

## **Konkordanzanalyse einer informationstechnischen Messmethode zur Lidschlagdetektion in der Beanspruchungsforschung**

Norman REßUT, Annette HOPPE

*Fachgebiet Arbeitswissenschaft/ Arbeitspsychologie  
Brandenburgische Technische Universität Cottbus-Senftenberg  
Siemens-Halske-Ring 14, D-03046 Cottbus*

**Kurzfassung:** Alle Diagnosen, Beurteilungen und Tests können fehlerbehaftet sein. Selbst technische Messungen sind nicht beliebig genau. Wenn im Bereich der Diagnostik ein neues Mess-, Test-, oder Beurteilungsverfahren eingesetzt oder zur Qualitätssicherung eine näherungsweise Übereinstimmung von mehreren Mess-, Test-, bzw. Beurteilungsverfahren aufgezeigt werden soll, sind Analysen zur Übereinstimmungsbewertung (sog. Konkordanzanalysen) notwendig. Methoden welche in diesem Zusammenhang eine Messgröße ohne Fehler (bzw. mit hinreichender Genauigkeit) messen, werden üblicherweise als Goldstandard bezeichnet. Bei der Einführung von neuen Verfahren, ist somit vorerst zu prüfen, inwieweit die erzielten Ergebnisse mit denen der etablierten Methoden übereinstimmen. Im Rahmen der psychischen (mental) Beanspruchungsforschung des Fachgebiets Arbeitswissenschaft und Arbeitspsychologie der BTU-Cottbus-Senftenberg, wurde ein eigens konzipiertes, digitales Lidschlagdetektionsverfahren zur Beanspruchungsermittlung eingesetzt und aus Gründen der Validität, zusätzlich einer Konkordanzanalyse unterzogen.

**Schlüsselwörter:** Konkordanzanalyse, Messtechnik, semiautomatisches Softwaresystem, videobasierte Lidschlagdetektion, Beanspruchungsforschung

### **1. Kurzvorstellung des Softwaresystems zur Lidschlagdetektion**

Das zugrundeliegende Softwaresystem zur Erfassung des Lidschlagverhaltens besteht aus mehreren Hardware (HW) - und Software (SW) - Komponenten. Aus den Eigenschaften des menschlichen Lidschlages (LS) und den Restriktionen der LS-forschung ergaben sich zwei wichtige Attribute für die videotechnischen Erhebungsinstrumente. Zum einen, die umfassenden Leistungsanforderungen, wie eine hinreichende Aufnahmegeschwindigkeit, eine hohe Bildqualität sowie eine kabellose Netzwerkintegration und zum anderen, eine praktische, handhabbare Kameratechnik mit überschaubaren Abmessungen für etwaig folgende praktikable Einsätze an Arbeitsplätzen. Diese Leistungs- und Abmessungsspezifikationen stehen sich in den Ausprägungen zumeist komplementär gegenüber. Mit der kleinen, digitalen 4K-Videokamera: FDR-X1000VR ist es möglich, 120 Bildern pro Sekunde in HD-Qualität (1920 x 1080 Pixel je Bild) aufzunehmen (Sony 2015). Wobei an dieser Stelle festgehalten werden muss, dass für längere Aufnahmesequenzen unter diesen Geschwindigkeiten und mit dieser Bildqualität, zusätzliche Vorkehrungen zur Kamerakühlung getroffen werden müssen. Des Weiteren können nur Speicherkarten verbaut werden, die sich in der „Ultra-High-Speed“ (UHS)-Klassifizierung einordnen lassen, d.h. die minimale Schreibgeschwindigkeit auf das Speichermedium beträgt 30 MB/s. Die genutzten Medien

(„microSDXC“-Karte) besitzen jeweils Speicherkapazitäten zwischen 64-128 GB und lassen sich mit dem UHS-I U3 Format klassifizieren. (SD Association 2019) Die benötigte Rechentechnik zur nach gelagerten LS-Detektion aus den Bildern der Videoaufnahmen beläuft sich - zum Einsatzzeitpunkt - auf einen handelsüblichen Rechner mit zusätzlicher HW zur Unterstützung der (grafischen) Rechenoperationen und einem Microsoft Windows 7 Professional Betriebssystem.

Im Kontext von Anwendungs- oder SW-Systemen wird der direkte SW-Bereich in Standard- und Individual-SW unterschieden. Standard-SW sind die Programme, welche für eine Vielzahl von Kunden mit ähnlichen Aufgaben produziert wurden. Demgegenüber wird die Individual-SW speziell auf den Bedarf eines Benutzers hin entwickelt und kann häufig nicht ohne weiteres von anderen Anwendern eingesetzt werden. (Mertens et al. 2017, S. 10)

Als Standard-SW, oberhalb des Betriebssystems, finden zwei Programme für die LS-Detektion ihren Einsatz. Das semiautomatische Zerlegen der Aufnahmen in die Einzelbilder mit gleichzeitigem Freistellen der interessierenden Augenpartie, wird über eine freiverfügbare Videobearbeitungs-SW durchgeführt (Lee 2015). Als weitere Unterstützungs-SW wird ein nicht lizenzfreies Standard-Kalkulationsprogramm mit einer zusätzlichen Programmierschnittstelle als Entwicklungsumgebung genutzt. Diese Entwicklungsumgebung erlaubt die Implementierung von individuellen Programmiererweiterungen in der skriptbasierten Programmiersprache „visual basic for applications“ (VBA), die sich aus der objektorientierten Programmiersprache „visual basic“ (VB) abgeleitet hat. (Benker 2014, S. 29; Microsoft 2010) Über die dort implementierten Algorithmen ist es folgend möglich die LS über die Bildpunktfärbung zu detektieren. Abbildung 1 veranschaulicht dies mittels eines Beispielgraphen, bei welchem das lokale Minima (*Bild\_ID*: 5) im funktionalen Verlauf dem geschlossenen Auge entspricht (Reßut & Hoppe 2019).

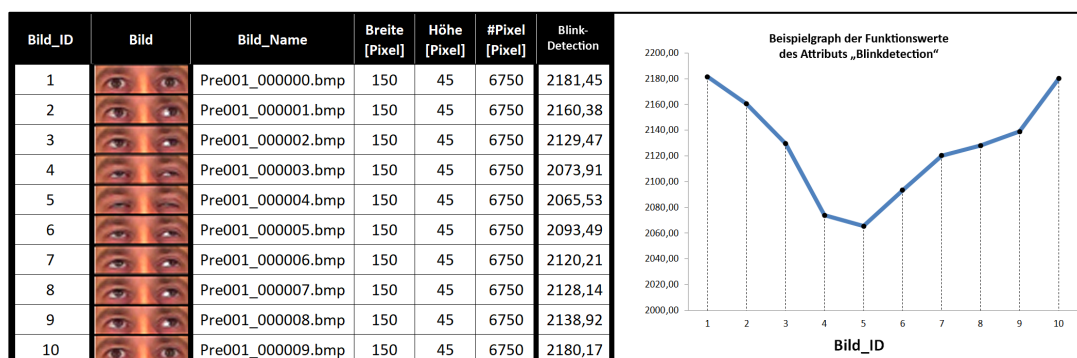


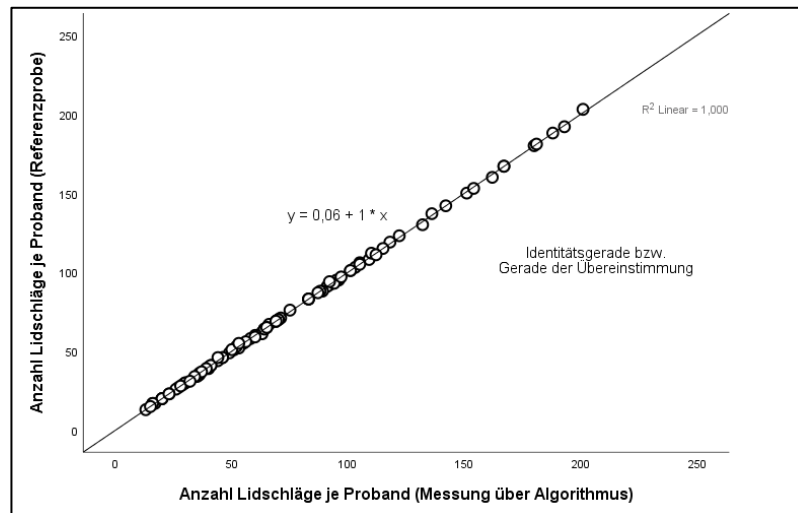
Abbildung 1: Beispielgraph des Lidschlagalgorithmus aus der Detektions-SW.

## 2. Konkordanzanalyse der Messmethode

Methoden welche in diesem Zusammenhang eine Messgröße ohne Fehler (bzw. hinreichend gut) messen, werden üblicherweise als Goldstandard bezeichnet. Bei der Einführung von neuen Verfahren, ist somit zuerst zu prüfen, inwieweit die erzielten Ergebnisse mit denen der etablierten Methode übereinstimmen. Die Güte von der Übereinstimmung zweier quantitativer Messungen lässt sich im ersten Schritt visuell veranschaulichen und zum Teil auch bewerten. Hierfür werden die Messwerte der beiden zu vergleichenden Methoden in einem Diagramm gegeneinander abgetragen. Bei

ähnlichen Messwerten befinden sich die markierenden Punkte nahe der Identitätsgeraden bzw. der Winkelhalbierenden. (Kwiecien et al. 2011, S. 515)

Abbildung 2 zeigt genau dieses übereinstimmende (lineare) Verhalten beim Vergleich der beiden angewandten Methoden in einer adäquaten Laboruntersuchung (Reßut & Hoppe 2019). Die erste Erfassungsmethode erzeugte ihre Messwerte über eine Expertensichtung und -analyse der Videoaufnahmen mit simultaner manueller Zählung der LS. Die Messmethode der manuellen Zählung wird für den Vergleichszweck als „Goldstandard“ bzw. Referenzmethode angenommen. Demgegenüber wurden die Werte der zu vergleichenden Messmethode über das zugrundeliegende und beschriebene SW-System gewonnen.



**Abbildung 2:** Streu-/Punktdiagramm von der Mess- und Referenzmethode zur LS-Erkennung.

Um das eingesetzte SW-System als messgenaues Datenerhebungsinstrument titulieren zu können, muss die Messgenauigkeit mit einem angemessen hohen Beobachterübereinstimmungs- bzw. Reliabilitätskoeffizienten nachgewiesen werden. Wobei „Beobachter“ in diesem Kontext als Synonym für die Messmethode steht. Wenn die Beobachterübereinstimmung („observer agreement“) hoch ist, wird dies als Hinweis darauf gedeutet, dass das Beobachtungssystem insgesamt problemlos auf die zu beobachtenden Fälle anwendbar ist und messgenaue Daten generiert. (Döring und Bortz 2016, S. 348) Auf dieser Grundlage und unter der oben erwähnten Annahme, dass jedes Diagnose-, Test-, Beurteilungs-, und Messverfahren fehlerbehaftet sein kann, werden im Folgenden, die Ergebnisse einer durchgeführten Konkordanzanalyse dargelegt.

Als ein Bestandteil der wissenschaftlichen Gütekriterien ordnet sich die Reliabilität (Zuverlässigkeit, Messgenauigkeit, Präzision, „reliability“) wiederum dem Qualitätskriterium der methodischen Strenge unter. Die - über Diagramme - visuell wahrnehmbare Übereinstimmung von Messmethoden kann diesen Kriterien nicht immer genügen. Daher sind für eine Qualitätsmessung der Gütekriterien, fassbare und vergleichbare Indikatoren notwendig. (Döring und Bortz 2016, S. 92) Aus diesem Grund schließt sich die Betrachtung adäquater Koeffizienten, der Analyse von der deskriptiven Statistik und des statistischen Zusammenhangs beider Messmethoden an.

Über die deskriptive Statistik ist es möglich, größere Datensätze in wenigen Maßzahlen oder Diagrammen zusammenzufassen, darzustellen und zu vergleichen. Als Fazit von diesem Vergleich ergibt sich der Sachverhalt einer fast identischen Ausprägung der jeweiligen Streumaße (vgl. Tabelle 1, SW-System vs. Referenzmethode).

**Tabelle 1:** Statistische Streumaße von den Messwertreihen der LS-Erfassungsmethoden.

Messwertreihen		SW-System	Referenzmethode	Differenz der Messwerte	Abweichung (in %)
N	Gültig	94	94	94	94
	Fehlend	0	0	0	0
Mittelwert		75,55	75,54	-0,01	0,69
Standardfehler des Mittelwerts		4,79	4,78	0,08	0,13
Std.-Abweichung		46,39	46,35	0,78	1,24
Varianz		2151,97	2148,29	0,61	1,55
Spannweite		188	190	4	5,88
Minimum		13	13	-2	0,00
Maximum		201	203	2	5,88
Median		65,00	65,00	0,00	0,00

Die Streumaße der Messwertdifferenzen beider Methoden deuten auf eine hohe Übereinstimmung der zugrundeliegenden Messwertreihen hin. Die Differenz des Mittelwertes ( $MW_{\Delta} = -0,01$ ) und die Differenz der Standardmessfehler des Mittelwertes ( $SEM_{\Delta} = 0,08$ ) befinden sich nahe der Null, die Differenzen der Standardabweichung ( $SD_{\Delta} = 0,78$ ) sowie die Differenz der Varianz ( $Var_{\Delta} = 0,61$ ) zeigen sich in ihren Ausprägungen auch relativ nahe der Null. Die maximale prozentuale Abweichung der Messwerte bei den Vergleichspaarungen ist kleiner 6%. Wobei die durchschnittliche Abweichung über alle Vergleichspaarungen hinweg bei 0,69% liegt. Die deskriptive Statistik widerspricht somit auch nicht der Annahme von einer hohen Messübereinstimmung beider Erfassungsmethoden.

Eine vollständig reliable Messung müsste bei gleichen Probanden\*innen und einer Wiederholungsmessung zu exakt den gleichen Ergebnissen kommen (perfekte Korrelation beider Messwertreihen). Weichen die Ergebnisse voneinander ab bzw. korrelieren sie nicht, so werden Messfehler dafür verantwortlich gemacht. (Döring und Bortz 2016, S. 467) Ein betragsmäßig großer Korrelationskoeffizient (nahe der 1 oder -1) ist ein guter Hinweis auf einen linearen Zusammenhang (Kwiecien et al. 2011, S. 518–519). Mit  $r = 1,000$  und  $p < 0,001$  (siehe Tabelle 2) kann die Annahme von einer hohen Messübereinstimmung der Erfassungsmethoden weiter gestützt werden.

**Tabelle 2:** Korrelation der Messwertreihen von den beiden LS-Erfassungsmethoden.

Korrelation		Referenzmethode	SW-System
Referenzmethode	Korrelation nach Bravais & Pearson	1	1,000**
	Signifikanz (2-seitig)		0,000
	Kovarianz	2148,294	2149,826
	N	94	94
Softwaresystem	Korrelation nach Bravais & Pearson	1,000**	1
	Signifikanz (2-seitig)	0,000	
	Kovarianz	2149,826	2151,970
	N	94	94

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Laut der aktuellen Literatur generiert sich des Öfteren über die Fehlinterpretation von Signifikanztests in Bezug auf den Korrelationskoeffizienten eine falsche Einschätzung. Dass die Korrelation von zwei Messmethoden signifikant von Null verschieden ist, genügt nicht, um die Übereinstimmung der Methoden zu belegen. (Kwiecien et al.

2011, S. 518–519) Da der Korrelationskoeffizient  $r$  für Verschiebungen im Niveau der Werte unsensibel ist und somit die wahre Übereinstimmung überschätzt, wird in der Literatur die Intra-Klassen-Korrelation ( $ICC_{unjustiert}$ , Typ: absolute Übereinstimmung) als das genauere Reliabilitätsmaß vorgeschlagen. (Döring und Bortz 2016, S. 570–571)

**Tabelle 3:** Intra-Klassen-Korrelation der Messwertreihen.

	Korrelation innerhalb der Klasse <sup>b</sup>	95%-Konfidenzintervall	
		Untergrenze	Obergrenze
<b>Einzelne Maße</b>	1,000 <sup>a</sup>	1,000	1,000
<b>Durchschnittliche Maße</b>	1,000 <sup>c</sup>	1,000	1,000

Modell mit gemischten Zwei-Weg-Effekten, bei dem die Personeneffekte zufällig und die Maßeffekte fest sind.

- a. Der Schätzer ist derselbe, unabhängig davon, ob ein Wechselwirkungseffekt vorliegt oder nicht.
- b. Korrelationskoeffizienten des Typs A innerhalb der Klasse unter Verwendung einer Definition der absoluten Übereinstimmung.
- c. Die Schätzung wird unter der Annahme berechnet, dass kein Wechselwirkungseffekt vorliegt, da anderenfalls keine Schätzung durchgeführt werden könnte.

Die ICC als Maß für intervall- und verhältnisskalierte Messwerte weist einen Wertebereich von  $-1$  bis  $+1$  auf. Ein Wert von Null sowie ein negativer Wert bedeuten, dass keine Übereinstimmung vorliegt. Je stärker sich der Wert der Intra-Klassen-Korrelation dem Wert  $+1$  nähert, desto besser sind die Übereinstimmung und die Reliabilität der untersuchten Messwerte. ICC-Werte über  $0,70$  werden konventionell als gute Reliabilität gewertet. (Döring und Bortz 2016, S. 571–572) Tabelle 3 zeigt den  $ICC_{unjustiert}$  mit  $1,000$  auf. Eine hohe Übereinstimmung der zugrundeliegenden Messwertreihen wird somit abermals nicht widerlegt.

Im Allgemeinen kann die Reliabilität ( $Rel$ ) als Anteil von der „wahren“ Varianz ( $Var(\tau)$ ) zu der beobachteten Varianz des Messwertes ( $Var(x)$ ) angesehen werden:

$Rel = Var(\tau)/Var(x)$ . Je größer der Anteil der wahren Varianz, umso geringer ist der Fehleranteil in den Messwerten. Laut DÖRING UND BORTZ (2016, S. 467) wird neben der empirisch ermittelbaren Varianz der Messwerte noch eine Schätzung für die (unbekannte) wahre Varianz benötigt. Unter der Annahme, dass die Methode von der Expertensichtung als „Goldstandard“ fungiert, wird für die Reliabilitätsbetrachtung die Varianz der Referenzmethode mit der wahren Varianz gleichgesetzt. Somit ergibt sich für die Reliabilitätsberechnung (mit den Werten aus Tabelle 2):

$$Rel = \frac{Var_{Expertensichtung}}{Var_{Softwaresystem}} = \frac{2148,294}{2151,970} = 0,998 \approx 1,0$$

Wird nicht die Annahme des Goldstandards (Messmethode ohne Fehler) zu Grunde gelegt, existieren, je nach Art der Schätzung von der wahren Varianz, vier weitere Methoden: Retest-Reliabilität, Paralleltest-Reliabilität, Testhalbierungs-Reliabilität und die Interne-Konsistenz. (Döring und Bortz 2016, S. 467) Die Ergebnisse aller für den Vergleich zweier Messmethoden relevanten Reliabilitätskoeffizienten gleichen sich dem dargelegten  $Rel$ -Wert bis auf die dritte Nachkommastelle und werden aus Platz- sowie Redundanzgründen nicht mit aufgeführt aber kräftigen weiter die Messübereinstimmungsannahme.

### 3. Diskussion & Fazit

Die durchgeführte Konkordanzanalyse konnte über die verschiedensten Methoden eine hohe Messübereinstimmung des entwickelten Softwaresystems zu der manuellen Expertensichtung (-zählung) aufzeigen. Zu konstatieren bleibt, dass solch eine Reliabilitätsbestimmung über zwei Messmethoden, entweder eine separate (Pretest-) Stichprobe oder - wie in dem zugrundeliegenden Fall - einen hinreichend großen Ausschnitt der Haupterhebung benötigt, wobei dieser konventionell 10 - 20% des gesamten Materials betragen sollte (Döring und Bortz 2016, S. 568). Darüber hinaus wird an dieser Stelle festgehalten, dass diese Analyse im Rahmen einer Laboruntersuchung mit klar definierten Bedingungsvariablen durchgeführt wurde. Inwieweit die entwickelte Lidschlagdetektionssoftware im Feldversuch bestehen kann und dort auch entsprechend valide und reliable Messwerte liefert, bleibt noch zu untersuchen.

### 4. Literatur

- Benker, Hans (2014): EXCEL in der Wirtschaftsmathematik. Anwendung von Tabellenkalkulationsprogrammen für Studenten, Dozenten und Praktiker. Wiesbaden: Springer Vieweg.
- Döring, Nicola; Bortz, Jürgen (2016): Forschungsmethoden und Evaluation. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kwicien, Robert; Kopp-Schneider, Annette; Blettner, Maria (2011): Concordance analysis: part 16 of a series on evaluation of scientific publications. In: Deutsches Arzteblatt international 108 (30), S. 515–521. DOI: 10.3238/arztebl.2011.0515.
- Lee, Avery (2015): What is VirtualDub? Online verfügbar unter <http://www.virtualdub.org/>, zuletzt geprüft am 04/2019.
- Mertens, Peter; Bodendorf, Freimut; König, Wolfgang; Schumann, Matthias; Hess, Thomas; Buxmann, Peter (2017): Grundzüge der Wirtschaftsinformatik. 12., grundlegend überarbeitete Auflage. Berlin: Springer Gabler.
- Microsoft (2010): Excel. Online verfügbar unter <https://products.office.com/en-us/excel>, zuletzt geprüft am 01/2020.
- Reßut, Norman; Hoppe, Annette (2019): Erfassung von individuellem Beanspruchungserleben bei kognitiven Belastungssituationen mittels Mustererkennung im Lidschlagverhalten. In: Zeitschrift für Arbeitswissenschaft, 74(4), S. 249-261. DOI: 10.1007/s41449-019-00165-y.
- SD Association (2019): SD Standards - Universal, High Performance Mobile Storage. Panasonic; San-Disk; Toshiba. Online verfügbar unter <https://www.sdcard.org>, zuletzt geprüft am 02/2020.
- Sony (2015): Vollständige technische Daten und Funktionen. X1000V / X1000VR 4K ACTION CAM MIT WI-FI & GPS. Online verfügbar unter <https://helpguide.sony.net>, zuletzt geprüft am 26.01.2019.





Gesellschaft für  
Arbeitswissenschaft e.V.

## Arbeit HUMAINE gestalten

67. Kongress der  
Gesellschaft für Arbeitswissenschaft

Lehrstuhl Wirtschaftspsychologie (WiPs)  
Ruhr-Universität Bochum

Institut für Arbeitswissenschaft (IAW)  
Ruhr-Universität Bochum

3. - 5. März 2021

---

## GfA-Press

---

**Bericht zum 67. Arbeitswissenschaftlichen Kongress vom 3. - 5. März 2021**

**Lehrstuhl Wirtschaftspsychologie, Ruhr-Universität Bochum  
Institut für Arbeitswissenschaft, Ruhr-Universität Bochum**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.  
Dortmund: GfA-Press, 2021  
ISBN 978-3-936804-29-4

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Dortmund**

**Schriftleitung: Matthias Jäger**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

**Screen design und Umsetzung**

© 2021 fröse multimedia, Frank Fröse

[office@internetkundenservice.de](mailto:office@internetkundenservice.de) · [www.internetkundenservice.de](http://www.internetkundenservice.de)