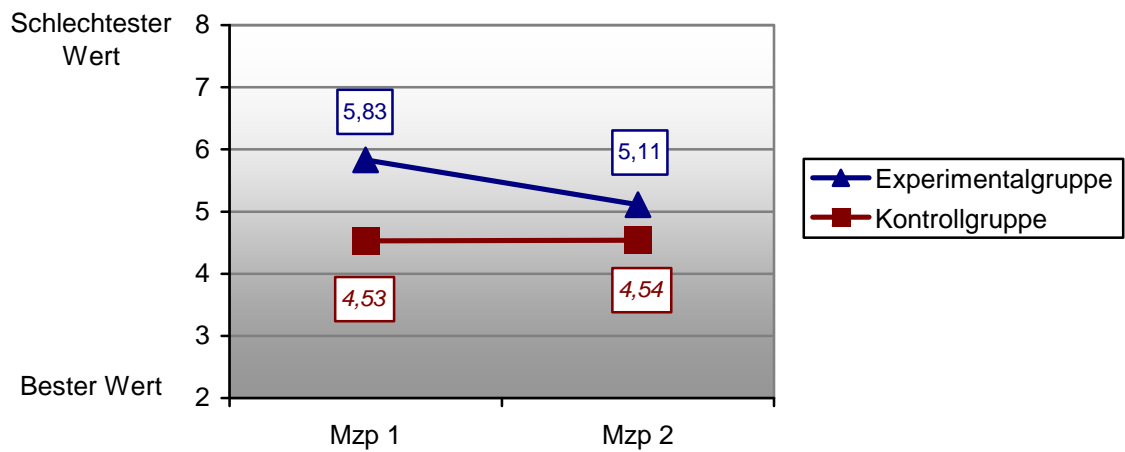


Abbildung 20: Fragebogen für Lehrer/innen Skala 6: Rückzug – Extrem- und Kontrollgruppe.

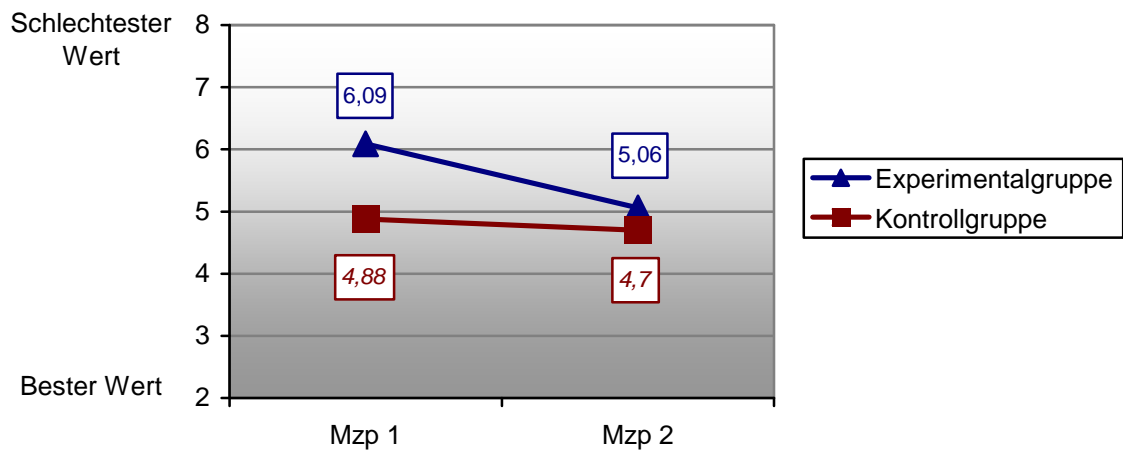


Abbildung 21: Fragebogen für Lehrer/innen Skala 7: Sprache – Extrem- und Kontrollgruppe.



#### **4.5 ERGEBNISSE DES ABSCHLUSSFRAGEBOGENS FÜR KINDER**

Im Abschlussfragebogen wurden die Kinder nach ihrer Bewertung des AUDIVA Trainings (Musikhören) anhand vorgegebener Antworten befragt sowie in offenen Fragen danach, was ihnen besonders bzw. nicht so gut gefallen hatte und was sie gern anders gehabt hätten. Der Fragebogen konnte aus organisatorischen Gründen nicht an alle Kinder ausgeteilt werden (teilnehmende Kinder: 44).

Von diesen gaben 37 an, das Musikhören hätte ihnen sehr gut gefallen. Drei Kindern hatte das Musikhören gut gefallen, vier Kinder kreuzten „geht so“ an. Die schlechteste Kategorie – „gar nicht“ gefallen – wurde von keinem Kind genannt (alle Antworten siehe Tabelle 10).

**Tabelle 10: Wie hat dir das Musikhören gefallen?**

Sehr gut		Gut		Geht so		Gar nicht		Gesamt
N	%	N	%	N	%	N	%	N
37	84,1	3	6,8	4	9,1	-	-	44

## **5 DISKUSSION**

### **5.1 LIMITIERUNGEN DER STUDIE UND EMPFEHLUNGEN FÜR DAS WEITERE VORGEHEN**

Wie in der Einführung beschrieben, konnte die Extremgruppe nicht mit einer entsprechenden Kontrollgruppe verglichen werden. Diese Ungleichheit der Gruppen bedeutet eine Einschränkung der Gültigkeit der Ergebnisse, da die unterschiedlichen Verläufe nicht mit 100%iger Sicherheit auf einen Treatmenteffekt zurückgeführt werden können. Zudem waren die Grundvoraussetzungen für die zweifaktorielle Varianzanalyse mit Messwiederholung nicht erfüllt, so dass kein direkter Vergleich der Verläufe der beiden Gruppen vorgenommen konnte. Hier sollte eine weitere Studie mit parallelisierten Gruppen und ausreichend vielen Kindern durchgeführt und dabei zusätzlich darauf geachtet werden, nur Kinder auszuwählen, die zur Extremgruppe zu zählen sind. Die Auswahl sollte möglichst durch geschulte Therapeuten anhand der Ergebnisse der Kinder in den Tests und Fragebögen erfolgen.

Des Weiteren sollte sichergestellt sein, dass das AUDIVA Training regelmäßig, d.h. möglichst täglich, durchgeführt wird und dass Lehrer/innen bzw. Eltern engagiert dabei sind. Hier waren in der Studie teilweise deutliche Einschränkungen vorhanden, die sich möglicherweise auf die Ergebnisse ausgewirkt haben.

In Zukunft sollte die anhand der Reliabilitätsberechnungen erstellte leicht veränderte Version des Fragebogens für Lehrer/innen (einschließlich mehrerer noch zu konstruierender neuer Items mit Schwerpunkt der „direkteren“ Abfrage von Konzentration und Hörfähigkeit) eingesetzt werden. Bei der Audiva-Test-CD sollten die Lautunterscheidungstests mit Cronbachs Alpha  $< ,600$  verbessert werden, z.B. indem der Schwierigkeitsgrad erhöht oder die Lautstärke bzw. Art des Störgeräusches verändert wird. Zudem wäre es interessant, mit einer größeren Gruppe von Kindern zusätzlich zum Audiva-Aufmerksamkeitstester und – Hörscreeener einen geeigneten „fremden“ Aufmerksamkeitstest und Hörtest durchzuführen, um eine externe Validierung vornehmen zu können.

In der vorliegenden Studie fehlte eine Katamnese, d.h. eine Nachtestung im Abstand von mehreren Monaten, mit der die Dauerhaftigkeit von Effekten überprüft werden

kann. In weiteren Studien sollte mindestens eine solche Nachtestung durchgeführt werden.

## **5.2 FAZIT**

Die Ergebnisse weisen auf positive Effekte des AUDIVA Trainings für die Extremgruppe in den verschiedenen Bereichen hin. Hierfür spricht vor allem, dass sich die durchschnittlichen Leistungen in der Kontrollgruppe vom ersten zum zweiten Messzeitpunkt fast nie signifikant verbesserten, in der Extremgruppe hingegen fast immer eine mindestens signifikante, oft auch sehr oder hoch signifikante Verbesserung der durchschnittlichen Leistung sichtbar wird.

Nach den Antworten der Kinder im Abschlussfragebogen der Gruppe 2 zum Messzeitpunkt 2, also am Ende des Musik-Hörtrainings ergaben sich sehr positive Bewertungen der dargebotenen Musik (meist Mozart), so dass es (zumindest aus Sicht der Kinder) in einer Folgestudie eine angenehme Fördersituation geben wird.