

„Wieviel Realität braucht der Mensch?“ – Kriterien zur Bestimmung und Quantifizierung der ökologischen Validität von Hörversuchen

Jochen Steffens^{1,2}, Brigitte Schulte-Fortkamp¹, Jörg Becker-Schweitzer²

¹ TU Berlin, Institut für Strömungsmechanik und Technische Akustik, 10587 Berlin

² FH Düsseldorf, Kompetenzplattform Sound and Vibration Engineering (SAVE), 40474 Düsseldorf

Einleitung

Im industriellen Sounddesign werden vermehrt Simulationen – Abbildungen und Reduktionen einer physikalischen Realität – zur Erforschung der Geräuschqualität von Produkten eingesetzt. In der Regel sind dabei Diskrepanzen von Geräuschbewertungen im Labor zu solchen im alltäglichen Kontext zu beobachten [1,2]. Sind Laborbedingungen repräsentativ für die individuellen Alltagsbedingungen von Menschen, spricht man von ökologischer Validität [3]. Ziel dieser und zukünftiger Untersuchungen ist es, Antworten auf die Frage „Wieviel Realität braucht der Mensch?“ zu erhalten und ein geeignetes Maß zur Quantifizierung der ökologischen Validität von Hörversuchen zu finden.

Testdesign

Anhand eines Fallbeispiels werden im Folgenden relevante Kriterien bei der Geräuschbewertung beschrieben. Untersucht wurde das „akustische Verhalten“ von Wäschetrocknern und die daraus resultierenden Empfindungen in einem dreistufigen Experiment. Durch Variationen der Versuchsumgebung galt es den Einfluss des Realitätsgrades auf die Geräuschbeurteilung aufzudecken. Im Versuchsteil I wurden die Geräusche über Kopfhörer wiedergegeben. Im Versuchsteil II wurden dieselben Geräusche in einer Simulation (Mock-up-Variante) präsentiert, die als statischer visueller Stimulus fungierte, aber keine Interaktion erlaubte. Im Versuchsteil III (Realversuch) konnten die Probanden mit den echten Geräten interagieren, indem sie diese selbst mit Wäsche befüllten, die Programme einstellten und die Trockner in der Betriebsphase bewerteten.

Insgesamt wurden sechs verschiedene Wäschetrockner beurteilt. Neben der sogenannten Normwäsche, die aus kleinen weißen Baumwollhandtüchern besteht, wurden zwei Maschinen auch mit einer akustisch auffälligeren Mischwäsche, bestehend aus Hosen und Jacken mit Knöpfen und Reißverschlüssen, beladen.

Als Messmethode fand zum einen ein 7-stufiges Semantisches Differential bestehend aus den Items Angenehmheit, Kraft, Schwankung, Wertigkeit und Gewohntheit Verwendung. Zum anderen wurden die Teilnehmer im Kopfhörer- und Simulator-Test gebeten, nach der Bewertung der Stimuli ihre Eindrücke, Empfindungen und Assoziationen zu artikulieren. Dies geschah in Anlehnung an das im EU-Projekt OBELICS entwickelten AISP (Brite-Euram-Projekt No. BE-96-3727). Beim Realtest wurden die Teilnehmer aufgefordert, während des gesamten Versuchs ihre spontanen Eindrücke zu äußern. In allen Tests wurden versuchsrelevante Themen im Gespräch aufgegriffen bzw. vertieft. Die Versuche I und II haben 25 bzw. 27 Probanden im Alter von 21-73 Jahren durchlaufen. Diese beiden Personengruppen haben zu einem großen Teil an dem darauffolgenden Realversuch (40 Personen) teilgenommen. Dadurch konnten auch intraindivi-

duelle Unterschiede in der Geräuschbewertung sichtbar gemacht werden.

Ergebnisse

Die Interviews wurden mit Hilfe der sozialwissenschaftlichen Analysemethode Grounded Theory ausgewertet [4]. Die Ergebnisse für das Item „Kraft“ sind bei der Auswertung des Semantischen Differentials nicht in die Untersuchung mit eingeflossen, da eine große Zahl der Teilnehmer Schwierigkeiten mit dieser Bewertung hatte.

Der Vergleich zwischen Kopfhörer- und Simulator-Test zeigt nur geringe Unterschiede. Ein über alle Items gebildeter und gemittelter Korrelationskoeffizient erzielt einen Wert von $r = 92\%$. Ein Residuum von 8% ist bei Berücksichtigung der Heterogenität der verschiedenen Stichproben als sehr gering einzuschätzen und daher vernachlässigbar. Daher werden diese beiden Versuche für die folgenden Untersuchungen zusammengefasst.

Die qualitative und quantitative Auswertung ergibt, dass in beiden Versuchen das „Wäscheklackern“ der Mischwäsche das maßgebliche Lästigkeitskriterium darstellt. Dies stellt zugleich einen Hauptunterschied im Vergleich zum Realversuch dar. In der realen Umgebung werden die Wäschegeräusche kaum oder gar nicht mehr als störend wahrgenommen. Dafür kommen mehrere Ursachen in Betracht. Zum einen „korrigiert“ das Auge durch den sichtbaren Vorgang die akustische Wahrnehmung, gerade was den Eindruck von Gleichförmigkeit angeht. Zum anderen wird durch die erlebte Informationsintegration klar, dass die Geräusche von der Wäsche und nicht vom Gerät selbst verursacht werden. Dies sorgt für ein verstärktes Verständnis für das Geräusch und so für ein geringeres Lästigkeitsempfinden.

Validität der Versuche

Zunächst wird der Realversuch als Maßstab für die Gültigkeit der beiden Simulationsversuche herangezogen. Im Zuge einer Korrelationsanalyse wird die Kohärenz sichtbar. Der durchschnittliche Korrelationskoeffizient r beträgt 41%, was die Diskrepanz der Versuche – bedingt durch die unterschiedlichen physikalischen Realitätsgrade – deutlich macht. Werden die zwei Stimuli der Mischwäsche, bei denen die Unterschiede aus den oben beschriebenen Gründen am größten waren, nicht mit eingerechnet, erhöht sich der mittlere Korrelationskoeffizient auf 58%. Die Diskrepanzen in den Ergebnissen sind in Abbildung 1 illustriert. Die arithmetischen Mittelwerte sind hierbei mit ihrem jeweiligen Standardfehler als Balken gekennzeichnet. Auf der x-Achse sind die untersuchten Geräte aufgeführt. Unterschiede in der Bewertung des „Wäscheklackerns“ werden bei den Stimuli T2a und T4a sichtbar. Die großen Differenzen in Bezug auf die Beurteilung der Geräte T2, T3 und T6 sind mit der Optik der Geräte, die im Realtest mutmaßlich mitbewertet wurde,

in Verbindung zu bringen. Die Geräte T2 und T3 wurden in den freien Äußerungen vermehrt als positiv beschrieben, T6 hingegen hinsichtlich seiner nichtakustischen Eigenschaften tendenziell abgewertet. Diese Tatsache unterstreicht die Notwendigkeit, Produkte in ihrer Gesamtheit und ihrer Umgebung und nicht nur rein akustisch bewerten zu lassen.

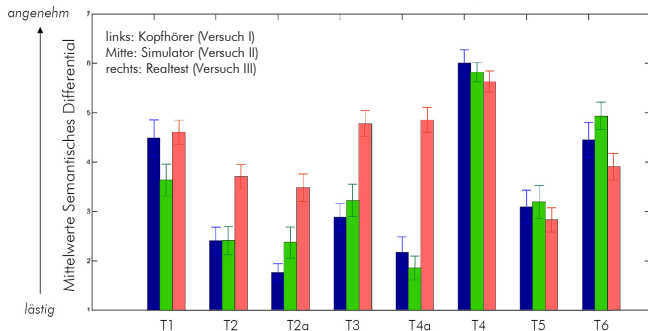


Abbildung 1: Bewertung des Items Angenehmheit

Um jedoch von ökologischer Validität sprechen zu können, müssen Alltagssituationen als Maßstab herangezogen werden. Als eine der wichtigen Alltagskategorien wurden in den Interviews Hintergrundgeräusche („Noisyness“) in der eigenen häuslichen Umgebung aufgeführt, die eine Bewertung beeinflussen. Des Weiteren werden Menschen durch die Testsituation in der Fokussierung ihrer Aufmerksamkeit gelenkt. Zudem sind Vorhersagen zu Geräuschwirkungen bezogen auf einen längeren Zeitraum im Rahmen reiner Laboruntersuchungen nahezu unmöglich.

Obgleich die beschriebenen Kategorien im Testdesign nicht berücksichtigt werden konnten, werden diese jedoch von den Teilnehmern zum Teil in den hypothetischen Kontext ihrer Beurteilung miteinbezogen. Es ist daher ein wesentliches Ziel, diese „konstruierte Realität“ bei der Geräuschbewertung zu erfassen.

„Validierender Konjunktiv“

Die Analyse der verbalen Äußerungen zeigt, dass der Gebrauch des Konjunktivs im Zuge der Geräuschbeurteilung als Indiz für das Verlassen des Versuchskontextes seitens der Probanden gewertet werden kann. Die Äußerung „das würde mich nicht stören“ unterscheidet sich in semantisch-informationeller Hinsicht wesentlich von der Aussage „das stört mich nicht“. Letztere liefert eine Information über die augenblickliche Geräuschwirkung unter den vorgegebenen Bedingungen. Durch den Gebrauch des Konjunktivs jedoch bezieht sich der Sprecher auf eine Situation, die nicht der des Versuches entspricht. Vielmehr richtet er den Fokus seiner Beurteilung auf die Alltagswahrnehmung. Aus diesem Grund wird dieses Phänomen im Rahmen der Untersuchung als „Validierender Konjunktiv“ bezeichnet.

In unserer Studie wird ebenfalls deutlich, dass die „konstruierte Realität“, die durch den Gebrauch des „Validierenden Konjunktivs“ offensichtlich wird, mit zunehmendem physikalischen Realitätsgrad steigt. Im Kopfhörertest verwendeten ihn ca. 8% der Versuchspersonen, im Simulator-test ca. 20% und im Realtest ca. 75%. Bei der Interpretation dieser Zahlen muss jedoch einschränkend bemerkt werden, dass die Gesprächsdauer im Realtest ungleich länger und damit die Wahrscheinlichkeit der Konjunktivverwendung deutlich erhöht war. Ferner ist beim Einsatz dieses „Instruments“ zu berücksichtigen, dass Äußerungen zu hypotheti-

schen Wahrnehmungen auf subjektiven Einschätzungen der Personen beruhen und nicht ohne Weiteres überprüfbar sind. Zu berücksichtigen sind auch regionale und personale Sprachgewohnheiten. Eine fehlende Verwendung des Konjunktivs als Invalidität eines Testes zu interpretieren, ist letztlich nicht zulässig. Weitere Forschungsarbeiten sind in diesem Zusammenhang bereits ins Auge gefasst.

„Realitäts“-Fragebogen

Ein mittelfristiges Ziel der Forschungsarbeiten des Autors ist die Ausarbeitung eines Fragebogens, mit dessen Hilfe die Validität von Hörversuchen allgemein quantifiziert werden kann. Derzeit stellt die explorative Evaluation den Stand der Technik dar, um die subjektive Wirklichkeit des Menschen zu ergründen [5]. Als Ergänzung hierzu kann der „Validierende Konjunktiv“ hilfreich sein, indem man bereits in der Rolle des Interviewers durch die Art der Fragestellung den Versuchskontext verlässt, z.B. durch die Frage „Würden Sie im Alltag genauso urteilen?“.

Zusammenfassung & Ausblick

In unserem Experiment wurden Wäschetrockner in einem dreiteiligen Versuch mit unterschiedlichen Realitätsgraden psychoakustisch untersucht. Der Vergleich der beiden Simulationsversuche mit dem Realtest bringt große Unterschiede in der Bewertung von Wäschegegeräuschen ans Licht. Außerdem ist der Einfluss anderer sensorischer Größen, v.a. des Aussehens eines Gerätes, auf die Geräuschbeurteilung offensichtlich. Ferner wurde ein qualitatives Validierungskonzept vorgestellt („Validierender Konjunktiv“), mithilfe dessen der Alltagskontext der Probanden bei einer Bewertung sprachlich aufgedeckt werden kann.

Zur Bestimmung der ökologischen Validität ist es notwendig, den Realitätsgrad der Versuchsumgebung weiter zu erhöhen. Unerlässlich werden in diesem Zusammenhang Feldforschungen sein, in denen die Menschen in ihrem alltäglichen Umfeld befragt werden. Außerdem sollten rein akustischen Untersuchungen Gesamtbewertungen an die Seite gestellt werden, um quantitative Aussagen zu Wechselwirkungen aller beteiligten sensorischen Einflussgrößen treffen zu können. Großes Augenmerk gilt der Verfeinerung des Konzeptes „Validierender Konjunktiv“, um diesen in der Zukunft zur potentiellen Quantifizierung der ökologischen Validität heranziehen zu können.

Literatur

- [1] Schulte-Fortkamp, B. (2000): Bewertung von Sound und Vibrationen in Fahrzeugen - ein komplexer Evaluationsprozess, DAGA 2000
- [2] Bodden, M., Heinrichs, R. (2003): Geräuschqualität im Kontext weiterer Fahrzeugattribute: Bewertung durch Kunden in Feld und Labor, DAGA 2003
- [3] Guski, R. (2001): About the validity of global noise annoyance judgments. Internoise 2001, Paper No. 305
- [4] Glaser, B., Strauss, A. (2005): Grounded Theory - Strategien qualitativer Forschung, Verlag Hans Huber, Bern
- [5] Fiebig, A., Bless, M., Schulte-Fortkamp, B., Genuit, K. (2005): Subjektive Evaluierung hat Methode, DAGA 2005