Validität und Usability von Wearables zur Erfassung körperlicher Aktivität in der Selbstevaluation gesundheitsbezogenen Verhaltens

Sebastian Schnieder[[1]](#footnote-1),[[2]](#footnote-2) (s.schnieder@uni-wuppertal.de), Guillermo Hidalgo1,2 (guillermo.hidalgo\_gadea@uni-wuppertal.de), Sonja-Dana Roelen1,2 (s.roelen@uni-wuppertal.de), André Wiggerich2,3 (a.wiggerich@ixp-wuppertal.de) und Jarek Krajewski1,2 (krajewsk@uni-wuppertal.de)

Zusammenfassung

Die Erfassung körperlicher Aktivität zur Evaluation gesundheitsbezogener Verhaltensweisen über Wearables wie Smartwatches und Fitness Tracker ist ein vielversprechender Ansatz zur Unterstützung gesundheitsförderlichen Verhaltens sowie zur Prävention und Therapie einer Vielzahl psychischer Störungen. Notwendig hierfür sind jedoch valide Verfahren mit hoher Usability. Ziel dieser Studie ist daher der Vergleich von multiplen Sensordaten von Smartwatches und Fitness Trackern mit kalibrierten wissenschaftlichen Messinstrumenten unter Berücksichtigung von Usabilitykriterien im Rahmen unterschiedlicher Use Cases. 60 Probanden wurden dabei mit Wearables sowie einer Life-logging App ausgestattet. Die multiplen Sensordaten wurden bei der Bearbeitung von Stressaufgaben sowie Aktivitäten auf dem Laufband mit Daten eines EKG-Messgeräts sowie Videoanalysen verglichen, um so Rückschlüsse über Validität und Robustheit der Stress- und Aktivitätsmessungen zu erhalten. Darüber hinaus wurden verschiedene fragebogenbasierte Usability und User Experience Instrumente eingesetzt. Ergebnisse legen nahe, dass die eingesetzten Wearables als valider und nutzerzentrierter Self-Assessment-Ansatz langfristig dazu beitragen könnten, das Bewusstsein auf gesundheitsbezogenes Verhalten zu lenken.

Schlüsselwörter: Wearables, Evaluation, Usability, User Experience, EKG

Körperliche Aktivität ist ein wichtiger Einflussfaktor bei der Betrachtung von physischer und psychischer Gesundheit sowie von allgemeinem Wohlbefinden und kann insbesondere im Rahmen der Prävention sowie der Therapie von psychischen Krankheiten eine entscheidende Rolle spielen (Paluska & Schwenk, 2000). Voraussetzung zur systematischen Nutzung von Informationen über entsprechende Verhaltensmuster ist jedoch die valide Erfassung der körperlichen Aktivität über verschiedene Kontexte im Alltag der Personen (Gabrys et al., 2014). Die valide Messung der körperlichen Aktivität über einen längeren Zeitraum ist dabei jedoch mit einer Vielzahl von Schwierigkeiten verbunden.

Da sich Personen ihrer körperlichen Aktivität besonders in wechselnden und dynamischen Kontexten nicht oder nur unzureichend bewusst sind, scheint der Einsatz von subjektiven Selbstauskunftsinstrumenten zur Erfassung der körperlichen Aktivität nicht angemessen. Subjektive Fragebogeninstrumente werden darüber hinaus auch im Hinblick auf einen möglichen Recall Bias, soziale Erwünschtheit sowie anderen Limitationen eher kritisch diskutiert (van Poppel et al., 2010). Auch die im Bereich der Sport- und Bewegungswissenschaften häufig als wissenschaftlicher Goldstandard bezeichnete Methode zur Bestimmung des aktivitätsinduzierten Energieumsatzes über die so genannte doubly-labeled-water-method, eine Methode die die Differenz in der Abgabe von Wasserstoff- und Sauerstoff-Isotopen nach Einnahme doppelt stabil markierten Wassers beschreibt, ist für den Alltagsgebrauch nicht anwendbar (Ainslie, Reilly & Westerterp, 2003). Ebenso sind intensive Verhaltensbeobachtungen für Langzeitmessungen nicht vertretbar.

Als angebracht erscheint daher ein Ansatz zur Erfassung der Aktivität über verschiedene sensorbasierte elektronische Messinstrumente. Messgenauigkeit und Zuverlässigkeit im Vergleich zu wissenschaftlichen Standards und Kriterien spielen hier jedoch ebenso eine Rolle wie Nutzerakzeptanz und Gebrauchstauglichkeit der eingesetzten Geräte. Um den hohen wissenschaftlichen Vergleichsstandards genügen sowie eine hohe Nutzerakzeptanz erzielen zu können, ist eine möglichst hohe Genauigkeit der einzusetzenden Geräte, also eine hohe Erkennungsrate bei gleichzeitig geringer False-Alarm-Rate, besonders relevant. Daher sollte auf einen multiplen Sensoransatz zurückgegriffen werden, um sowohl intraindividuelle wie auch interindividuelle Unterschiede der Nutzer bei der Nutzung der Geräte erfassen zu können. Besondere Relevanz sollte zudem dem Komfort der Nutzer zukommen. Um einen erfolgreichen Einsatz im Langzeit-Monitoring zu gewährleisten, ist eine hohe compliance der Nutzer unumgänglich. Dies kann nur gelingen, wenn den Geräten eine hohe Usability zugeschrieben wird. So wird deutlich, dass sich bisherige wissenschaftliche Messinstrumente wie z.B. EKG nicht für Langzeiterhebungen im Alltagskontext von Personen eignen.

Neuere Entwicklungen im Bereich von Wearables zur Messung körperlicher Aktivität über multiple Sensoren erscheinen hier vielversprechend. Smartwatches und smart bands werden analog einer Armbanduhr am Handgelenk getragen und mit einem Smartphone gekoppelt. Sie verfügen über einer Reihe von Sensoren, wie z.B. einem Umgebungslichtsensor, Beschleunigungssensor, Kompass, Gyroskop, Photoplethysmographie Sensor oder GPS. Die Bedienung kann dabei über Sprach-, Berührungs- oder Gesteneingabe erfolgen. In Kombination mit verschiedenen Apps kann so Verhalten von Personen in unterschiedlichen Kontexten über einen längeren Zeitraum differenziert erfasst werden.

Das Ziel dieser Studie ist daher der Vergleich von multiplen Sensordaten von Wearables mit kalibrierten wissenschaftlichen Messinstrumenten zur Aktivitätsmessung unter Berücksichtigung von Usability Kriterien im Rahmen unterschiedlicher Use Cases, um so Informationen über die Validität, die Robustheit der Messung sowie die Usability der eingesetzten Wearables zur Analyse gesundheitsbezogenener Verhaltensweisen zu bekommen und so langfristig ein möglichst präzises und für den Nutzer komfortables Self-Assessment von Verhalten in verschiedenen Nutzungskontexten mittels Einsatz von Wearables bzw. von Life-Tracking-Apps zu ermöglichen.

**Methode**

Im Rahmen der Untersuchung wurden bisher 60 Probanden untersucht, die über Aushänge an der Universität Wuppertal rekrutiert wurden, um so die multiplen Sensordaten der Fitness Tracker in verschiedenen standardisierten Use Cases mit wissenschaftlichen Messverfahren vergleichen zu können. Den Probanden wurden zu Beginn die Smartwatches am Handgelenk der nicht-dominanten Hand angelegt. Zudem wurde den Probanden ein gekoppeltes Smartphone mit einer Life-logging App, ausgehändigt. Einen Überblick über den zeitlichen Ablauf der Untersuchung mit den verschiedenen Use Cases bietet Abbildung 1.

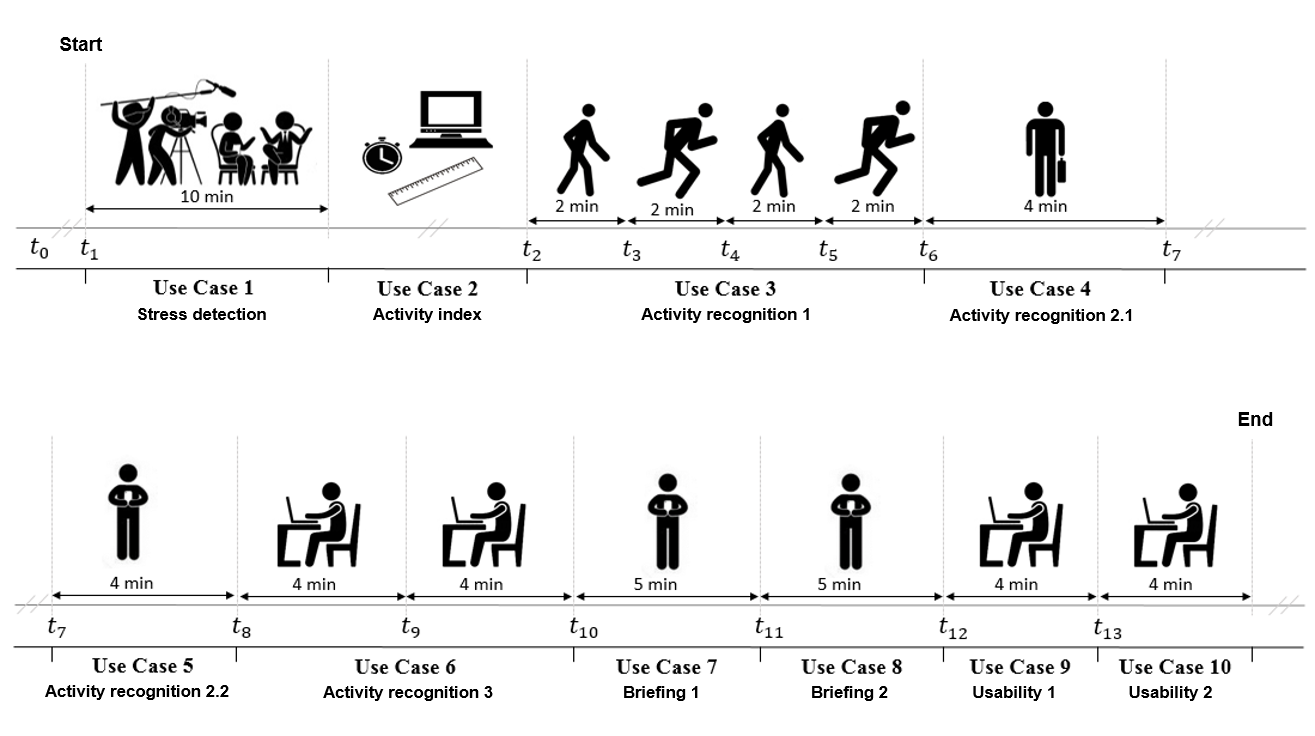


Abb. 1: Zeitlicher Ablauf der Untersuchung mit Use Cases. Use Case 1: Stresserkennung (Trier Social Stress Test), Use Case 2: Aktivitätsindex (Herzrate, Schrittzahl, Entfernung, Geschwindigkeit), Use Case 3-6: Aktivitätserkennung (Gehen, Laufen, Gehen mit Koffer, Gehen mit Smartphone, Computer Tippen, Sortieraufgabe), Use Case 7-8: Einweisung (Smartwatch, Life-logging App), Use Case 9-10: Usability (Smartwatch, Life-logging App).

Im ersten Use Case zur Stresserkennung wurde ein Teil des Trier Social Stress Test (Kirschbaum, Pirke & Hellhammer, 1993) zur Stressinduktion eingesetzt. Aufgabe der Probanden war es dabei insgesamt 5 Minuten lang beginnend bei der Zahl 1022 in 13er Schritten zu subtrahieren. Bei falschen Lösungen musste dabei wieder von neuem begonnen werden. Die stressbezogenen Herzratenvariabilitätsdaten der Smartwatches wurden hierbei mit Daten eines EKG- und Aktivitätssensors (ekgMove, movisens GmbH), der über einen Brustgurt getragen wurde, verglichen. Nach einer fünfminütigen Erholungsphase wurde im Use Case 3 die Aktivitätserkennung der Fitness Tracker überprüft (siehe Abbildung 2). Hierzu befanden sich die Probanden auf einem Laufband und sollten in zwei Geschwindigkeitsstufen jeweils zwei Minuten gehen (ca. 4 km/h) bzw. laufen (ca. 6 km/h).



Abb.2: Protokollierung des Use Case 3: Laufen und Gehen auf dem Laufband.

Dieser Ablauf wurde anschließend einmal wiederholt. Die Aktivitätsdaten der Wearables wurden dabei sowohl über den EKG- und Aktivitätssensor, über videobasierte Schrittzählung mit Hilfe der Videobearbeitungssoftware Kinovea, sowie die Entfernungs- und Geschwindigkeitsangaben des Laufbandes überprüft (siehe Abbildung 3).

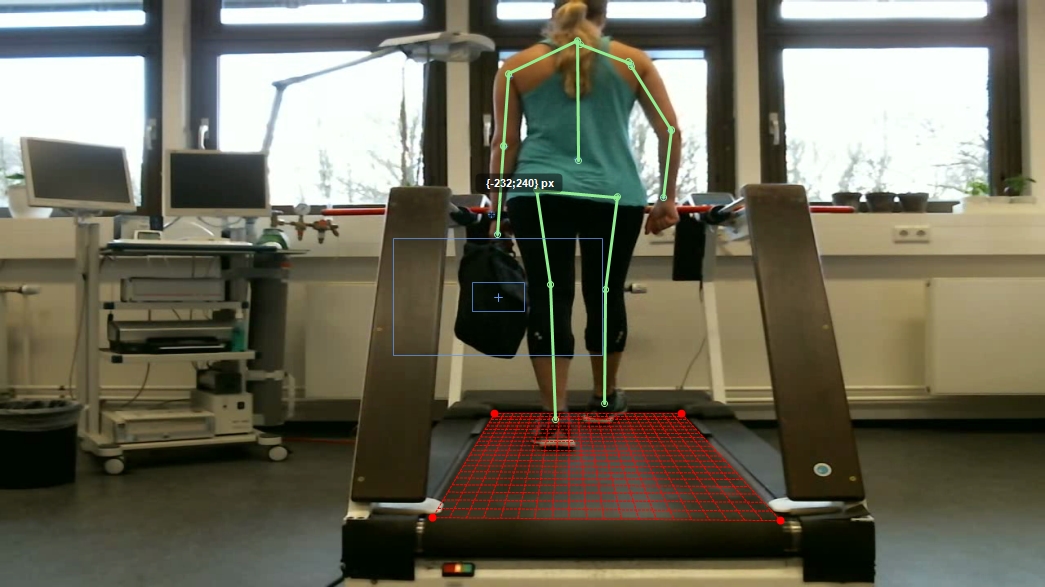


Abb. 3: Auswertung des Bewegungsablaufes in Use Case 4.

Zur weiteren Überprüfung der Messgenauigkeit und Robustheit der Messung gegen Störeinflüsse der am Handgelenk befindlichen Smartwatches, trugen die Probanden in Abänderung dieses Use Case während des Laufens einen Koffer in der Hand oder beschäftigten sich während des Gehens. So sollten mögliche Interferenzen mit den die eigentliche Aktivität überlagernden Tätigkeiten untersucht werden (siehe Abbildung 4).



Abb. 4: Use Case 5: Aktivitätserkennung durch Handgelenksbewegung ohne übliche Armschwingung während des Gehens.

In einem weiteren Use Case zur Aktivitätserkennung sollten die Probanden an einem PC sitzend einen Text über Tastatur in ein Textverarbeitungsprogramm eingeben sowie Karten eines Kartenspiels vier Minuten lang mischen. Auch hierbei sollte die Interferenz der sitzenden Aktivität mit der Bewegung des Handgelenks mittels Videoanalysen untersucht werden.

Nach Bearbeitung dieser Use Cases bekamen die Probanden die Aufgabe, sich fünf Minuten mit den unterschiedlichen Funktionen der Smartwatch selbstständig vertraut zu machen. Der Versuchsleiter kommunizierte dabei über Kurznachrichtendienste sowie Email mit den Probanden, sodass verschiedene Funktionen der Wearables deutlich wurden. Anschließend sollten sich die Probanden ebenfalls fünf Minuten mit der Life-logging App (vgl. Abbildung 5) des Smartphones beschäftigen, um einen Überblick über Bedienung und Funktion dieses Tools zu erhalten.

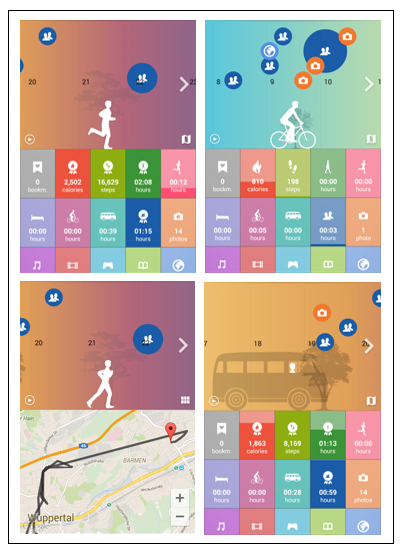


Abb. 5: Screenshots der Aktivitätserkennung der verwendeten Life-logging App während des Laufens und Fahrradfahrens sowie ein GPS-Verlauf und Transport.

Zur Erfassung der Usability sowie User Experience der Smartwatches sowie der Life-logging App wurden dabei drei fragebogenbasierte Instrumente eingesetzt, um unterschiedliche Aspekte des Nutzererlebens abzubilden. Bei den Instrumenten handelt es sich um den User Experience Questionnaire (UEQ) von Laugwitz, Schrepp und Held (2006), sowie den Mobile Phone Usability Questionnaire (MPUQ) von Ryu und Smith-Jackson (2006) und den ISONORM 9241/110- S von Prümper (2007).

# Ergebnisse

Anhand erster Untersuchungen zeigten sich folgende vorläufige Ergebnisse.

Sowohl die stressbezogenen Herzratenvariabilitätsparameter des EKG-Messgeräts als auch das Stress-Scoring der Smartwatches zeigten einen Anstieg des Stresslevels im Use Case Stresserkennung sowie einen Abfall des Stresslevels in der anschließenden Erholungsphase. Signifikante Unterschiede im relativen Stress-Scoring zwischen den beiden Erhebungsarten (Elektrokardiogramm und Photoplethysmographie) müssen im Laufe der Studie genauer untersucht werden. Anschließend sollen die PPG Messungen der unterschiedlichen Fitness Tracker verglichen werden. Die über der Smartwatches bestimmte Anzahl zurückgelegter Schritte zeigte eine hohe Pearson-Korrelation von r = 0.75 zur Messung des Referenzsystems Movisens. Darüber hinaus zeigten auch die Aktivitätsklasse sowie der Energieumsatz der durch das EKG-Messgerät bestimmt wurde hohe Übereinstimmung von den über die life-logging App bestimmten Werten. Die Messwerte erwiesen sich zudem als robust gegenüber Störeinflüssen über verschiedene Handgelenksbewegungsarten. So konnten über Mittelwertsvergleiche auch in den hierfür abgeänderten Use Cases 4 und 5 keine signifikanten Unterschiede in der Bestimmung der Aktivitätsklasse zwischen den Geräten festgestellt werden.

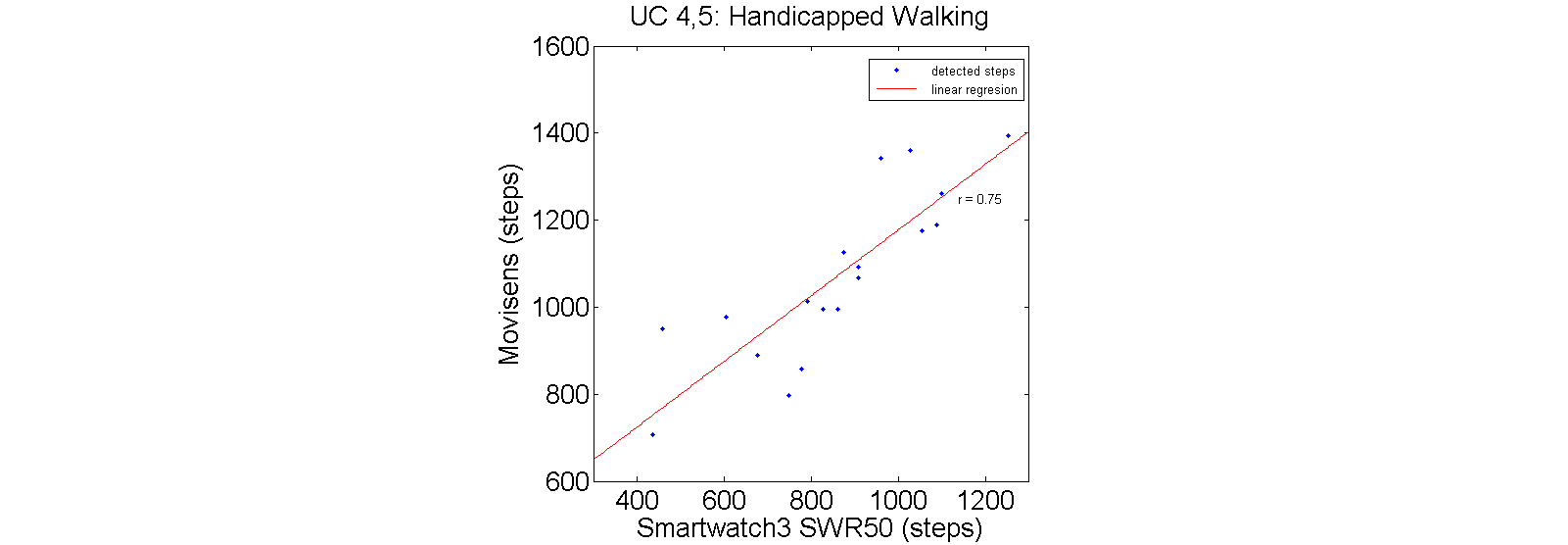


Abb. 6: Lineare Regression der Schrittzählung eines Fitnesstrackers zum Referenzsystem in Use Case 4 und 5 (Gehen mit Störeinflüßen).

Ergebnisse der Usability Fragebögen zeigten auf den unterschiedlichen Skalen der Erhebungsinstrumente eine zufriedenstellende bis sehr gute Usability der Smartwatch sowie der Life-logging App. Auch im User Experience Questionnaire wurde die Smartwatch durch die Nutzer als attraktiv und effizient bewertet.

Die Ergebnisse lassen somit darauf schließen, dass durch Wearables das Verhalten von Personen in unterschiedlichen Use Cases äußerst zufriedenstellend im Vergleich zu wissenschaftlichen Erhebungsinstrumenten eingeschätzt werden kann. Diese vorläufige Erkenntnisse in Kombination mit den positiven Ergebnissen zur Usability zeigen, dass die Technologie als valider und nutzerzentrierter Self-Assessment-Ansatz einsetzungsfähig ist und dazu beitragen könnte, gesundheitsbezogenes Verhalten zu untersuchen.

# Literatur

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health. *Sports medicine*, *29*(3), 167-180.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The ‘Trier Social Stress Test’–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76-81.

Laugwitz, B., Schrepp, M., & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In A.M. Heinecke & H. Paul (Hrsg.): Mensch & Computer 2006, 125–134. München: Oldenbourg Verlag.

Ryu, Y. S., & Smith-Jackson, T. L. (2006). Reliability and validity of the mobile phone usability questionnaire (MPUQ). *Journal of Usability Studies, 2*(1), 39-53.

Prümper, J. (2007). Fragebogen ISONORM 9241/110- S. http://www.seikumu.de/de/dok/dok-echtbetrieb/Fragebogen-ISONORM- 9241-110-S.pdf. Accessed 7 May 2015.

van Poppel, M. N., Chinapaw, M. J., Mokkink, L. B., Van Mechelen, W., & Terwee, C. B. (2010). Physical activity questionnaires for adults. *Sports Medicine*, *40*(7), 565-600.

Ainslie, P. N., Reilly, T., & Westerterp, K. R. (2003). Estimating human energy expenditure. *Sports Medicine*, *33*(9), 683-698.

Gabrys, L., Thiel, C., Tallner, A., Wilms, B., Müller, C., Kahlert, D., ... & Vogt, L. (2014). Akzelerometrie zur Erfassung körperlicher Aktivität. *Sportwissenschaft*, *45*(1), 1-9.

1. Bergische Universität Wuppertal, Institut für Sicherheitstechnik, Experimental Industrial Psychology [↑](#footnote-ref-1)
2. Institut für experimentelle Psychophysiologie GmbH, Wuppertal

   3 Technische Universität Dresden [↑](#footnote-ref-2)