

Gibt es raumakustische Hörexpertise?

Markus von Berg^{1,2}, Jochen Steffens^{1,2}, Daniel Müllensiefen³, Stefan Weinzierl¹

¹TU Berlin, Fachgebiet Audiokommunikation, Einsteinufer 17c, 10587 Berlin

²Hochschule Düsseldorf, Institute of Sound and Vibration Engineering (ISAVE), Münsterstraße 156, 40476 Düsseldorf

³Goldsmiths College, University of London, New Cross, London, SE14 6NW

Email: markus.vonberg@hs-duesseldorf.de

Einleitung

Die „Expertise“ ist ein häufig in Anspruch genommenes Kriterium bei der Auswahl von Teilnehmern für Hörversuche. Hierbei wird regelmäßig auf eine professionelle Beschäftigung mit Akustik, Tontechnik oder Musik verwiesen, unter der Annahme, dass Personen aufgrund ihres Sachverständnisses zum einen mit den im Hörversuch verwendeten Begriffen und Konzepten vertraut sind und zum anderen aus ihrer professionellen Erfahrung eine erhöhte Sensitivität für die präsentierten Stimuli resultiert. Eine einheitliche Definition von Experten und Laien existiert im Kontext von Hörversuchsdurchführungen allerdings nicht [1] – anders als etwa im Bereich der Lebensmittelindustrie, in der es einen internationalen Standard zur Kategorisierung und Schulung von Teilnehmern an sensorischen Experimenten gibt (ISO 8586 [2]). Darin werden Sachverstand und sensorische Sensibilität als zwei unabhängige Kompetenzen unterschieden.

In Hörversuchen zur Raumakustik wurden verschiedene Dimensionen wie Stärke, Halligkeit oder Klarheit herausgearbeitet, um die perzeptive Wirkung von Raumklang zu erfassen [3]. Im Zuge dessen stellt sich einerseits die Frage, inwieweit Teilnehmer eines raumakustischen Hörversuchs auch kleine Unterschiede in diesen klanglichen Dimensionen erkennen können. Andererseits ist die Frage von Interesse, ob aus dem Klang eines Raums Informationen über dessen physikalischen Eigenschaften gewonnen werden können, inwiefern Hörer also nicht nur zwei Räume anhand ihrer Stärke und Halligkeit unterscheiden können, sondern daraus auch die richtigen Schlüsse im Hinblick auf die Größe der Räume und die darin verbauten Materialien ziehen können ([4],[5]). Eine hohe sensorische Sensibilität in Kombination mit der Kompetenz, die sensorische Empfindung bestimmten Eigenschaften ihrer Quelle zuschreiben zu können, wurde im Rahmen dieser Studie als *raumakustische Hörexpertise* definiert. Es sollte untersucht werden, inwiefern sich die Unterscheidung von Sachverstand und perzeptiven Erkennungsleistung aus der ISO 8586 auch auf den Audiobereich übertragen lässt und ob, wie bisher üblich, die perzeptiven Fähigkeiten auf Basis der musikalischen und fachlichen Qualifikationen abgeschätzt werden können.

Methode

Zur Beantwortung dieser Frage wurde ein Hörversuch mit 102 Personen im Alter zwischen 16 und 59 Jahren (\bar{x} =25,8 Jahre; SD=7,9; 69,6% männlich) durchgeführt. Dies erfolgte an der TU Berlin, der Hochschule

Düsseldorf, der Goldsmiths, University of London und dem Albert-Einstein-Gymnasium in Kaarst.

Der Versuch kombinierte verschiedene Hörtests mit Selbstauskünften und expliziten Fragen zu Fachwissen, Erfahrung mit Raumakustik und Musikalität. Die Hörtests sollten verschiedene Aspekte der auditorischen Verarbeitung erfassen, welche bei der Wahrnehmung und anschließender Interpretation raumakustischer Klangmerkmale als erforderlich angesehen werden. Die Selbsteinschätzung von Wissen, Erfahrung und musikalischem Hintergrund sollte ferner zeigen, inwiefern Merkmale, welche in bisheriger Forschung als Eignungskriterien für Versuchspersonen genutzt wurden [4, 6], tatsächlich mit perzeptiven Erkennungsleistungen assoziiert sind.

In den Hörtests wurden die jeweils untersuchten Fähigkeiten der Probanden nach der probabilistischen Testtheorie (*item response theory*, IRT) [7] als latente Variablen modelliert. Gemäß dieser Theorie wird diese Fähigkeit einer Person als linearer Prädiktor für die Antworten auf verschiedene Testitems geschätzt, die im Modell durch spezifische Itemparameter charakterisiert werden. Da die Ratewahrscheinlichkeit in den Itemparametern modelliert wird und die Fähigkeiten der Probanden von der Anzahl an Testdurchführungen unabhängig ist, sind keine Wiederholungen eines Items notwendig, um die Reliabilität der Ergebnisse zu gewährleisten [7]. Stattdessen bestand der Test aus einer gewissen Anzahl von im Prinzip identischen Testitems mit variablem Schwierigkeitsgrad.

Für die im folgenden dargestellten Räume wurden binaurale Raumimpulsantworten (BRIRs) mit der raumakustischen Simulations-Software RAVEN [8] simuliert. In der Simulation wurden Spiegelschallquellen erster und zweiter Ordnung berechnet und ein Raytracing mit 200 000 Partikeln durchgeführt. In den ersten beiden Hörtests wurde ein einfacher Quaderraum mit einem Volumen von 1200 m² und einer Bühne in Form eines um 1 m erhöhten, 4,5 m tiefen Plateaus verwendet, auf dem die Quelle mittig platziert wurde. Im dritten Hörtest wurden simple Quaderräume eingesetzt, in denen Quelle und Empfänger mittig platziert wurden. Alle Wandflächen wurden mit dem gleichen Material simuliert und die Absorptions- und Streugrade iterativ für die gegebenen Zielnachhallzeiten (vgl. Abbildung 1) approximiert. Da nicht für alle Schallquellen Richtcharakteristika verfügbar waren, wurde in allen Fällen eine omnidirektionale Schallabstrahlung simuliert. Der Abstand zwischen Quelle und Empfänger

betrug in den Hörtests 1+2 10 m und im Hörtest 3 5 m.

Die BRIRs wurden mit nachhallfreien Instrumentalaufnahmen gefaltet und über unterschiedliche Kopfhörer wiedergegeben (Sennheiser HD650, STAX SR-303, AKG K702, Beyerdynamic DT770 Pro), welche durch regulierte Inversion jeweils individuell entzerrt waren. Die Abhörlautstärke wurde in einem Vorversuch von drei Experten auf einen für die Kombination von Quelle und Raum plausiblen Wert eingestellt und an allen Standorten reproduziert. In allen Tests wurden die Items in zufälliger Reihenfolge präsentiert.

Der gesamte Versuch wurde über ein Web-Interface implementiert, um flexibel auf das Versuchsinterface und die Antwortdaten zugreifen zu können. Von den ersten beiden Hörtests gab es zwei Varianten, von denen jeder Versuchsperson immer nur jeweils eine präsentiert wurde, um die Dauer der Versuchsdurchführung zu begrenzen. An allen Durchführungsorten wurden insgesamt 54 Personen mit der ersten und 48 mit der zweiten Variante getestet.

Hörtest 1

Im ersten Hörtest sollte zunächst die generelle Sensibilität für raumakustische Klangmerkmale untersucht werden. Dazu wurden 21 Stimuli auf einer Skala angeordnet. Es handelte sich jeweils um das gleiche, kurze Cello-Stück (Dauer: 1,6 Sekunden) als Quellsignal und einen Raumklang, der auf jeder Stufe der Skala so moduliert wurde, dass ein bestimmtes raumakustisches Merkmal von links nach rechts gleichmäßig zunahm. Zusätzlich wurde den Versuchspersonen eine Referenz präsentiert, welche auf dieser Skala wiedergefunden werden musste. Die Abweichung der Antwort der Versuchsperson von der tatsächlichen Position der Referenz auf der Skala zeigte, welche Unterschiede im jeweiligen raumakustischen Merkmal von der Person nicht mehr wahrgenommen werden konnten.

Es existierten jeweils fünf zufällig angeordnete Items. In der ersten Variante dieses Tests wurde die Nachhallzeit auf der Skala gleichmäßig von links nach rechts frequenzunabhängig erhöht. In der zweiten Variante variierte stattdessen die Klangfarbe des Nachhalls. Dazu wurde ausgehend von einer frequenzunabhängigen Nachhallzeit das Spektrum derselben schrittweise so moduliert, dass sie zunehmend zu hohen Frequenzen abfiel und zu tiefen anstieg. Abbildung 1 zeigt eine Auswahl der jeweils generierten Nachhallzeiten.

Hörtest 2

Der zweite Hörtest sollte die Fähigkeit der Teilnehmer erfassen, etwaige klangliche Effekte des Raumes in einem Audiosignal zu erkennen. Dazu wurden drei Stimuli präsentiert, bei denen es sich jeweils um verschiedene Auszüge eines Klavierstücks (Dauer: 5–7 Sekunden) handelte. Zwei dieser Klavierpassagen wurden mit der gleichen, die dritte mit einer anderen Raumimpulsantwort gefaltet. Die Aufgabe der Versuchsperson war es nun zu erkennen, bei welchen der drei Stimuli der

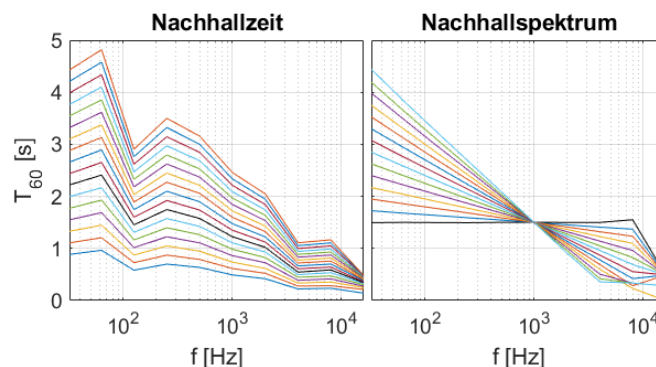


Abbildung 1: Modulation der Nachhallzeit (links) bzw. des Nachhallzeitspektrums (rechts) in den ersten beiden Hörtests. Aus Gründen der Lesbarkeit ist bei der variablen Nachhallzeit nur jede zweite und bei der spektralen Variation nur jede dritte Modulationsstufe dargestellt.

Raumklang gleich war. Dabei wurde, wie im ersten Test, entweder mit variabler Nachhallzeit oder mit variabler spektraler Verteilung des Nachhalls getestet. Es gab 20 zufällig angeordnete Testitems, bei denen ein Raum stets der gleiche blieb und der andere sich in der Modulation des jeweiligen Parameters zunehmend vom ersten unterschied, um so variable Schwierigkeitsgrade zu erzeugen. Innerhalb einer Durchführung wurde in den ersten beiden Hörtests immer der gleiche Parameter variiert.

Hörtest 3

Im dritten Test mussten drei Räume, die alle die gleiche Nachhallzeit besaßen, der Größe nach sortiert werden. Wie im ersten Hörtest wurde dreimal die gleiche Cellosequenz mit unterschiedlichem Raumklang als Stimulus präsentiert. Es gab 14 Testitems, bei denen die Unterschiede in den Raumvolumina gleichmäßig anstiegen.

Da sich die Raumgrößen nicht nur unterschieden, sondern in eine explizite Reihenfolge gebracht werden mussten, lieferte dieser Test zwei Resultate. Zunächst mussten die Versuchspersonen die Räume anhand klanglicher Merkmale in eine bestimmte Reihenfolge bringen. Wurden die Stimuli beispielsweise (wie von vielen Teilnehmern berichtet) der Lautheit nach angeordnet, musste die Versuchsperson in einem zweiten Schritt entscheiden, welcher mutmaßliche Zusammenhang zwischen Lautheit und Raumgröße besteht, also z. B. ob kleine oder größere Räume die Lautheit eines Schallsignals erhöhen. Auf diese Weise wurde einerseits die „Sortierleistung“ einer Person erfasst und andererseits die Fähigkeit, den Zusammenhang zwischen Raumgröße und klanglichen Unterschieden korrekt herstellen zu können.

Abfrage von Erfahrung und Wissen

Um die fachliche Expertise zu erfassen, sollten die Probanden zunächst die bisherige Erfahrung mit Raumakustik im Rahmen akademischer, musikalischer und aufnahmebezogener Tätigkeiten auf einer fünfstufigen Skala von ‚noch nie‘ bis ‚sehr oft‘ angeben. Zudem wurden in einem kurzen Quiz acht spezifische Fragen zu verschiedenen

Aspekten der Raumakustik gestellt. Die Fragen konzentrierten sich sowohl auf allgemeinere Themen, wie z. B. Einflussgrößen auf die Schallgeschwindigkeit, als auch auf spezielle Sachverhalte wie die statische Schalldruckverteilung von Raumresonanzen. Bei jeder Frage gab es jeweils vier Antwortoptionen, von denen nur eine richtig war.

Die Musikalität der Teilnehmer wurde mithilfe des Gold-MSI zur Erfassung der musikalischen Erfahrungheit [9] erhoben. Ausgewählt wurden die 25 Fragen, welche den Facetten *Musikalische Wahrnehmungsfähigkeiten*, *Aktiver Umgang mit Musik* und *Musikalisches Training* zuzuordnen sind.

Ergebnisse

Erfahrung mit Raumakustik

In jeder Kategorie raumakustischer Erfahrung gaben ca. 40–50 % der Probanden an, noch nie in dieser Form mit Raumakustik Kontakt gehabt zu haben. Tabelle 1 zeigt die Antwortverteilungen in allen drei Kategorien.

Tabelle 1: Selbstausskunft zur Erfahrung mit Raumakustik in verschiedenen Kontexten.

Kontakt	akademisch	musikalisch	Aufnahmen
noch nie	51,96 %	39,22 %	44,12 %
selten	13,73 %	22,55 %	16,67 %
ab und zu	12,75 %	22,55 %	25,49 %
regelmäßig	18,63 %	13,73 %	8,82 %
sehr oft	2,94 %	1,96 %	4,90 %

Hörtests

Aus den Hörtests und dem Quiz wurden nach der probabilistischen Testtheorie Fähigkeitsgrade berechnet. Im dritten Hörtest wurde dabei nur die Sortierleistung betrachtet und zusätzlich für jeden Probanden in einer binären Variable gespeichert, ob er oder sie den Zusammenhang zwischen Raumklang und -größe in mindestens vier Items korrekt herstellen konnte oder nicht. Weder für die Probandengruppe, die mit variabler Nachhallzeit getestet wurde, noch für die, die mit variabler Nachhallklangfarbe getestet wurde, ergaben sich signifikante Korrelationen zwischen den Fähigkeitsgraden aus den einzelnen Hörtests. Die Ergebnisse der Hörtests waren für eine Faktorenanalyse zur Extraktion einer zentralen Kompetenz aus den drei Tests ungeeignet (Kaiser-Meyer-Olkin-Kriterium [KMO] = 0,45 bei der Nachhallzeit und KMO = 0,44 bei der Klangfarbe).

Daher wurde für die einzelnen Fähigkeitsgrade aus den Hörtests und die korrekte Interpretationsleistung der Personen im dritten Hörtest separat getestet, inwiefern sich diese durch die Ergebnisse des Quizes, die Gold-MSI-Faktorwerte oder die Angaben zur Erfahrung vorhersagen lassen. Mit Hilfe einer *Step-Forward-Selection* wurden, ausgehend von der Variable mit der höchsten Korrelation mit dem jeweiligen Hörtestergebnis, schrittweise weitere Prädiktoren hinzugefügt, um das Akaike-Informationskriterium (AIC) des Regressionsmodells zu minimieren. Weder für die Sortierleistung noch für die korrekte Interpretation der Raumgröße im dritten Hörtest ergaben sich signifikante Prädiktionseffekte.

Tabelle 2 zeigt, welche Variablen in den übrigen Modellen enthalten waren und welche davon signifikante Prädiktionseffekte ergaben. Zudem ist das korrigierte R^2 als Maß der Effektstärke angegeben. Im Fall der variablen Nachhallzeit war im ersten Hörtest die Leistung im Quiz und im zweiten die Selbstausskunft zur akademischen Auseinandersetzung der einzige signifikante Prädiktor, wobei die Varianzaufklärung im zweiten Test mit $R^2 = 0,30$ deutlich höher lag als beim ersten Test ($R^2 = 0,17$). Beim Test mit variabler Klangfarbe ergab hingegen der Gold-MSI-Faktor *Aktiver Umgang mit Musik* bei beiden Hörtests einen signifikanten Effekt und im ersten Hörtest zudem ebenfalls das Ergebnis der Fachfragen. In diesem Fall war die Effektstärke beim ersten Hörtest mit $R^2 = 0,24$ größer als das $R^2 = 0,13$ des zweiten Tests.

Tabelle 2: Prädiktionseffekte für die ersten beiden Hörtests (HT1 bzw. HT2), getestet mit variabler Nachhallzeit (NHZ) und Nachhallklangfarbe (NHK). Als Prädiktoren wurden neben dem Quiz die Gold-MSI Faktoren *Musikalische Wahrnehmungsfähigkeiten* (MSI-W), *Aktiver Umgang mit Musik* (MSI-U) und *Musikalisches Training* (MSI-T) und die akademische (Erf-A), musikalische (Erf-M) und aufnahmebezogene Erfahrung mit Raumakustik (Erf-R) getestet. Das Symbol markiert das Vorzeichen des Regressionskoeffizienten. Signifikante Effekte sind mit (*) markiert.

Var.	Hörtestergebnis			
	HT1 NHZ	HT2 NHZ	HT1 NHK	HT2 NHK
Quiz	+	+	+	
MSI-W				
MSI-U		+	+	+
MSI-T				
Erf-A	+	+	–	
Erf-M				
Erf-R				
R^2	0,17	0,30	0,25	0,13

Diskussion

In dieser Pilotstudie zur Messung raumakustischer Hörexpertise wurde untersucht, wie die Leistung in Hörtests zur Raumakustik mit Erfahrung, Fachwissen und musikalischem Hintergrund der Person zusammenhängen. Es wurde beobachtet, dass die Erkennungsleistung bei einer Veränderung der Nachhallzeit mit dem theoretischen Fachwissen einer Person und die Erkennung von spektralen Variationen der Nachhallklangzeit vor allem mit dem aktiven Umgang mit Musik im Alltag zusammenhängen. Für die konkrete inhaltliche Interpretation eines gehörten raumakustischen Effektes, in diesem Fall die korrekte Zuordnung eines Größenunterschieds zweier Räume zu Differenzen in deren Nachhallcharakteristik, konnte kein aussagekräftiger Prädiktor gefunden werden. Eine zentrale, übergeordnete Kompetenz konnte auf Grundlage der vorliegenden Daten nicht ermittelt werden.

Dass Fachkenntnisse und akademische Erfahrung die Leistung in Hörtests verbessern, erscheint ebenso plausibel wie die Tatsache, dass eine aktive Beschäftigung mit Musik förderlich ist [6]. Bemerkenswert ist jedoch, dass das formale musikalische Training keine relevante Einflussgröße jenseits der vorhandenen Interkorrelation des entsprechenden Gold-MSI-Faktor mit dem zum aktiven Umgang mit Musik ($r = 0,34$) zu sein scheint. Die Effekte von Wissen und musikalischem Hintergrund hängen offenbar selbst dann, wenn in beiden Fällen Raumklänge verglichen werden, vom konkreten klanglichen Parameter ab. Inwiefern Fachkenntnisse und Musikalität als Eignungskriterien angenommen werden können, variiert also mit der konkreten Aufgabenstellung und den klanglichen Parametern im Versuch. Die insgesamt eher niedrigen R^2 -Indizes der Prädiktionseffekte deuten zudem darauf hin, dass die Ergebnisse der Hörtests neben den betrachteten Modellparametern von weiteren Faktoren abhängen, die sowohl personenbezogener als auch situativer Natur sein können.

Aus zeitlichen Gründen konnte jede Person in den ersten beiden Hörtests nur mit einem Parameter getestet werden. Aufgrund der limitierten Anzahl verfügbarer Probanden wurden außerdem nur zwei verschiedene Parameter eingesetzt. Es bleibt offen, wie das Verhältnis der Ergebnisse derselben Person beim Test mit Nachhallzeit und dem mit Nachhallklangfarbe ist und wie sich Fachwissen und die Ergebnisse des Gold-MSI auf Tests mit anderen raumakustischen Klangparametern auswirken. Beides sollte in nachfolgenden Studien untersucht werden.

Für die inhaltliche Interpretation klanglicher Eindrücke, die über die Sortieraufgabe des dritten Hörtests abgefragt wurde, ließen sich bei keinem dieser Merkmale signifikante Prädiktionseffekte erkennen. Nach den Versuchsdurchführungen wurde in kurzen Nachbesprechungen versucht, die Vorgehensweise der Probanden in diesem Test nachzuvollziehen. Dabei zeigte sich einerseits, dass auch fachlich ungebildete Personen, die z.B. in der Kindheit praktische Erfahrungen mit musikalischen Darbietungen gemacht haben, die Raumgrößen tendenziell korrekt dem Klang zuordnen konnten. In Zukunft sollte daher diese Erfahrung präziser und umfassender abgefragt werden, um deren Rolle bei der Interpretation von Raumklang besser abbilden und verstehen zu können. Darüber hinaus ergab das qualitative Feedback im Anschluss an die individuellen Tests, dass einige, tendenziell unerfahrene Probanden die Deutungshypothese nach den ersten Items geändert haben. Der Versuch, die konkreten physikalischen Ursachen der Klangunterschiede zu bestimmen, führte dabei zu Verunsicherung und zu Änderungen der Antworten. Auch wenn dies teils zu inkorrekten Schlussfolgerungen führte, entstanden offenbar Trainingseffekte, die ebenfalls in Zukunft näher untersucht werden sollten. Diese Trainingseffekte und die grobe Abfrage von Erfahrung könnten mögliche Gründe sein, warum sich keine signifikanten Prädiktionseffekte in diesem Test erkennen ließen.

Die Ausgangsfrage war, ob raumakustische Hörexpertise als eine zentrale Kompetenz existiert. Die Ergebnisse dieser Pilotstudie deuten darauf hin, dass es sich hierbei eher um ein komplexeres Konstrukt handelt, das sich einerseits in rein perzeptive und kognitive Leistungen und andererseits in verschiedene Parameter raumakustischer Klangwahrnehmung aufteilt. Diese können durch jeweils unterschiedliche fachliche oder musikalische Kompetenzen vorhergesagt werden.

Literatur

- [1] Zacharov, N.; Lorho, G.: What are the requirements of a listening panel for evaluating spatial audio quality? Proc. Int. Workshop on Spatial Audio and Sensory Evaluation Techniques (2004)
- [2] DIN EN ISO 8586: Sensorische Analyse - Allgemeiner Leitfaden für die Auswahl, Schulung und Überprüfung ausgewählter Prüfer und Sensoriker. Beuth-Verlag (2014)
- [3] Weinzierl, S.; Lepa, S.; Ackermann, D.: A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI). Journal of the Acoustical Society of America 144(3) (2018) S. 1245–1257
- [4] Hameed, S.; Pakarinen, J.; Valde, K.; Pulkki, V.: Psychoacoustic Cues in Room Size Perception. Audio Engineering Society Convention 116 (2004)
- [5] Shinn-Cunningham, B.: Acoustics and perception of sound in everyday environments. Proceedings of the 3rd Int Workshop on Spatial Media (2003)
- [6] Lokki T.; Pätynen, J.; Kuusinen, A.; Tervo, S.: Disentangling the preference ratings of concert hall acoustics using subjective sensory profiles. Journal of the Acoustical Society of America 132(35) (2012) S. 3148–3161
- [7] Hambleton, R. K.; van der Linden, W. J.: Advances in Item Response Theory and Applications: An Introduction. Applied Psychological Measurement 6(4) (1982) S. 373–378
- [8] Schröder D.; Vorländer, M.: RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments. Forum Acusticum (2011)
- [9] Müllensiefen, D. Gingras, B.; Musil, J.; Stewart, L.: The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. PLoS ONE 9(2) (2014)