

Anwendung und Auswertung

Der Fragebogen besteht aus sechs Subskalen, welche in Tabelle 1 mit den dazugehörigen Items gelistet sind. Das theoretische Modell ist in Körber (2019) beschrieben und in Abbildung 1 visualisiert. Die Antworten zu invers formulierten Items (Item 5, 7, 10, 15 und 16) müssen vor der Auswertung invertiert werden, sodass eine höhere Zustimmung einer niedrigeren Ausprägung entspricht. Beispiel: Eine Antwort von 5 bei Item 16 wird zu einer 1 rekodiert, eine Antwort 4 zu einer 2 (und umgekehrt). Die Kategorie *keine Angabe* wird als fehlender Wert kodiert. Je nach Anteil der fehlenden Werte können diese beispielsweise auf Item-Ebene durch multiple Imputation aufbereitet werden (Mazza et al., 2015; Rioux & Little, 2021).

Jede dieser Subskalen wird einzeln ausgewertet. Die Subskalen können dementsprechend auch einzeln ohne den restlichen Fragebogen verwendet werden. Um Vertrauen in Automation in seiner Gesamtheit zu erfassen, sollte trotzdem der gesamte Fragebogen verwendet werden. Die Berechnung eines Gesamtscores aus allen Items ist auf Grund der Mehrdimensionalität mit dem derzeitigen Kenntnisstand nicht eindeutig interpretierbar.

Tabelle 1
Skalen und Items des Fragebogens

#	Item	Subskala
<i>Reliabilität/Kompetenz</i>		
1	R/K1	Das System ist imstande, Situationen richtig einzuschätzen
6	R/K2	Das System arbeitet zuverlässig
10	R/K3*	Ein Ausfall des Systems ist wahrscheinlich
13	R/K4	Das System kann wirklich komplizierte Aufgaben übernehmen
15	R/K5*	Das System könnte stellenweise einen Fehler machen
19	R/K6	Ich bin überzeugt von den Fähigkeiten des Systems
<i>Verständlichkeit/Vorhersagbarkeit</i>		
2	V/V1	Mir war durchgehend klar, in welchem Zustand sich das System befindet
7	V/V2*	Das System reagiert unvorhersehbar
11	V/V3	Ich konnte nachvollziehen, warum etwas passiert ist
16	V/V4*	Zu erkennen, was das System als nächstes macht, ist schwer
<i>Vertrautheit</i>		
3	Ve1	Ich kenne bereits ähnliche Systeme
17	Ve2	Ich habe ähnliche System bereits genutzt
<i>Intention der Entwickler</i>		
4	I1	Die Entwickler sind vertrauenswürdig
8	I2	Die Entwickler nehmen mein Wohlergehen ernst
<i>Neigung zu vertrauen</i>		
5	N1*	Bei unbekannten automatisierten Systemen sollte man eher vorsichtig sein
12	N2	Ich vertraue einem System eher als dass ich ihm misstraue
18	N3	Automatisierte Systeme funktionieren generell gut
<i>Vertrauen in Automation</i>		
9	ViA1	Ich vertraue dem System
14	ViA2	Ich kann mich auf das System verlassen

Anmerkung. *: invers formuliert.

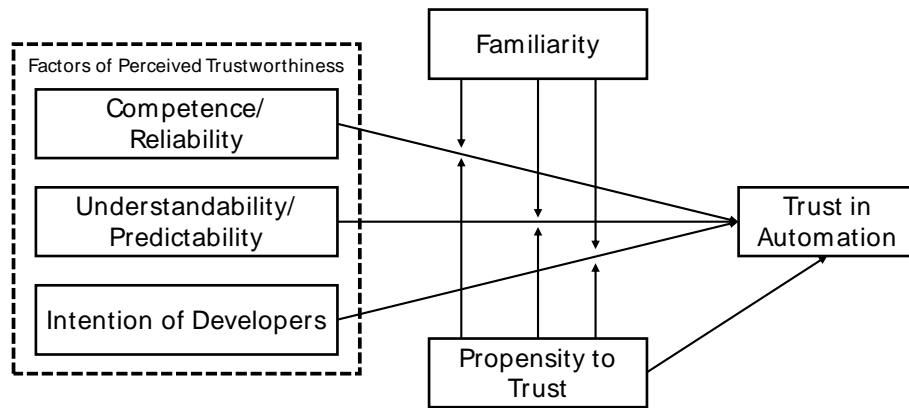


Abbildung 1. Theoretisches Modell des Vertrauens in Automation (Körper, 2019).

Psychometrische Qualität

In diesem Abschnitt werden die Ergebnisse aus zwei Studien zur Untersuchung der psychometrischen Qualität und der Konstruktvalidität präsentiert. Eine detailliertere Ausführung, welche ebenfalls den Entwicklungsprozess beschreibt, findet sich in Körper (2018) und Körper (2019). Alle berichteten Kennwerte beziehen sich auf die deutsche Version des Fragebogens.

In einer Online-Studie sahen $n = 58$ Probanden (im Alter zwischen 17 und 72 Jahren, $M = 34.00$, $SD = 15.10$, 59 % männlich, 38 % weiblich) zwei Videos entweder eines perfekt funktionierenden automatisierten Fahrzeugs (Bedingung *reliable*) oder dasselbe Video um einen *take-over request* erweitert (Bedingung *unreliable*). Die ermittelten Reliabilitätskoeffizienten der einzelnen Skalen reichen von .77 bis .95 und finden sich in Tabelle 2.

Tabelle 2
Indizes der internen Konsistenz der einzelnen Subskalen

	Omega Total	Revelles Omega
Vertrautheit	.83 ^a	-
Intention der Entwickler	.79 ^a	-
Neigung zu vertrauen	.78	.77
Reliabilität/Kompetenz	.92	.95
Verständlichkeit/Vorhersagbarkeit	.81	.88

Anmerkung. ^a: Da Omega total und Revelles Omega nicht für Skalen mit weniger als drei Items berechnet werden können, wurde hier gemäß Eisinga et al. (2013) der Spearman-Brown-Koeffizient berichtet.

Die Ergebnisse einer explorativen Faktorenanalyse (Hauptachsenfaktorenanalyse) mit obliquer Rotation (oblimin) ergaben eine Lösung mit vier Faktoren, wobei jeder Faktor mit einer ausreichenden Anzahl von Items mit hoher Ladung ($> .50$) und ohne substanzielle Querladung ($> .35$) repräsentiert wurde, sowie mittlere bis hohe Kommunalitäten. Detailliertere Kennzahlen finden sich in Körper (2019). Probanden in der Bedingung *reliable* bewerteten die Fahrzeugautomation als zuverlässiger ($t(41.32) = 3.76$, $p < .001$, $d = 1.05$). Alle Skalen korrelierten positiv mit dem Rating des Items „Ich vertraue diesem System“ auf einer 5-stufigen Ratingskala, 1 (= stimme gar nicht zu) bis 5 (= stimme sehr zu), wobei die niedrigsten Werte für die Subskala *Vertrautheit* ($r = .33$) und die höchsten Werte für die Subskala *Reliabilität* ($r = .85$) ermittelt wurden.

Die Kriteriumsvalidität wurde in einer Fahrsimulatorstudie von Körber et al. (2018) bestätigt. In dieser Studie erlebten $n = 40$ Probanden drei kritische Situationen, während sie in einem hochautomatisierten Fahrzeug (SAE Level 3) fuhren. Eye Tracking wurde benutzt, um zu messen, wie sehr sich die Probanden auf die Fahrzeugautomation verlassen. Je nach Bedingung erhielten die Probanden zu Beginn des Versuchs entweder vertrauensfördernde oder vertrauensmindernde Informationen zur Fahrzeugautomation. Der Fragebogen wurde an drei Stellen abgefragt: 1) nach dem Einführungsvideo, 2) nach der Einführungsfahrt, 3) nach der Experimentalfahrt. Die Analyse umfasste den gesamten Fragebogen *Vertrauen in Automation* (*Trust in Automation*; TiA) sowie eine genauere Betrachtung der Subskalen *Kompetenz* und *Vertrauen in Automation*. Die Subskala *Vertrauen in Automation* zeigte eine interne Konsistenz von $\alpha = .63$ (Einführungsvideo), $\alpha = .70$ (Einführungsfahrt) und $\alpha = .85$ (Experimentalfahrt). Die Subskala *Kompetenz* zeigte eine interne Konsistenz von $\alpha = .71$ (Einführungsvideo), $\alpha = .71$ (Einführungsfahrt) und $\alpha = .83$ (Experimentalfahrt). Probanden mit höherem berichteten Vertrauen zeigten konsistent in allen behavioralen Messungen mehr Verlass auf die Fahrzeugautomation: Das berichtete Vertrauen korrelierte positiv mit der Übernahmezeit ($r = .27$ bis $r = .33$) und negativ mit der minimalen time-to-collision ($r = -.29$ bis $r = -.35$). Mit steigendem Vertrauen blickten die Probanden länger auf die fahrfremde Tätigkeit ($r = .34$ bis $.35$), während sie gleichzeitig kürzer ($r = -.33$ bis $-.42$) und seltener ($r = -.40$ bis $-.44$) auf die Straße und das Kombi-Instrument blickten. Darüber hinaus berichteten Probanden, die in engen aber nicht-kritischen Situationen die Fahrzeugautomation überstimmten, niedrigeres Vertrauen ($d = 0.23$ bis $d = 0.51$).

Verwendung und Weiterentwicklung

Der Fragebogen kann im Rahmen der Lizenz (CC BY-SA 4.0) frei verwendet und weiterentwickelt werden. Mögliche Punkte zur Weiterentwicklung werden in Körber (2019) und in der README des dazugehörigen repository ([github.com/moritzkoerber/TiA Trust in Automation Questionnaire](https://github.com/moritzkoerber/TiA_Trust_in_Automation_Questionnaire)) auf Github genannt. Eine detailliertere Diskussion bezüglich des Einsatzes eines einzelnen Items gegenüber dem kompletten Fragebogen findet sich in Körber (2018) und Körber (2019). Die Referenz zur Zitation des Fragebogens lautet (APA 7th edition):

Körber, M. (2019). Theoretical considerations and development of a questionnaire to measure trust in automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (1st ed., pp. 13–30). Springer.

Literaturverzeichnis

Eisinga, R., Grotenhuis, M. t. & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International journal of public health*, 58(4), 637–642.
<https://doi.org/10.1007/s00038-012-0416-3>

Körber, M. (2018). *Individual differences in human-automation interaction: A driver-centered perspective on the introduction of automated vehicles* [Doctoral dissertation]. Technical University of Munich, Munich. <https://mediatum.ub.tum.de/1432904>

Körber, M. (2019). Theoretical considerations and development of a questionnaire to measure trust in automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander & Y. Fujita (Hg.), *Proceedings*

of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics (1. Aufl., S. 13–30). Springer.

Körber, M., Baseler, E. & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, 66, 18–31.

<https://doi.org/10.1016/j.apergo.2017.07.006>

Mazza, G. L., Enders, C. K. & Ruehlman, L. S. (2015). Addressing Item-Level Missing Data: A Comparison of Proration and Full Information Maximum Likelihood Estimation. *Multivariate behavioral research*, 50(5), 504–519. <https://doi.org/10.1080/00273171.2015.1068157>

Rioux, C. & Little, T. D. (2021). Missing data treatments in intervention studies: What was, what is, and what should be. *International Journal of Behavioral Development*, 45(1), 51–58.

<https://doi.org/10.1177/0165025419880609>