

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/346198032>

Messung der Qualität des Benutzererlebnisses bei der Verwendung von Sprachassistenten

Presentation · November 2020

CITATIONS

0

2 authors, including:



Andreas M. Klein
Universidad de Sevilla

6 PUBLICATIONS 3 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Voice Assistants [View project](#)

Messung der Qualität des Benutzererlebnisses bei der Verwendung von Sprachassistenten

Andreas M. Klein

Abteilung für Computersprachen und -systeme
Universität Sevilla, Spanien
andreas.klein@iwt2.org

Prof. Dr. Maria Rauschenberger

Fachbereich Technik
Hochschule Emden/Leer, Deutschland
maria.rauschenberger@hs-emden-leer.de

1. Vorwort (Preface)
2. Einleitung (Introduction)
3. Methodik (Methodology)
4. Studie (Study for scale construction)
5. Ergebnisse und Diskussion (Results and discussion)
6. Fazit und Ausblick (Conclusion and future work)
7. Literaturhinweise (References)

Measuring User Experience Quality of Voice Assistants

2020 15th Iberian Conference on Information Systems and Technologies (CISTI)

Andreas M. Klein

University of Applied Sciences Emden/Leer, Germany

Martin Schrepp

SAP SE, Walldorf, Germany

Andreas Hinderks

University of Seville, Spain

Jörg Thomaschewski

University of Applied Sciences Emden/Leer, Germany

DOI: [10.23919/CISTI49556.2020.9140966](https://doi.org/10.23919/CISTI49556.2020.9140966)

Publisher: IEEE

Construction of UEQ+ Scales for Voice Quality

2020 Mensch und Computer (MuC '20)

Andreas M. Klein

University of Applied Sciences Emden/Leer, Germany

Martin Schrepp

SAP SE, Walldorf, Germany

Andreas Hinderks

University of Seville, Spain

Jörg Thomaschewski

University of Applied Sciences Emden/Leer, Germany

DOI: <https://dl.acm.org/doi/10.1145/3404983.3410003>

Publisher: ACM

Exploring Voice Assistant Risks and Potential with Technology-based Users

2020 16th International Conference on Web Information Systems and
Technologies (WEBIST)

Andreas M. Klein

University of Seville, Spain

Andreas Hinderks

University of Seville, Spain

Maria Rauschenberger

Social Computing Systems, Max Planck Institute for
Software Systems, Saarbrücken, Germany

Jörg Thomaschewski

University of Applied Sciences Emden/Leer, Germany

Publisher: SCITEPRESS

https://www.researchgate.net/publication/345253241_Exploring_Voice_Assistant_Risks_and_Potential_with_Technology-based_Users

- **Voice assistants (VAs):**
 - **Great future prognoses [1]**
 - **Wide range of application [2]**
 - **General-purpose assistants**
 - 'Adaptive Voice (Vision) Assistants' [3]
 - **Siri** (Apple), **Alexa** (Amazon), **Google Assistant** (Google), **Cortana** (Microsoft), **Bixby** (Samsung),...

- **Voice User Interfaces (VUI)**

“VUI is what a person interacts with when communicating with spoken language application.” [4]

- **User Experience (UX)**

- ISO 9241-210 [5]: Holistic concept includes all types of reactions (*emotional, cognitive, physical*) before, during and after use of the product.
- Set of distinct quality criteria [6] includes the classical usability criteria (*e.g. Efficiency*) and non-goal directed *criteria* (*e.g. Stimulation*).

- **Why measuring UX quality of VAs?**
 - Great future prognoses, wide range of application [1], [2]
 - Useful tool for evaluating [7]
 - Existing tools measure usability [8]
 - Consideration of complete UX
 - Exploring improvements [9]

- **UEQ+ Framework [10]** (Modular questionnaire concept)
 - Contains 16 scales to measure different UX aspects.
 - Scales can be combined to create product-related questionnaire.
 - Scales type 1: User interaction with graphical user interfaces (GUI) [7]
- **Research gap**
 - UEQ+ lacks scales for VUI.
- **UX aspects of VUI** differ significantly from those of GUI.
 - Hearing and voice function differently than the eyes.
- **Construction of scales** to measure UX aspects of VAs.



<http://ueqplus.ueq-research.org>

- **UX aspects of VAs**

Component	Scale
User (Expectations) <ul style="list-style-type: none"> - Naturalness - Confidentiality 	Response behaviour <ul style="list-style-type: none"> - According to social norms - Human conversationalist
System (Properties) <ul style="list-style-type: none"> - Functionality - Purpose 	Response quality <ul style="list-style-type: none"> - Current information - Intention of the user fulfilled
Context <ul style="list-style-type: none"> - Environment - Purpose of activity 	Comprehensibility <ul style="list-style-type: none"> - No special formulations (syntax) - Intention of the user recognized

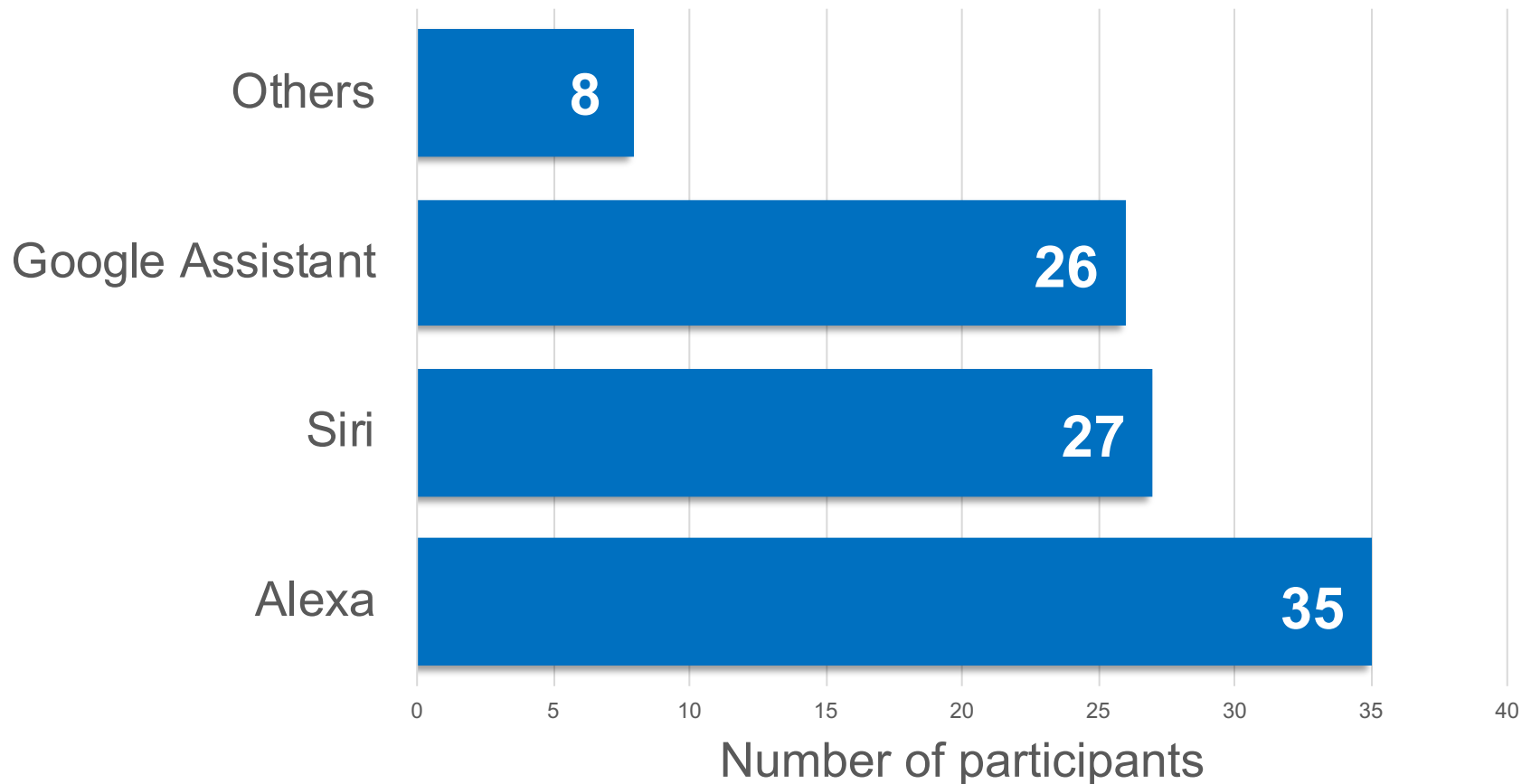
Components that influence UX in HCI [11] compared to derived scales for VUIs.

Study for scale construction

- **Scale construction**
 - Pool of candidate items for the three UX aspects for VAs.
 - Integration into the UEQ+ framework scale format.
- **Study**
 - Online questionnaire in German language.
 - Students and members of the University of Applied Sciences in Emden/Leer (Germany).
 - 96 persons participated voluntarily.
 - Participant average age: 35 years (59 male, 35 female, 2 no answer)

Results and discussion

- **Voice assistants rated by the participants:**



Results and discussion

- **Factorial analysis** of item sets (*a total of 30 candidate items*)
- Principal component analysis [13] (*Varimax rotation*)
- Assumption of **three factors** confirmed:
 - **Response behavior** loaded high on factor 2 (*low on the others*)
 - **Response quality** loaded high on factor 3 (*low on the others*)
 - **Comprehensibility** loaded high on factor 1 (*low on the others*)
- Details analysis is described in the research protocol [11].

Results and discussion

- **Response behaviour:**
 - Set of candidate items and loadings on factor 2 (varimax rotation)

No.	Items (German – Original Version)		Items (English Translation)		Loading
	1	technisch	menschlich	technical	
2	künstlich	natürlich	artificial	natural	0.80
3	fremd	vertraut	unfamiliar	familiar	0.66
4	ungewöhnlich	gewöhnlich	unusual	usual	0.25
5	langsam	schnell	slow	fast	0.48
6	unangenehm	angenehm	unpleasant	pleasant	0.75
7	unsympathisch	sympathisch	unlikeable	likable	0.81
8	unfreundlich	freundlich	unfriendly	friendly	0.66
9	langweilig	unterhaltsam	boring	entertaining	0.68

Results and discussion

- **Response behaviour:**
 - Items with highest loadings (factor 2) and introducing sentence:

<i>In my opinion the response behaviour of the voice assistant is</i>								
<i>artificial</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>natural</i>
<i>unpleasant</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>pleasant</i>
<i>unlikable</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>likable</i>
<i>boring</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>entertaining</i>

Results and discussion

- **Response quality:**

- Items with highest loadings (factor 3) and introducing sentence:

<i>The answers and questions asked by the voice assistant are</i>								
<i>inappropriate</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>suitable</i>
<i>useless</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>useful</i>
<i>not helpful</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>helpful</i>
<i>unintelligent</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>intelligent</i>

- **Comprehensibility:**

- Items with highest loadings (factor 1) and introducing sentence:

<i>In my opinion the voice assistant has understood my voice commands</i>								
<i>complicated</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>simple</i>
<i>inaccurate</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>accurate</i>
<i>unambiguous</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>ambiguous</i>
<i>enigmatic</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<i>explainable</i>

Results and discussion

- Product-related questionnaire
 - Select relevant scales
 - Product-specific UX aspects first [14]
 - Further criteria e.g. for marketing [14]
 - More information: UEQ+ Handbook (<http://ueqplus.ueq-research.org>)

Questionnaire example 1: Users applying smart home	Questionnaire example 2 Voice assistant customer service
<ul style="list-style-type: none"> ▪ Perspicuity ▪ Trust ▪ Trustworthiness of content ▪ Quality of content <p style="text-align: right;">UEQ+</p>	<ul style="list-style-type: none"> ▪ Efficiency ▪ Perspicuity ▪ Dependability ▪ Trust <p style="text-align: right;">UEQ+</p>
<div style="display: flex; align-items: center;"> <div style="background-color: yellow; padding: 5px; margin-right: 10px;"> <p style="margin: 0;">Voice quality scales</p> </div> <ul style="list-style-type: none"> ▪ Response behaviour ▪ Response quality ▪ Comprehensibility </div>	

Conclusion and future work

- **Conclusion:**
 - Approach based on UEQ+
 - Construction of VA scales
 - Study for scale construction
 - The factorial analysis confirmed:
 - **Response behaviour**
 - **Response quality**
 - **Comprehensibility**
 - Two example using new VA scales.
- **Future work:**
 - Validation of scales specific to VAs.

- [1] Tuzovic, S. and Paluch, S. (2018). Conversational Commerce – A New Era for Service Business Development?, pages 81–100. Springer Fachmedien Wiesbaden, Wiesbaden.
- [2] Tractica (2020). Tractica. <https://tractica.omdia.com/newsroom/press-releases/voice-and-speechrecognition-software-market-to-reach-6-9-billionby-2025/>.
- [3] Knote, R., Janson, A., Söllner, M., and Leimeister, J. M. (2019). Classifying smart personal assistants: An empirical cluster analysis. In Proceedings of the 52nd Hawaii International Conference on System Sciences.
- [4] Cohen, M. H., Giangola, J. P., and Balogh, J. (2004). Voice User Interface Design. Addison Wesley Longman Publishing Co., Inc., USA.
- [5] ISO9241-210, Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (2010).
- [6] Preece, J., Rogers, Y., & Sharp, H. (2015). Interaction Design: Beyond HCI, 4th edn. Wiley, Chichester, 2015.
- [7] Klein, A. M., Hinderks, A., Schrepp, M., and Thomaschewski, J. (2020b). Construction of UEQ+ Scales for Voice Quality. In Proceedings of the Conference on Mensch und Computer, MuC '20, page 1–5, New York, NY, USA. ACM.
- [8] Hone, K. S., & Graham, R. (2000). Towards a tool for the subjective assessment of speech system interfaces (SASSI). *Natural Language Engineering*, 6(3-4), 287–303.
- [9] Klein, A. M., Hinderks, A., Rauschenberger, M., & Thomaschewski, J., (2020a). Exploring Voice Assistant Risks and Potential with Technology-based Users. In Proceedings of 16th International Conference on Web Information Systems and technology (WEBIST), pages 1–8, SCITEPRESS.
- [10] Schrepp, M. and Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(7):88–95.
- [11] Hassenzahl, Tractinsky, User experience—a research agenda, *Behaviour & Information Technology*, Vol. 25, No. 2, March-April 2006, 91–97.
- [12] Klein, A. M., Hinderks, A., Schrepp, M., & Thomaschewski, J., (2020b). Protocol for Measuring User Experience Quality of Voice Assistants. DOI: 10.13140/RG.2.2.12816.35848
- [13] Revelle, W. (2018). Psych: Procedures for personality and psychological research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 1.8.12.
- [14] Winter, D., Hinderks, A., Schrepp, M. & Thomaschewski, J., (2017). Welche UX Faktoren sind für mein Produkt wichtig? In: S. Hess & H. Fischer (Eds.), *Mensch und Computer 2017—Usability Professionals*. Gesellschaft für Informatik e.V. (pp. 191–200).