

Untersuchung der Kontextvariablen bei der Bewertung von Fahrzeuggeräuschen

Ewald Strasser¹, David Maiberger², Reinhard Weber¹, Uwe Letens² und Steven van de Par¹

¹ CvO Universität Oldenburg Institut für Akustik, 26129 Oldenburg, E-Mail: Ewald.Strasser@uni-oldenburg.de

² Daimler AG, 71063 Sindelfingen

Abstract

In dieser Studie wird exemplarisch dargestellt, welchen Einflüssen subjektive Bewertungen der Fahrgeräuschqualität unterliegen können. Subjektive Bewertungen können neben dem interessierenden Signal, der tatsächlichen Geräuschqualität, auch Störquellen enthalten. In der beschriebenen Studie wurden 6 unterschiedliche Fahrzeuggeräusche durch 84 Probanden beurteilt. Zwei der Fahrzeuggeräusche wurden durch elektronische Zuspielungen verändert. Die Daten zeigen einen eindeutigen Zusammenhang der Bewertungen mit objektiv unterschiedlichen Fahrzeuggeräuschen.

Die Daten zeigen aber auch die Abhängigkeit der Bewertungen von nicht mit der tatsächlichen Geräuschqualität interagierenden Persönlichkeitsvariablen. Zur Lösung dieses Problems wird vorgeschlagen Items zu verwenden, bei denen relevante Fahrzustände auf den Dimensionen Komfort und Sportlichkeit bewertet werden sollen. Damit wird die Validität der Itemliste zur Vorhersage der Fahrgeräuschqualität verbessert.

Einleitung

Für Autobauer galt lange Zeit die Grundregel Störgeräusche möglichst zu reduzieren, um die Gunst des Kunden zu gewinnen. In den letzten Jahren näherte sich dieser frühe Sounddesign-Ansatz allerdings seinen Grenzen an. Die einfache Formel „leiser ist besser“ ist nicht mehr uneingeschränkt gültig. Ein modernes Sounddesign muss zunehmend helfen, die Markenidentität zu prägen und dem positiven Bild des Fahrzeuges zuträglich sein [1].

Um diese Anforderungen zu bewältigen, wird unter anderem ein psychoakustischer Ansatz verwendet. Dabei werden subjektive Urteile wie beispielsweise Antworten auf semantische Differentiale in Beziehung zu objektiv operationalisierbaren Merkmalen des Geräusches gesetzt. Dieser Ansatz stößt an seine Grenzen, wenn Merkmale des Kontextes, in dem das Fahrgeräusch normalerweise gehört wird, die Wahrnehmung des Geräusches signifikant verändert [2]. Deshalb wurde vorgeschlagen, die Fahrgeräusche unter realen Bedingungen zu beurteilen [2]. Die Anwendung dieses Ansatzes brachte eine deutliche Erweiterung des Wissens über den Einfluss des Kontextes auf das subjektive Erleben des Fahrgeräusches.

Ein Aspekt, der bisher allerdings noch wenig beachtet wurde, ist der Aspekt der ökologischen Validität von so gewonnenen Ergebnissen. Ökologische Validität ist dann relevant, wenn ein erhobener Wert zur Vorhersage eines anderen Wertes verwendet werden soll. Können beispielsweise physikalische Eigenschaften eines Fahrgeräusches dazu genutzt werden, die Bewertungen

desselben Signals auf einem semantischen Differential genau vorherzusagen, so sind diese physikalischen Eigenschaften ein valides Maß für die Abschätzung der Bewertung des semantischen Differentials.

Abschätzungen der Validität wie im eben erwähnten Beispiel werden in der klassischen Psychoakustik häufig vorgenommen und sind für sich genommen unproblematisch. Der Grund für diese Abschätzungen scheint implizit allerdings oft der zu sein, über den Umweg des semantischen Differentials Aussagen über eine dritte Variable machen zu können. In Bezug auf die Fahrgeräuschakustik könnte man beispielsweise annehmen, dass Fahrgeräusche, die höhere Werte beim Item „angenehm“ erzielen, dem Kunden besser gefallen und in weiterer Folge gar annehmen, dass Fahrzeuge mit solchen Geräuschen tendenziell häufiger gekauft werden.

Obwohl eine gewisse Augenschein-Validität in diesem Beispiel nicht von der Hand zu weisen ist, darf hier aber nicht vergessen werden, dass solche Aussagen gesondert geprüft werden müssen.

Ein kurzes Beispiel soll dies illustrieren. Wird ein Fahrgeräusch als angenehm bewertet, so wird es wahrscheinlich auch so empfunden. Klarer wird das Problem, wenn relationale Aussagen gemacht werden sollen. Denn ein Fahrgeräusch mit zwei Skaleneinheiten beim Item „angenehm“ wird sehr wahrscheinlich nicht doppelt so angenehm empfunden wie ein Fahrgeräusch mit einem Skaleneinheit. Hier wird die Willkürlichkeit von solchen Bewertungen offensichtlich [3]. Diese Willkürlichkeit kann nur umgangen werden, indem konkrete Skaleneinheiten mit empirisch beobachtbarem Verhalten oder objektiven Sachverhalten in Verbindung gebracht werden.

Wenn man die subjektive Bewertung des Fahrgeräusches allgemein als Signal betrachten würde, so wäre die Aufgabenstellung dieser Studie festzustellen, wie hoch das Signal-Rausch-Verhältnis ist. Der Signalanteil wäre somit als jener Anteil in den subjektiven Bewertungen definiert, der geeignet ist objektive Unterschiede im Fahrgeräusch wiederzugeben. Rauschen wäre weiter jener Anteil der nicht durch die objektiven Unterschiede zustande kommt. Mittels einer abgestimmten Itemliste sollte die Geräuschqualität verschiedener Fahrzeuggeräusche durch naive Probanden unter realen Fahrbedingungen beurteilt werden. Die Fahrgeräusche stammten zum Teil von sehr unterschiedlichen Fahrzeugen und wurden zum Teil durch elektroakustische Modifikationen so stark verändert, dass ihnen durch Experten eine eindeutig unterschiedliche Geräuschqualität attestiert wurde. Um die weitere Diskussion zu vereinfachen, wird an dieser Stelle der Unterschied in der tatsächlichen Geräuschqualität als gegeben angesehen. Ein valides Messinstrument zur Messung der Geräuschqualität sollte somit fähig sein, die

beobachten Unterschiede zwischen den Fahrgeräuschen genau wiederzugeben.

Aus Erfahrung wissen wir allerdings, dass hier aber noch weitere Einflussfaktoren im Spiel sind (z.B.: [4]). Es kann beispielsweise sein, dass Variablen, die das Subjekt betreffen, mit in die subjektive Bewertung der Geräuschqualität eingehen. Diese können die Wahrnehmung der tatsächlichen Geräuschqualität verändern oder auch nicht. Verändern sie die Wahrnehmung nicht, so können diese Variablen gesondert gemessen und später herausgerechnet werden, um die tatsächliche Geräuschqualität zu bestimmen. Wird die Bewertung durch subjektbezogene Variablen verändert, so müssen auch in diesem Fall diese Variablen erhoben werden, damit das Ausmaß des Einflusses festgestellt werden kann. Aus dem Ausmaß des Einflusses können Rückschlüsse auf die tatsächliche Geräuschqualität gezogen werden, aber bei Vorhersagen, die das Subjekt betreffen, muss nun berücksichtigt werden, dass die subjektive Wahrnehmung nur mehr moderiert durch die subjektbezogenen Variablen verstanden werden kann. Grundsätzlich muss also in der Praxis überlegt werden, ob im jeweiligen Anwendungsfall nicht die Interaktion zwischen Subjekt und Fahrgeräusch die nützlicheren Vorhersagen treffen kann als die ideale, tatsächliche Fahrgeräuschperformance.

In dieser Studie werden wir als Beispiel für subjektbezogenen Variablen die Big 5 [5] Persönlichkeitsdimensionen heranziehen. Grund dafür ist, dass diese nachweislich für jeden Menschen verlässlich dargestellt werden können [6].

Zuletzt muss beachtet werden, dass jede Messung mit einem Restfehler behaftet ist. Dieser darf keine Korrelation mit den zu messenden Variablen aufweisen, sodass er bei ausreichend großer Datenbasis gegen Null strebt.

Die hier getroffenen Aussagen lassen sich wie folgt ausdrücken:

$$SQ_{xy} = t_{SQ_y} + SV_x + t_{SQ_y} * SV_x + Fehler \quad (1)$$

SQ_{xy} ist der von Person x gemessene Wert der Geräuschqualität für das Fahrzeug y , t_{SQ_y} ist die ideell angenommene, tatsächliche Geräuschqualität des Fahrzeugs y , SV_x sind jene subjektbezogenen Variablen von Subjekt x , die auf SQ_{xy} Einfluss nehmen und Fehler ist der auftretende Messfehler. Der Term $t_{SQ_y} * SV_x$ entspricht jenem Teil der Wahrnehmung der tatsächlichen Geräuschqualität, der durch interindividuelle Merkmale intraindividuell verschiedene Ausprägungen annimmt.

Für die Auswertung der Daten sind mehrere Ausgänge denkbar. Sollte die vorgenommene Bewertung der Geräuschqualität die tatsächliche Geräuschqualität messen, wird in unserem Beispielfall die meiste Varianz durch die unterschiedlichen Fahrzeuge mit ihren unterschiedlichen Fahrzeuggeräuschen erklärt. Würde die Bewertung hauptsächlich durch den Big 5 erklärt dann wäre sie natürlich keine gutes Maß für Geräuschqualität sondern eine gutes Maß für die Big 5 Persönlichkeitsdimensionen. Würde der Term $t_{SQ_y} * SV_x$ bzw. Fahrzeuggeräusches*Big 5 die Bewertungen erklären so müssten wir annehmen, dass (zumindest bei diesen Bewertungen) die Wahrnehmung des

Fahrzeuggeräusches wesentlich durch subjektive Variablen verändert ist. Falls keine der genannten Varianten zutrifft und nur der Fehler bzw. das Rauschen die Varianz erklären, so wäre die Bewertung nicht geeignet die Geräuschqualität abzuschätzen.

Methode

Zur Überprüfung dieser Annahmen wurden Daten von insgesamt 84 Personen herangezogen. Die entsprechenden Versuche verteilten sich auf zwei Standorte mit unterschiedlichen Testkonditionen. Am ersten Standort wurden drei verschiedene Autos (Kompaktklasse, Mittelklasse, Sportwagen) im Serienzustand genutzt. Pro Versuchsperson wurde eine Fahrt mit einem der Autos durchgeführt. Am zweiten Standort fuhr jede Versuchsperson drei Mal mit demselben Fahrzeug (Mittelklasse), allerdings wurde hier der Motorsound mittels Soundgenerator verändert. Zwischen den Fahrten wurde ein Abstand von mindestens einer Woche eingehalten, um eventuelle Übertragungseffekte zu vermeiden.

Zirka 40% der Probanden waren weiblich und 60% männlich. Das Alter der Probanden reichte von 20 bis 69 Jahren und lag im Schnitt bei zirka 43 Jahren.

Ablauf

Alle Probanden wurden derselben Testprozedur unterzogen. Nach einer Begrüßung musste eine Einwilligungserklärung zur Datenverwertung unterzeichnet werden. Die Versuchsperson wurde daraufhin über den weiteren Ablauf des Versuchs unterrichtet. Sobald der Proband sich mit dem Fahrzeug vertraut gemacht hatte, begann der erste Teil der Testfahrt.

Während der Fahrt sollten die Versuchspersonen all ihre Eindrücke zu den akustischen Phänomenen des Fahrzeuges schildern.

Auf diese erste Fahrt folgte ein Leitfadeninterview, bei dem näher auf die akustischen Phänomene eingegangen wurde. Ins Leitfadeninterview eingebettet waren auch einige quantitative Items, die die Geräuschqualität betrafen. Nach Beendigung des Interviews folgte der zweite Teil der Fahrt, bei der die Probanden wieder ihre Eindrücke über die Fahrzeuggeräusche schildern sollten.

Auch auf diese Fahrt folgten ein Leitfadeninterview und die Beurteilung einiger quantitativer Items. Zusätzlich wurden anschließend noch die „Itemliste zu Hörgewohnheiten und Erleben der akustischen Umwelt“ [7] und BFI-10 Persönlichkeitsfragebogen [8] zur Bestimmung der Big 5 vorgegeben. Abschließend folgte noch ein go/no-go Association Test [9] in dem auf implizite Statusaspekte eingegangen wurde.

Messinstrumente

Zur Messung der Big 5 Persönlichkeitsdimensionen wurde der BFI-10 [8] verwendet. Die Einteilung der Persönlichkeit des Menschen in fünf kulturell stabile Faktoren geht bis in die 30er Jahre des 20. Jahrhunderts zurück. Etabliert wurde das Modell durch [6] und gilt noch immer als State of the Art-Methode um die charakteristischen Persönlichkeitsmerkmale des Menschen zu erheben.

Zur Operationalisierung der Geräuschqualität wurden die Probanden direkt gefragt: „Welche Bewertung geben Sie dem Fahrgeräusch dieses Fahrzeugs?“. Diese Frage wurde direkt nach der ersten und direkt nach der zweiten Testfahrt gestellt, und zur Beantwortung wurde eine 10-stufige Skala mit den Endpunkten „sehr gut“ und „sehr schlecht“ vorgelegt. Zur weiteren Analyse wurde der Mittelwert beider Werte herangezogen.

Die unterschiedlichen Fahrzeuge und Fahrzeuggeräusche werden im weiteren Verlauf als Faktor „Fahrzeuge“ operationalisiert. Der Faktor „Fahrzeuge“ wird als Maß für die Validität der Bewertung herangezogen. Kann die Bewertung die unterschiedlichen Fahrzeuggeräusche gut unterscheiden, so ist die Validität gut und die Antwort der Probanden spiegelt die objektiven Unterschiede in der Fahrgeräuschqualität gut wieder. Um dies zu quantifizieren, wird betrachtet, wieviel Varianz der Faktor „Fahrzeuge“ in Bezug auf die Fahrzeuggeräuschbewertung aufklärt.

Ergebnisse

Zusammengefasst zeigen die Ergebnisse, dass ein Modell, das die Variablen *Fahrzeuggeräusche*, *Neurotizismus* und *Offenheit* beinhaltet, am besten die Bewertungen der Probanden wiedergibt. Es zeigt sich auch, dass ein großer Anteil der Messung nicht durch die verwendeten Variablen (*Fahrzeuggeräusche* und Big 5 Dimensionen) erklärt werden kann. Es folgt die statistische Begründung dieser Ergebnisse.

Zur Testung der in Gleichung (1) zusammengefassten Annahmen wurde mittels eines Linear Mixed Models in SPSS [10] ein Modell gesucht, das am besten die subjektive Bewertung der Geräuschqualität beschreibt. Zur Bewertung der Passung wurde ein Grundmodell herangezogen. Dieses Grundmodell wurde im Verlauf der Modellanpassung mit weiteren Parametern angereichert, um die Bewertung der Geräuschqualität besser wiederzugeben. Um zu bewerten, ob ein Parameter zusätzlichen Nutzen bringt, wird die von SPSS angegebene χ^2 verteilte -2 Log Likelihood herangezogen. Wenn der kritische χ^2 Wert überschritten wird, so kann der eben eingeführte Faktor als sinnvoll erachtet werden.

Die vollständige Darstellung der Modellanpassung würde den Rahmen dieses Beitrags sprengen und kann hier nur ausschnittsweise dargestellt werden.

Das Grundmodell beinhaltet lediglich die verschiedenen „Fahrzeuge“ als Faktoren. Es wurde somit untersucht wie Unterschiede in der Fahrzeuggeräuschbewertungen dadurch erklärt werden können, dass unterschiedliche Fahrzeuge mit unterschiedlichen Fahrgeräuschen bewertet wurden. Der Faktor „Fahrzeuge“ hat einen signifikanten Einfluss auf die Geräuschqualität ($F(5, 129) = 2.32, p < .05$). Der -2 Log-Likelihood-Wert (617.2) des Modells wurde als Richtschnur für potentielle Alternativmodelle herangezogen. Das Alternativmodell mit der besten Passung beinhaltet die „Fahrzeuge“ als Faktoren und die Skalen *Offenheit* und *Neurotizismus* aus dem Big 5 Fragebogen als unabhängige Kovariablen. Dieses Modell weist einen -2 Log-Likelihood-Wert von 576.8 auf, was einer Verbesserung von 40.5 bei

zusätzlich 2 Freiheitsgraden entspricht. Das $\chi^2(2) = 40.5$ überschreitet den kritischen Wert für zwei Freiheitsgrade und ist mit $p < .001$ hoch signifikant. Der Faktor „Fahrzeug“ wird mit $F(5, 122) = 3.04$ mit einem $p < .05$ signifikant. *Neurotizismus* erreicht mit $F(1, 122) = 5.28$ ebenso wie *Offenheit* mit $F(1, 122) = 4.45$ ein Signifikanzniveau von $p < .05$ (siehe Tabelle 1).

Tabelle 1: Tests auf feste Effekte

Quelle	Zähler df	Nenner df	F-Wert	p
Konstanter Term	1	122	2028,92	,000
Fahrzeuge	5	122	3,04	,013
Neurotizismus	1	122	5,28	,023
Offenheit	1	122	4,45	,037

Eine genauere Betrachtung des Faktors „Fahrzeuge“ zeigte, dass das Fahrzeug der Kompaktklasse sich mit einem $t(122) = -2.55$ signifikant ($p < .05$) von allen anderen Fahrzeugen unterscheidet (siehe Abbildung 1).

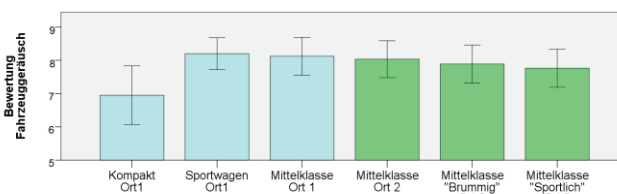


Abbildung 1: Bewertung des Fahrgeräusches der unterschiedlichen Fahrzeuge. Fehlerbalken zeigen 95% Konfidenzintervall.

Das Modell mit dem Faktor „Fahrzeuge“ und den Kovariablen *Neurotizismus* und *Offenheit* kann somit als das Modell mit der besten Anpassung an die subjektive Bewertung der Fahrzeuggeräuschqualität angesehen werden. Es muss allerdings angemerkt werden, dass ein alternatives Modell, das nur eine Interaktion aus „Fahrzeuge“ und *Offenheit* beinhaltet ebenfalls eine sehr gute Anpassung aufwies ($F(6, 122) = 2.24, p < .05$). Da der -2 -Log-Likelihood-Wert aber um 6.1 höher und bei nur 1 Freiheitsgrad unter dem letztendlich gewählten Modell lag, musste dieses Modell verworfen werden ($\chi^2(1) = 6.1, p < .05$).

Mittels einer Kovarianzanalyse wurde für das gewählte Modell eine Abschätzung der Effektstärke vorgenommen. Für das Modell mit dem Faktor *Fahrzeuge* und den Kovariaten *Neurotizismus* und *Offenheit* wurde ein korrigiertes $R^2 = .091$ errechnet. Laut Cohen [11] kann dies als kleiner Effekt bewertet werden.

Die Regressionskoeffizienten von *Neurotizismus* ($b = .25, t(116) = 2.22, p < .05$) und *Offenheit* ($b = .23, t(116) = 2.04,$

$p < .05$) sind jeweils positiv. Dies bedeutet, dass mit steigenden Werten in *Neurotizismus* und *Offenheit* auch mit höheren Bewertungen des Fahrzeuggeräusches gerechnet werden kann. Abbildung 2 und Abbildung 3 zeigen die Probanden unterteilt in Gruppen mit hohem und niedrigem Neurotizismus und Offenheitswert.

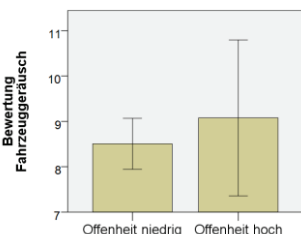


Abbildung 2: Mittelwert der Bewertung der Fahrgeräuschqualität von Probanden mit hohem und niedrigem Wert für Offenheit.

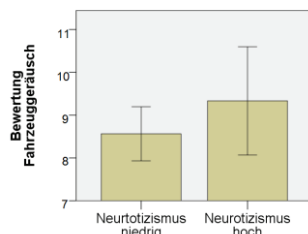


Abbildung 3: Mittelwert der Bewertung der Fahrgeräuschqualität von Probanden mit hohem und niedrigem Wert für Offenheit.

Diskussion

Aus den vorliegenden Daten lässt sich ablesen, dass sowohl der Faktor „Fahrzeuge“ als auch die Kovariaten *Offenheit* und *Neurotizismus* einen signifikanten Anteil an der subjektiven Bewertung der Geräuschqualität erklären können.

Der Faktor „Fahrzeuge“ bezieht sich auf unterschiedliche Fahrzeuge und auf sehr unterschiedlich modifizierte Fahrzeuggeräusche. Es kann daher angenommen werden, dass jener Anteil an der Varianz der Fahrzeuggeräuschbewertung, der durch den Faktor „Fahrzeug“ erklärt wird, sich auf die latente, tatsächliche Geräuschqualität bezieht.

Im Kontrast dazu zeigt der signifikante Einfluss der Big 5-Variablen *Offenheit* und *Neurotizismus* an, dass die Bewertung noch Varianzanteile enthält, die nichts mit der Geräuschqualität zu tun haben.

Aus diesen Daten könnte geschlossen werden, dass die Geräuschqualität von Fahrzeugen direkt und unverändert aus den Urteilen von Probanden abgeleitet werden kann. Somit könnten mittels Laborversuchen, bei ausreichender Kontrolle von Subjektvariablen, valide Urteile über die tatsächliche Geräuschqualität gemacht werden. Hier muss allerdings eingewandt werden, dass bereits die vorliegende Analyse eine Interaktion zwischen der Geräuschqualität und dem Big 5 Faktor *Offenheit* nicht vollkommen ausschließen konnte. Zudem ist das Ausmaß der gesamten Varianzaufklärung zu gering, um zu beurteilen, ob noch weitere Faktoren einen signifikanten Einfluss auf die subjektive Bewertung der Geräuschqualität haben.

Ausblick

Mit der vorliegenden Studie wird gezeigt, dass Augenschein-Validität nicht mit ökologischer Validität gleichgesetzt werden kann. Wenn wir Aussagen über Verhalten oder Erfahrungen von Personen vorhersagen

wollen, so muss das Vorhersageinstrument genau an den vorherzusagenden Variablen geeicht werden.

Ziel der an diese Auswertung anknüpfenden Arbeit wird es sein, ein besseres Verständnis der Bewertung der Geräuschqualität von Fahrzeugen zu entwickeln. Erste Auswertungen zeigen, dass eine Bewertung basierend auf Bewertungen der Geräuschqualität in Bezug auf einzelne Fahrzustände auf den Dimensionen Komfort und Sportlichkeit die tatsächlichen Unterschiede bezüglich Geräuschqualität besser abbilden werden als es mit der hier verwendeten Itemliste der Fall war.

Interessant könnten auch weitere Studien bezüglich der Interaktion zwischen der Geräuschqualität und dem Persönlichkeitsfaktor *Offenheit* sein. Für eine solche Interaktion wurden in dieser Arbeit Indizien gefunden. Diese deuten darauf hin, dass Personen mit unterschiedlichen Ausprägungen in Bezug auf Offenheit bestimmte Fahrzeuge / Fahrzeuggeräusche anders beurteilen.

Es deutet sich an, dass Fahrzeuge, die zu sehr von einer noch zu bestimmenden Norm abweichen, von Personen mit niedriger Ausprägung in *Offenheit* schlechte Noten für die Geräuschqualität bekommen. Sollte sich dieser Verdacht bestätigen, so wäre der Bereich, den diese Personen als „optimales Sounddesign“ wahrnehmen, viel enger als für Personen, die eine hohe Ausprägung des Faktors *Offenheit* aufweisen.

Literatur

- Genuit, I.K. Parameters Influencing the Benchmarking of Vehicle Interior Noise. in *Proceedings CFA/DAGA*. 2004.
- Blauert, J. and U. Jekosch, Sound-quality evaluation—a multi-layered problem. *Acta acustica united with acustica*, 83(5) (1997). p. 747-753.
- Blanton, H. and J. Jaccard, Arbitrary metrics in psychology. *American Psychologist*, 61(1) (2006). p. 27.
- Susini, P., et al., Instruction's effect on semantic scale ratings of interior car sounds. *Applied Acoustics*, 70(3) (2009). p. 389-403.
- Rammstedt, B., et al., Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit. *Methoden, Daten, Analysen*, 7 (2013). p. 233-249.
- Costa, P.T. and R.R. McCrae, The NEO personality inventory. (1985).
- Notbohm, G., Die Erfassung individueller Unterschiede im Erleben der akustischen Umwelt. *Lärmbekämpfung*, 3(3) (2008). p. 122-129.
- Rammstedt, B., et al., Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit. *Methoden, Daten, Analysen*, 7(2) (2013). p. 233-249.
- Nosek, B.A. and M.R. Banaji, The go/no-go association task. *Social cognition*, 19(6) (2001). p. 625-666.
- IBM SPSS Statistics for Windows. 2015; 23:
- Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, New Jersey: L. (1988).