

# Eine Leseaufgabe für Lärmwirkungsstudien im Arbeitskontext: Analysen zu Trainingseffekten und Bearbeitungsfehlern

Helga Sukowski

*Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA), Dortmund  
E-Mail: Sukowski.Helga@baua.bund.de*

## Einleitung

In einem BAuA-Forschungsprojekt wird der Einfluss der akustischen Arbeitsumgebung auf die Leseleistung und das Wohlbefinden von Beschäftigten untersucht. In einer ersten von zwei empirischen Studien stand die Untersuchung von Trainingseffekten bei zweimaliger Bearbeitung einer Leseaufgabe im Mittelpunkt. Zudem wurden die beiden Testversionen der speziell für diese Studie entwickelten Leseaufgabe genau betrachtet, indem sowohl die Schwierigkeit der einzelnen Items und der gesamten Testversionen als auch die Art der Bearbeitungsfehler ermittelt wurden.

Die Fragestellung des gesamten Forschungsprojektes ist darauf ausgerichtet, die Wirkung unterschiedlicher Geräusche auf die Leseleistung zu ermitteln, um Aussagen über den Einfluss der akustischen Arbeitsumgebung ableiten zu können. Ergeben sich konkrete Hinweise, dass bestimmte Geräusche besonders nachteilige Wirkungen haben, können diese Erkenntnisse in die Diskussion zur Optimierung der akustischen Gestaltung in Arbeitsräumen einfließen.

Um Aussagen über mögliche Effekte von Geräuschen zuverlässig ermitteln zu können, ist es grundsätzlich wichtig, bei der Erhebung von Leistungsdaten Aufgaben einzusetzen, die das erfassen, was erfasst werden soll, die aber zudem für den Einsatz in den Zielgruppen tatsächlich geeignet sind. Das bedeutet auch, dass die Aufgaben einen angemessenen Schwierigkeitsgrad haben sollten, damit sie in unterschiedlichen Gruppen von Beschäftigten eingesetzt werden können und damit die Aussagekraft der Ergebnisse nicht durch Boden- oder Deckeneffekte eingeschränkt ist. Speziell im Hinblick auf den in Lärmwirkungsstudien häufig angestrebten Vergleich der Effekte verschiedener Geräuschbedingungen (in verschiedenen Gruppen, z. B. [1, 2]; in wiederholten Messungen, z. B. [3, 4]) ist es darüber hinaus wichtig, gleich schwierige Versionen einer Aufgabe einzusetzen und den Trainingseffekt bei mehrmaligem Einsatz zu kennen.

In einer Pilotstudie waren die ersten Versionen (A und B) einer selbst entwickelten Leseaufgabe eingesetzt worden (vgl. [5, 6]). Bereits bei der Gestaltung der Aufgabe für die Pilotstudie war es ein wesentliches Anliegen gewesen, die Erkenntnisse aus bisherigen Lärmwirkungsstudien zu berücksichtigen, um eine Aufgabe zu schaffen, die hinreichend lärmsensitiv ist. Zudem erfolgte die Konstruktion vor dem Hintergrund der oben beschriebenen Aspekte. Bei der Auswertung der Daten aus der Pilotstudie war zusätzlich zur Frage nach dem Geräuscheffekt auch auf diese Aspekte geachtet worden.

Die Auswertung zum Geräuscheffekt hatte bereits bei einer kleinen Stichprobe (N=12) nachteilige Effekte eines sprachlichen Hintergrundgeräusches im Vergleich zu einer Ruhe-

bedingung gezeigt [5, 6]. Auf der Grundlage der Erkenntnisse aus der Pilotstudie war die Aufgabe vor dem Einsatz in der aktuellen Studie leicht modifiziert worden. Ein wesentliches Ziel der hier beschriebenen Studie war u. a., die aufgabenspezifischen Aspekte bei einer größeren Stichprobe zu untersuchen, um die Aufgabe für zukünftige Anwendungen noch weiter optimieren zu können. Zusätzlich zur Auswertung der Gesamtleistung in den verschiedenen Durchgängen (vgl. [7]) wurden daher auch Analysen durchgeführt, in denen die Bearbeitungszeiten und die Arten der Bearbeitungsfehler im Fokus standen. In diesem Beitrag werden Ergebnisse dieser weiterführenden Analysen berichtet.

## Methode

Die nachfolgend beschriebene Studie wurde als Laborstudie mit externen Teilnehmer(inne)n an der BAuA in Dortmund durchgeführt.

*Teilnehmer und Teilnehmerinnen:* An der Studie haben insgesamt 51 berufstätige Personen teilgenommen. In die Ergebnisauswertung gingen die Daten von 45 Personen (29 Männer, 16 Frauen) im Alter zwischen 30 und 59 Jahren ein. (Zum Ausschluss von Proband(inn)en: siehe [7].)

*Leseaufgabe:* Die Teilnehmer(innen) hatten die Aufgabe, in geschriebenen Sätzen, die am Bildschirm präsentiert wurden, Fehler zu finden. Derzeit hat die Aufgabe zwei Versionen (A und B) mit jeweils 52 Items. Die Items bestehen aus ein oder zwei Sätzen unterschiedlicher Länge und Komplexität. Die Items sind entweder korrekt oder sie enthalten einen Fehler. Die Fehler variieren in der Art und im Schwierigkeitsgrad. Die Items wurden jeweils einzeln auf dem Bildschirm dargestellt, und das fehlerhafte Wort konnte jeweils mit einem Eingabestift oder mit dem Finger auf dem Touchscreen markiert werden. Die Entscheidung, ein Wort, das einen Fehler enthält bzw. den Fehler darstellt, zu markieren, oder die Entscheidung, dass das Item fehlerfrei ist (am Bildschirm gegebenes Feld „alles richtig“), war jeweils mit dem Antippen eines Feldes „fertig“ am Bildschirm zu bestätigen. Zwischen den einzelnen Items waren für vier Sekunden die Textfelder am Bildschirm leer, bis ein neues Item erschien. Die Bearbeitung der Aufgabe war mit einem gewissen Zeitdruck verbunden, da die vorab festgelegten Darbietungszeiten für jedes Item mit 8 bis 22 Sekunden relativ kurz waren. Die Aufgabe war den Teilnehmer(inne)n vorher nicht bekannt.

*Akustische Bedingungen:* Insgesamt wurden in der Studie zwei akustische Bedingungen realisiert.

(1) Ruhebedingung: Die Teilnehmer(innen) trugen während der Aufgabebearbeitung einen geschlossenen, ohraufliegenden Kopfhörer (Sennheiser HD 25), über den jedoch kein Geräusch eingespielt wurde.

(2) Geräuschbedingung „Sprache“: In dieser Bedingung wurden über den Kopfhörer (s. o.) stereophon Ausschnitte eines Hörspiels eingespielt. In den Ausschnitten waren mehrere Sprecher in unterschiedlichen Räumen zu hören. Das Hintergrundgeräusch entsprach dem Geräusch, das in der Pilotstudie eingesetzt worden war [5, 6]. Die Schalldruckpegel lagen an den beiden für die Studie eingerichteten Arbeitsplätzen (AP) bei:

- AP 1: links  $L_{eq} = 57.5$  dB(A), rechts  $L_{eq} = 60.4$  dB(A)
- AP 2: links  $L_{eq} = 57.9$  dB(A), rechts  $L_{eq} = 61.1$  dB(A)

**Untersuchungsdesign:** Die Teilnehmer(innen) haben die Leseaufgabe zweimal in der Ruhebedingung (Gruppe 1, RR,  $n = 22$ ) oder zweimal in der Geräuschbedingung (Gruppe 2, GG,  $n = 23$ ) bearbeitet (erster Messtermin = D1; zweiter Messtermin = D2). Jede Person hat dabei einmal Version A und einmal Version B der Leseaufgabe erhalten. Die Abfolge der Testversionen war in beiden Gruppen ausbalanciert. Zwischen den beiden Messzeitpunkten lag ein Zeitraum von mindestens einer Woche und maximal 10 Tagen.

**Ort der Untersuchung:** Die Studie fand in einem großen schallgedämmten Labor (L: 8,4 m x B: 6,4 m x H: 4,2 m) mit geringer Nachhallzeit statt. Im Labor waren zwei gleiche Arbeitsplätze eingerichtet, so dass die Testung mit zwei Teilnehmer(inne)n gleichzeitig durchgeführt werden konnte. Die beiden Arbeitsplätze waren durch eine undurchsichtige Stellwand voneinander getrennt.

**Ablauf:** Nach einer mündlichen und schriftlichen Instruktion sowie Übungsdurchgängen mit und ohne Zeitbegrenzung haben die Teilnehmer(innen) die Aufgabe selbständig bearbeitet. Sie waren darüber informiert, dass die Items nur für eine begrenzte Zeit sichtbar sind und nur in dieser Zeit bearbeitet werden können. Im Anschluss an die Bearbeitung der Leseaufgabe wurden alle Teilnehmer(innen) in jedem Durchgang gebeten, die erlebte Anstrengung, die Konzentration, die eigene Leistung bei der Aufgabenbearbeitung und die empfundene Störung durch die akustische Bedingung einzuschätzen. Ihr Urteil gaben sie jeweils auf einer Skala von 0 bis 100 ab. (Weitere Informationen zum Ablauf: siehe [7].)

Für das Forschungsprojekt, in dem die beschriebene Studie stattfand, liegt ein positives Votum der Ethikkommission der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin vor. Alle Teilnehmer(innen) wurden vor Beginn der Untersuchung über die Inhalte der Studie sowie über datenschutzrechtliche Aspekte informiert, und alle Teilnehmer(innen) gaben schriftlich ihr Einverständnis, an der Studie teilzunehmen.

## Ergebnisse

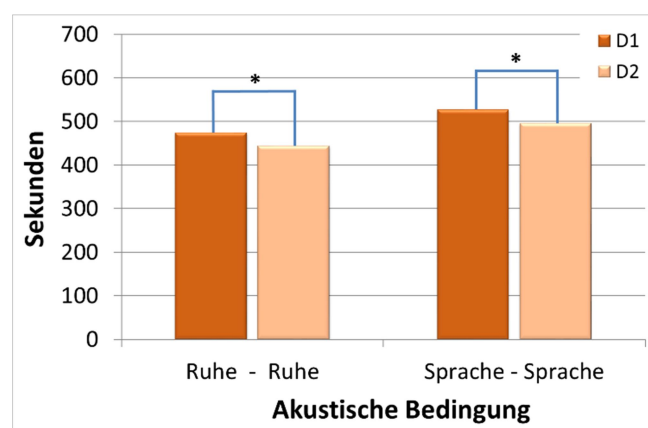
### Gesamtleistung im ersten und im zweiten Testdurchgang

Die Ergebnisse für die „Anzahl richtig bearbeiteter Items“ (Gesamtleistung) in den beiden Testdurchgängen D1 und D2 wurden beim ICSV26 präsentiert [7]. Es hatte sich ein Trainingseffekt gezeigt. Im Mittel ( $M$ ) wurden im zweiten Durchgang signifikant mehr Items richtig bearbeitet als im ersten Durchgang (D1 = 32,93; D2 = 35,89). Auch bei einer getrennten Betrachtung der beiden Gruppen zeigte sich

jeweils ein signifikanter Unterschied zwischen den Mittelwerten im ersten und im zweiten Testdurchgang (RR: D1 = 34,73; D2 = 37,09; GG: D1 = 31,22; D2 = 34,74).

### Bearbeitungszeiten

Die Darbietungszeiten für die einzelnen Items sowie die Zeiten zwischen den Darbietungen waren festgelegt. Die Darbietungszeit entspricht der max. Bearbeitungszeit. Für die nachfolgenden Berechnungen wurden die Bearbeitungszeiten für die einzelnen Items (ohne Zwischenzeiten) aufsummiert. Diese Summe wird nachfolgend als „Gesamtbearbeitungszeit“ bezeichnet. In dieser Summe sind die Zeiten für alle Items enthalten, unabhängig davon, ob ein Item richtig, nicht richtig oder gar nicht bearbeitet wurde. Im letzten Fall entspricht die Bearbeitungszeit der festgelegten Darbietungszeit. Das Ergebnis der Auswertung ist in Abbildung 1 zusammengefasst.



**Abbildung 1:** Mittelwerte der Gesamtbearbeitungszeit in Sekunden im ersten und im zweiten Durchgang, getrennt für die beiden akustischen Bedingungen (Ruhe, Ruhe; Sprache, Sprache).

Sowohl für die Gesamtgruppe als auch für die beiden Untergruppen zeigte sich ein signifikanter Unterschied zwischen den Mittelwerten im ersten und zweiten Durchgang. Die Bearbeitungszeit war im Mittel im zweiten Durchgang signifikant kürzer als im ersten Durchgang:

Gesamt: [ $t(44) = 6,1, p < 0.05$ ];  $M_{D1} = 501$  s;  $M_{D2} = 470$  s

RR: [ $t(21) = 4,399, p < 0.05$ ];  $M_{D1} = 474$  s;  $M_{D2} = 444$  s

GG: [ $t(22) = 4,168, p < 0.05$ ];  $M_{D1} = 526$  s;  $M_{D2} = 495$  s

### Bearbeitungsfehler

Für die Auswertung der Gesamtleistung wurden alle Items, die vollständig richtig bearbeitet wurden, als „richtig“ bewertet und alle anderen Items als „nicht richtig bearbeitet“. Die Analyse der Bearbeitungsfehler klärt nun auf, durch welchen ausgeführten oder nicht ausgeführten Schritt bei der Bearbeitung die Kategorisierung als „nicht richtig bearbeitet“ zustande gekommen ist. So gilt beispielsweise ein Item, das keinen konstruierten Fehler enthält, als „nicht richtig bearbeitet“, wenn trotzdem ein Wort als Fehler markiert wurde, oder wenn die Bearbeitung („alles richtig“) nicht mit „fertig“ abgeschlossen wurde. Bei den Items mit einem Fehler gibt es ebenfalls mehrere Gründe, warum die Bearbeitung eines Items in die Kategorie „nicht richtig bearbeitet“ fällt. Dies war beispielsweise der Fall, wenn der

Fehler nicht entdeckt wurde, wenn ein anderes Wort als das Zielwort (konstruierter Fehler) markiert wurde, oder ebenfalls, wenn zwar der Fehler gefunden wurde, die Bearbeitung aber nicht mit „fertig“ abgeschlossen wurde.

Für die Analyse wurden zunächst alle Arten von möglichen Bearbeitungsfehlern aufgelistet, und zwar getrennt für die Items ohne einen konstruierten Fehler und die Items mit einem konstruierten Fehler. Für die folgende Auswertung wurde ausgezählt, wie häufig welcher Bearbeitungsfehler in jeder Testversion vorlag und wie häufig welcher Bearbeitungsfehler im ersten und im zweiten Testdurchgang aufzufinden war. In Tabelle 1 sind die Summen der Bearbeitungsfehler getrennt für die beiden akustischen Bedingungen (RR und GG) in jedem Durchgang (D1 und D2) für Items mit einem konstruierten Fehler aufgelistet.

**Tabelle 1:** Häufigkeit der Bearbeitungsfehler bei Items mit einem konstruierten Fehler, getrennt für die akustischen Bedingungen (Ruhe, Ruhe; Sprache, Sprache) sowie für den ersten und den zweiten Messzeitpunkt (D1, D2). „Alles richtig“ und „fertig“ sind gegebene Felder auf dem Präsentationsbildschirm.

Bearbeitungsfehler	Akustische Bedingung			
	Ruhe - Ruhe (n = 22)		Sprache - Sprache (n = 23)	
	D1	D2	D1	D2
nicht bearbeitet	13	1	25	16
„alles richtig“; „fertig“	305	290	359	311
„alles richtig“; kein Abschluss	13	4	19	12
Zielwort markiert; kein Abschluss	6	0	8	7
anderes Wort markiert; „fertig“	20	26	21	25
anderes Wort markiert; kein Abschluss	1	1	6	1
Summe	358	322	438	372

Die Tabelle zeigt: Der häufigste Bearbeitungsfehler war, dass der konstruierte Fehler nicht gefunden wurde. Das Item wurde daher als „alles richtig“ markiert und mit „fertig“ abgeschlossen. Alle anderen Fehler finden sich deutlich seltener, und bezogen auf die Gesamtzahl der in einem Durchgang in einer Gruppe bearbeiteten Items sind die Anteile sehr gering (< 3 Prozent).

## Zusammenfassung und Ausblick

Die Ergebnisse zur Bearbeitungszeit zeigen, dass auch bei dieser abhängigen Variable ein Trainingseffekt vorliegt. Im Mittel wird die Aufgabe im zweiten Durchgang schneller bearbeitet als im ersten Durchgang. Da die Auswertung für die Gesamtleistung einen Anstieg der Anzahl richtig bearbeiteter Items vom ersten zum zweiten Durchgang gezeigt hatte, führte die schnellere Bearbeitung offenbar nicht zu einem Absinken der Arbeitsgenauigkeit. Sowohl der Anstieg der Anzahl richtig bearbeiteter Items als auch die kürzere Bearbeitungszeit in D2 finden sich auch bei der getrennten Betrachtung der beiden Untergruppen RR und GG. Der Trainingseffekt scheint somit relativ stabil zu sein, und für weitere Studien mit gleichem Untersuchungsdesign wäre er somit kalkulierbar. Ob der Effekt bei längeren oder kürzeren Zeitintervallen zwischen beiden Messterminen oder bei längerer Testdauer anders ausfallen würde, lässt sich aus diesen Ergebnissen nicht schließen. Dazu wären weitere

Studien erforderlich, in denen dies gezielt untersucht wird. Um den Trainingseffekt in der im Forschungsprojekt geplanten zweiten Studie möglichst noch weiter zu reduzieren, wird der Ablauf dahingehend verändert, dass eine längere Übungsphase eingeplant wird und ein Übungsdurchgang adaptiv gestaltet wird, so dass sich zumindest hinsichtlich der Handhabung beim Fehlersuchen alle Personen auf ähnlichem Einstiegsniveau fühlen.

Die Ergebnisse, die bei der Analyse der Bearbeitungsfehler ermittelt wurden, haben gezeigt, dass der häufigste Bearbeitungsfehler der Fehler ist, der „beabsichtigt“ ist, das heißt, der konstruierte Fehler wird übersehen und das Item wird fälschlich als „alles richtig“ markiert und mit „fertig“ abgeschlossen. Die Beobachtung, dass es verhältnismäßig wenige andere Bearbeitungsfehler gibt, lässt den Schluss zu, dass die Instruktion zur Bearbeitung der Items im Großen und Ganzen verständlich und umsetzbar ist. Eine große Zahl von nicht mit „fertig“ abgeschlossenen Items könnte beispielsweise auf Probleme mit der Instruktion oder auf unrealistisch kurze Darbietungszeiten hindeuten. Dafür gibt es aber auf der Grundlage der beschriebenen Ergebnisse keine Hinweise. Tabelle 1 hat gezeigt, dass in beiden Gruppen und beiden Durchgängen jeweils einige Male ein anderes Wort als das Zielwort markiert und das Item mit „fertig“ abgeschlossen wurde (Bearbeitungsfehler: anderes Wort markiert; „fertig“, siehe Tabelle 1). Im Verhältnis zur Gesamtzahl der in jedem Durchgang bearbeiteten Items ist der Anteil dieses Bearbeitungsfehlers sehr gering. Da es im oben genannten Forschungsprojekt jedoch ein Anliegen ist, die Ergebnisse der Studie zu nutzen, um die Aufgabe zu optimieren, wird auch diesem Ergebnis Beachtung geschenkt. Items, bei denen mehrmals ein bestimmtes anderes Wort als das Zielwort markiert wurde, werden daher ebenfalls überarbeitet.

Zusammenfassend war das Ziel der hier vorgestellten Studie, einen Überblick darüber zu erhalten, wie die Ergebnisse in den beiden Durchgängen ausfallen, wenn die Aufgabe zweimal in der gleichen akustischen Bedingung bearbeitet wird. Die Informationen aus den hier präsentierten Ergebnissen sowie aus weiteren Analysen (z. B. Schwierigkeitsindex der einzelnen Items) bilden die Basis für die Vorbereitung der nächsten empirischen Studie. In der kommenden Studie ist beabsichtigt, die Effekte unterschiedlicher, für den Arbeitskontext realistischer Geräusche zu vergleichen. Dabei wird die überarbeitete Leseaufgabe eingesetzt sowie eine weitere kognitive Aufgabe.

Das Ziel der Forschung in diesem Bereich ist, zuverlässige Aussagen darüber ableiten zu können, ob und wenn ja, in welchem Ausmaß ungünstige akustische Bedingungen die Leseleistung beeinflussen. Ein wesentliches Anliegen dabei ist, die Studien, auch wenn sie als Laborstudien durchgeführt werden, hinsichtlich möglichst vieler Aspekte mit Blick auf reale Arbeitssituationen zu gestalten. Das bedeutet, die Studien sollten mit Beschäftigten durchgeführt werden, die Aufgaben sollten vergleichbar mit Anforderungen im Arbeitskontext sein und die Geräuschbedingungen sollten auch reale Arbeitssituationen repräsentieren. Damit soll ermöglicht werden, der tatsächlichen Beeinträchtigung von Beschäftigten durch unerwünschte Geräusche bei der Arbeit

näher zu kommen. Die Erkenntnisse aus diesen Studien wären dann eine wesentliche Grundlage, um nachweislich ungünstige akustische Bedingungen gezielt verändern zu können und damit Beschäftigte vor möglichen Beeinträchtigungen besser zu schützen.

## Danksagung

Vielen Dank an alle, die an der Studie teilgenommen haben und die die Durchführung der Studie auf vielfältige Weise unterstützt haben. Insbesondere danke ich Ulrich Hold für die Programmierung der computerbasierten Leseaufgabe, Ilka Arendt für die umfangreiche Unterstützung bei der Datenerhebung und Nina Ahrweiler für die Assistenz bei der Datenauswertung.

## Literatur

- [1] Weinstein, N. D.: Effect of noise on intellectual performance. *Journal of Applied Psychology*, 59 (1974) Nr. 5, S. 548-554.
- [2] Liebl, A.: Auswirkungen von Hintergrundschall auf das Lesen und Verstehen von Texten. Dissertation, Katholische Universität Eichstätt-Ingolstadt 2006.
- [3] Venetjoki, N.; Kaarlela-Tuomaala, A.; Keskinen, E.; Hongisto, V.: The effect of speech and speech intelligibility on task performance. *Ergonomics*, 49 (2006) Nr. 11, S. 1068-1091.
- [4] Haka, M.; Haapakangas, A.; Keränen, J.; Hakala, J.; Keskinen, E.; Hongisto, V.: Performance effects and subjective disturbance of speech in acoustically different office types - a laboratory experiment. *Indoor Air*, 19 (2009) Nr. 6, S. 454-467.
- [5] Sukowski, H.; Romanus, E.: Effects of background speech on reading performance in adults. *Proceedings of Meetings on Acoustics*, 28 (2017), 050002.
- [6] Sukowski, H.: Wirkungen von Lärm auf das Lesen. Bisherige Erkenntnisse und Ergebnisse einer Pilotstudie mit Blick auf den Arbeitskontext. *Lärmbekämpfung*, 12 (2017) Nr.1, S. 19-26.
- [7] Sukowski, H.: Effects of the acoustical work environment on reading performance in employees: A laboratory study on the evaluation of a reading task. In: *Proceedings of the 26th International Congress on Sound and Vibration (ICSV26)*, Montreal, Kanada. Canadian Acoustical Association, 2019, S. 698-705.