

Validation of a usability questionnaire for summative evaluation of robotic systems*

E. Heinold, P.H. Rosen and S. Wischniewski

Abstract— Robotic systems are increasingly prevalent in different work domains and allow new forms of interaction between robot and worker due to their increasing autonomy. Usability plays in this context an important role for the interaction quality. The IsoMetrics Usability Questionnaire is an established survey instrument for the user-centered usability assessment of interactive systems. It is based on seven principles for the ergonomic design of interactive systems from DIN EN ISO standard 9241-110. In 2020, a new version of the standard was published, in which two principles (customizability and controllability) were merged and a new principle (user engagement) was included.

The aim of this study was to develop and validate a usability questionnaire for the summative evaluation of interactive robotic systems, based on the IsoMetrics and the interaction principles of the DIN EN ISO 9241-110. An initial item pool of 100 robot specific items was assessed in a video-based online study. A ten-factorial structure was identified as part of an exploratory factor analysis ($N = 219$). Subsequently, the item pool was reduced from 100 to 50 items. The confirmatory factor analysis ($N = 84$) could not replicate these results. To test the construct validity, convergent and divergent constructs were collected with mixed results.

The questionnaire in its preliminary form has high content validity, and indicates that the structure of the DIN EN ISO standard is replicable in the form of ten factors. However, further research is needed to identify reasons for the deviating results of the confirmatory factor analysis.

I. INTRODUCTION

In their 2021 report on industrial robots, the International Federation of Robotics (IFR) reports over 3 million robotic systems in use in 2020 [1]. Constantly advancing technological innovations produce new types of robots that can be integrated into various work environments. This enables novel forms of interaction between humans and robots [2]. While previous generations of robots were caged for workers safety, innovations in sensor input and artificial intelligence have paved the way for robotic systems that share a workspace with human workers. In that workspace the two parties can coexist, cooperate on their working task or even collaborate [2]. Workflow is also becoming increasingly flexible and instructions or queries can be transmitted to the robot via multimodal interfaces in the form of speech, augmented reality, or gestures [3]. This new variety in interaction with robotic systems offers the

chance to develop robots with an increased focus on ergonomic system design. DIN EN ISO 9241-110 is the German version of this international standard for the ergonomic design of interactive systems [4]. It includes the so called interaction principles, seven facets of a systems design relevant to its overall usability. The current seven interaction principles are:

- suitability for the task
(the dialogue is suitable for the user's task and skill level)
- self-descriptiveness
(the dialogue makes it clear how the user should proceed)
- controllability
(the user is in control of the pace and sequence of the interaction)
- conformity with user expectations (consistent performance)
- error tolerance (the dialogue is forgiving to a reasonable degree)
- suitability for learning (the dialogue supports learning)
- user engagement (the system motivates the user to proceed in the interaction)

Based on these principles, survey instruments regarding the usability of interactive systems have already been developed. An established questionnaire in this field is the IsoMetrics [5]. It was originally created for software systems, but has already successfully been applied to other types of technology, including robotic systems [6].

Since its development, robots in particular have evolved in their capabilities and functions. In addition, DIN EN ISO 9241-110 has changed in terms of content and structure since the development of the IsoMetrics questionnaire. The introduction of a new interaction principle, user engagement, and the merging of controllability and customizability change the structure of the norm [5]. Hence, there is both the need and opportunity to develop a questionnaire based on the current form of the DIN EN ISO 9241-110, incorporating established elements of the IsoMetrics, and adjusting them to fit robotic

* This article was created as part of the SOPHIA project (which is covered by funding agreement no. 871237). This project is funded by the European Union's research and innovation program "Horizon 2020" (H2020-ICT-2019-2/ 2019-2023). Eva Heinold, Patricia H. Rosen and Sascha Wischniewski are with the Federal Institute of Occupational Health and Safety, 44149, Dortmund, Germany (corresponding author to provide e-mail: Heinold.eva@baua.bund.de.)

systems specifically. Subsequently, a validation needs to be performed to ensure the reliability and quality of the new assessment tool.

II. METHODOLOGY

In preparation for the questionnaire development, an expert workshop was held at the Federal Institute for Occupational Safety and Health Germany (BAuA), in which the IsoMetrics was evaluated with regard to its suitability for assessing usability of robotic systems. This was done for service robots, as well as industrial robots. Furthermore, the need for additional items was identified to reflect new functions of robotic systems. This generated five new robot-specific items. Additional 20 items were generated for the new factor user engagement, as in the current form of the IsoMetrics, this interaction principle is not represented. The new items are closely modeled along the phrasing of the DIN EN ISO 9241-110:2020-10, to represent its content as accurately as possible. The new items were added to the items of the IsoMetrics. The original IsoMetrics items were also adjusted to fit robotic systems in their phrasing. This concluded in an initial item pool of 100 items.

The item pool created in this way was assessed using a video-based online survey. Participants were shown a work situation with human-robot interaction in two two-minute videos. The participants were then asked to rate the robots usability along the items on a five point Likert scale from „does not apply“ (1) to „completely applies“ (5). The original scale contains the option to abstain from answering individual items. This option was excluded for the purpose of data consistency during the evaluation. It can however be added for future field uses. In addition, the overall experience level participants had with robotic systems was surveyed as a control variable. An exploratory factor analysis (EFA) was performed to examine the factor structure of the items and to perform item reduction. In a second survey, the remaining items were assessed by new participants using the same video material and then evaluated using a confirmatory factor analysis (CFA). In addition, the questionnaire IsoNorm 9241/10 [7] was also included to check the convergent validity. The IsoNorm 9241/10 is based on the same interaction principles as the original IsoMetrics, hence their factors are expected to correlate. Three items per interaction principle are presented and to be judged on a seven point, bi-polar scale between --- and +++, indicating scores from -3 to +3. On each side of the scale, the same item is presented in either a positive or negative version.

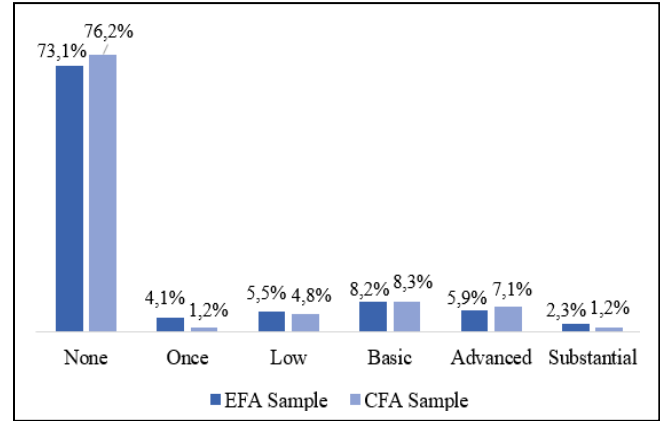
To assess the divergent validity, the Affinity for Technology Interaction-Short (ATI-S) questionnaire was included [8]. The ATI-S, is a unidimensional, four item scale with high internal consistency (Cronbach's alpha .83-.92) and good construct validity [8]. In the short version the four items are scored on a 6-point Likert scale with "Not at all true"(1), "Not true"(2), "Rather not true" (3), "Rather true" (4), "Agree to a large extent" (5) and "Agree to a great extent" (6). The ATI-S is a questionnaire regarding the affinity for technology interaction, which is not inherently related to the usability of a technology. Hence, it is not expected to correlate with the IsoMetrics factors.

A. Samples

The sample of the EFA (N = 219) is 54.8% female, 42.9% male and 2.3% diverse. The participants stated that they are between 18 and 63 years old (M = 27.85, SD = 10.19). The majority were students (55%) and employees (32%), and to a lesser extent self-employed (7.8%), unemployed (1.4%), trainees (1%) and other employees (2.7%). The majority of participants did not have any experience with robotic systems prior to the experiment. However, 21.9% of participants have interacted with a robotic system more than once before. The exact distribution can be seen in Fig. 1.

The CFA sample (N = 84) was 61% female, 38% male and 1% diverse. The majority were students (65.5%) and employees (21.4%), while the rest were self-employed (8.3%), jobseekers (2.4%), and other employees (2.4%). This sample contains no trainees. The average age was 27.58 years (SD = 8.83). This sample, too reported low levels of prior robot experience, with 21.4% having interacted with a robotic system more than once.

Figure 1. Level of experience with robotic systems



B. Statistical analysis

Before carrying out the EFA, the admissibility was confirmed using the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO), Bartlett test and correlation matrix. All tests confirmed data suitability for EFA.

This was carried out in eleven rounds based on principal component analysis and principal axis analysis with Promax rotation. In addition, scree plot and parallel analysis were used to determine the factors more precisely (Table 1). Once a preliminary structure was determined, communalities of the items within each factor were calculated. Finally, Cronbach's α was assessed for internal factor consistency. All statistical analyses for the EFA were performed with SPSS - Statistical Package for Social Sciences version 27 [9].

The CFA was carried out using the analysis program R 4.1.2 [10] in order to assess and possibly confirm the factors resulting from the EFA in a second independent sample. In addition, the packages foreign_0.8-81 [11], lavaanPlot_0.6.2 [12], lavaan_0.69 [13] and readxl_1.3.1 [14] were used. Prior to performing the confirmatory factor analysis the initial Cronbach's α of both, the base ten factors as well as all factors of convergent and divergent questionnaires were calculated.

The convergent and divergent validities were calculated using factor correlations.

III. RESULTS

A. Exploratory factor analysis

Before carrying out the EFA, the admissibility was checked using the KMO coefficient ($KMO=.785$), Bartlett test ($df(4950)=114080.519$, $p<.001$) and correlation matrix. The correlation matrix shows that each item has a correlation of .3 but no greater than .9 with at least one other item in the questionnaire. Accordingly, the prerequisites for carrying out an EFA are given.

This was performed based on principal component analysis and principal axis analysis with Promax rotation. In addition, the scree plot and parallel analysis were used to determine the factors more precisely.

The indicators show slightly deviating results from each other regarding the final structure of the questionnaire.

TABLE I. ROUND BASED RESULTS OF THE EXPLORATORY FACTOR ANALYSIS

Round	Items	Principal Component Analysis			Principal Axis Analysis	Reduced Items
		Eigenvalue >1	Explained variance	Scree-plot		
1	100	24	62.311	6	13	5
2	95	23	50.279	6	13	11
3	84	20	57.423	13	12	7
4	77	19	58.655	12/7	12	3
5	74	15	53.402	7	12	3
6	72	15	47.467	7	9	7
7	65	14	42.939	7	7	6
8	59	14	48.385	9	9	6
9	53	14	50.695	9	10	1
10	52	13	51.228	9	10	1
11	51	13	51.650	9	10	0

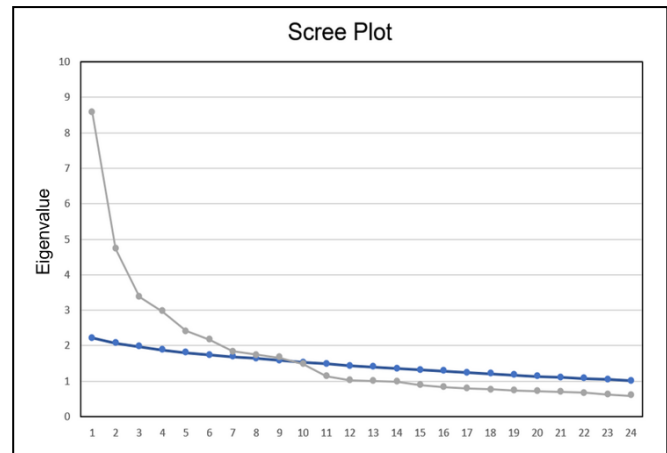
The Kaiser-Guttman Eigenvalue criterion indicates thirteen factors with eigenvalues with >1, the scree plot shows its most significant cut off point at ten factors, whereas the parallel analysis indicates nine factors, bordering on ten (Fig. 2).

Considering the factor content, and the fact that neither the thirteen nor nine factor solution noticeably improved the results, ten factors were defined as the final form for the questionnaire, as it provided the most accurate representation of the interaction principles structures in the DIN EN ISO 9241-110.

The original 100 items were reduced to 50 in eleven iterations and form the ten-factorial structure (Table 1). Items were excluded based on their factor loadings and possible cross loadings onto other factors.

The final KMO coefficient is .780 and the Bartlett test is significant ($df(1275)=5354.921$, $p<.001$).

Figure 2. Scree plot and parallel analysis



The four factors were named according to their counterparts in the DIN EN ISO 9241-110 or,

- conformity with user expectations,
- self-descriptiveness,
- suitability for the task,
- suitability for learning

Three of the interaction principles are each represented by two separate factors. The interaction principle of error tolerance is divided into functional robustness against user errors and effort required to correct user errors. The new interaction principle user engagement forms two factors: functional user engagement and motivational user engagement. The factor controllability splits into two factors named customizability and controllability.

The communalities of the items are between $h^2 = .51-.31$ for conformity with user expectations, between $h^2 = .66-.38$, for the factor functional user engagement, between $h^2 = .76-.40$ for self-descriptiveness and for customizability between $h^2 = .80-.50$. The communalities of suitability for the task are between $h^2 = .57-.46$, functional robustness against user errors between $h^2 = .60-.43$, suitability for learning $h^2 = .65-.41$, effort required to correct user errors between $h^2 = .62-.40$, motivational user engagement between $h^2 = .64-.40$ and controllability between $h^2 = .53-.36$.

The ten factors achieve a variance explanation of 51.65% and have a mostly good internal consistency according to Cronbach's α (Table 2).

A correlation analysis also showed that each of the identified factors had a significant low to moderate correlation with at least two of the remaining factors.

TABLE II. FACTOR VALUES AND CRONBACH'S α AFTER EXPLORATORY FACTOR ANALYSIS.

Factor name	M	SD	Items	Cronbach's α
Conformity with user expectations	3.62	0.95	9	.805
Functional user engagement	2.71	1.12	6	.824
Self-descriptiveness	3.18	1.05	5	.853
Customizability	3.21	1.16	4	.858
Suitability for the task	3.55	0.98	4	.791
Functional robustness against user errors	3.64	1.02	5	.793
Suitability for learning	2.98	1.15	4	.774
Effort required to correct user errors	2.54	1.13	4	.795
Motivational user engagement	3.75	1.04	5	.770
Controllability	3.17	1.08	4	.655

B. Confirmatory factor analysis

The ten-factorial structure of the EFA is used as a basis for the confirmatory factor analysis. Within the CFA a structural equation model for the factorial structure of the questionnaire is calculated. Whether the model fit is acceptable is indicated by several criteria.

In the initial ten factor model the χ^2 test is significant, χ^2 (1130, N = 85) = 2754.65, $p < .001$, where the value χ^2/df = 2.57 is acceptable. With a value of 0.337, the CFI is in the unacceptable range. The value of the SRMR is .138. The RMSEA is significant with $p < .001$ and the value is .133. Neither are in range of an acceptable model fit. In order to improve the fit of the model, the items with the weakest factor loadings and high residual variances were excluded, taking content related importance into account. In the first model, the loadings of six items are below $\beta < .300$, which is why they are excluded from further analysis [15]. This procedure was repeated five times without achieving significant improvements in the overall model fit. In the sixth calculation, correlations between the factors were allowed to improve the model fit. In total ten multifactorial models were calculated, without reaching an acceptable model fit.

The best model fit achieves a χ^2 (114, N = 85) = 277.02, $p < .001$, and χ^2/df has a value of 2.43. The RMSEA is .130 and the SRMR is .104. The CFI is .737. This model solution contains 17 items divided into four factors. This is the best model fit achieved, but does not meet the necessary criteria to be accepted.

To confirm this result, a one-factorial competition model was calculated. The model fit is not acceptable. The χ^2 test is highly significant with $p < .001$ and χ^2/df with 3.07 higher than that of the four-factorial solution. The RMSEA was significant with a value of $p < .001$, whereby the value of .156 indicates an unacceptable fit. The CFI and SRMR values of .601 and .123 were not in the acceptable range. Hence, the one-factor solution was rejected as well.

The IsoNorm 9241/10 was used to record the convergent validity. This is also based on the interaction principles. The divergent validity is checked using the ATI-S. Table 2 shows mean values, standard deviations and Cronbach's α for all survey instruments. The IsoNorm9241/10 factors correlate significantly with three matching factors: Conformity with expectations ($r=.315$), error tolerance with functional robustness against user errors ($r=.337$), and controllability ($r=.222$), with small to moderate effect sizes. Suitability for the task, customizability, suitability for learning and self-descriptiveness do not correlate with the respective corresponding factors of IsoNorm9241/10. The two factors related to user engagement are not represented in IsoNorm 9241/10. Functional user engagement correlates with IsoNorm 9241/10 conformity with user expectations ($r=.315$) and suitability for the task ($r=.301$) with a moderate effect sizes. Motivational user engagement also correlates with conformity with user expectations ($r=.362$), suitability for the task ($r=.298$), and controllability ($r=.302$).

Regarding the divergent constructs, affinity for technology interaction shows significant correlations with four of the ten factors. It correlates positively with motivational user engagement ($r=.215$) and suitability for learning ($r=.347$) with moderate effect sizes and moderately negatively with customizability ($r=-.344$) and the effort required to correct user errors ($r=-.264$).

TABLE III. FACTOR MEANS AND STANDARD DEVIATIONS FOR THE CONFIRMATORY FACTOR ANALYSIS.

Factor Name	M	SD	Cronbach's α
Conformity with user expectations	3.35	0.51	.706
Functional user engagement	3.42	0.68	.677
Self-descriptiveness	3.24	0.63	.643
Customizability	2.89	0.81	.706
Suitability for the task	3.66	0.68	.648
Functional robustness against user errors	3.17	0.61	.508

Suitability for learning	2.75	0.71	.517
Effort required to correct user errors	3.50	0.53	.240
Motivational user engagement	3.69	0.71	.749
Controllability	3.32	0.69	.567
IsoNorm 9241/10 Expectation conformity	5.63	1.03	.702
IsoNorm 9241/10 Self-descriptiveness	5.06	0.89	.632
IsoNorm 9241/10 Customisability	4.55	1.05	.485
IsoNorm 9241/10 Task adequacy	5.32	1.06	.435
IsoNorm 9241/10 Learnability	5.09	1.37	.875
IsoNorm 9241/10 Error Tolerance	5.21	0.91	.629
IsoNorm 9241/10 Controllability	4.53	0.68	.303
ATI-S	3.90	1.11	.808

IV. DISCUSSION

The overall results of the EFA and the CFA present diverting results regarding the factor structure underlying the created questionnaire.

While the EFA indicated a factor structure largely congruent with the DIN EN ISO 9241-110:2020-10, its deviations were still in line with the content of the norm. Controllability splitting into two factors more closely resembling the norm in its previous form, where customizability formed a separate interaction principle, as well as two more principles separating into subfacets. The robustness against user errors and the user engagement are divided into two factors each. These were named according to their focus on functional robustness against user errors and the effort required to correct user errors, as well as functional user engagement and motivational user engagement. Internal consistency values of the factors and explained variance of the instrument in the EFA are good. The communalities of the items within each factor were also within an acceptable range. These results supports the assumption that the DIN EN ISO 9241-110 can be represented in from of a structured questionnaire used to evaluate robotic systems. They might indicate that the construct of usability as represented in the interaction principles contains more distinct constructs on a factor basis than previous questionnaires represented. Of course, there is valuable information gained from assessing the overall interaction principles. However, for the sake of an overview, that includes each factor, crucial details within each one might be lost.

An additional outcome of the study is the first time assessment of the newly introduced factor user engagement in the context of factor analysis. The results indicate that it can be represented in form of two distinct factors. One focusing on the functional abilities of the robotic system to engage the user in the process, while the other largely focusses on the individuals' internal perception of the interaction in terms of engagement in the process. While these two factors show a clear separation between them, they do not represent the original three sub facets formulated in the norm (Motivation, Integration, and trustworthiness) [4]. All three facets were present in both factors. However, as this is the first factor analysis of this specific interaction principle, more data needs to be gathered to make more concrete conclusion on differences between how the norm is formulated and how it represents in a factor analysis.

The CFA based on the ten-factorial solution of the EFA did not achieve a sufficient model fit. Various adjustments could not achieve an acceptable model fit, even if individual indicators improved in the process. The best model contains 17 items divided into four factors. These, too, did not achieve an acceptable model fit. In this form, the remaining items also no longer reflects the content of DIN EN ISO 9241-110 [4]. The calculated models do not allow confirmation of the ten-factorial solution to the EFA, however, as the one factorial solution also did not achieve an acceptable fit, we cannot conclude any definitive structure for the questionnaire. Explaining the diverting results of the CFA poses a challenge. The methodology was identical to the EFA, and the samples did not differ significantly in their makeup from each other. Several of the model fit indicators have a documented tendency either to overestimate model fit in small samples [16] or be comparatively robust to sample size [17]. Hence, it remains unclear what caused this deviation. While data quality control measures were taken before the EFA and CFA, it is possible that the remaining datasets of the CFA contained low quality data, contributing to these results. To gain certainty on the underlying structure of the questionnaire, another independent data set should be gathered and analyzed.

The existence of other questionnaires based on DIN EN ISO 9241-110, the many years of successful use of IsoMetrics, the high internal consistency of the ten factors of the EFA, with a variance explanation of over 50%, give reason to assume that there are not completely unstructured concepts.

To check the validity of the questionnaire, bivariate correlations between the factors and the convergent and divergent variables were calculated. The ATI-S construct of affinity for technology interaction, which is assumed to be divergent, is found to be largely independent of the factors. The significant correlations with motivational user loyalty and learnability can be explained by considering that increased technology affinity has a positive effect on the continuous intention to learn and use a technology [18]. The perceived customizability correlates negatively with the results of the ATI-S, which can be explained by the fact that people with a high technology interaction affinity may not have perceived sufficient customization options on the robotic system. Convergent constructs of IsoNorm 9241/10 show mixed results. Although the factors are based on largely the same interaction principles, they do not

consistently correlate with each other. A possible explanation for this is that in the process of item reduction, the questionnaire developed focused on a different facet of the interaction principle. Thus, both the factors of the IsoNorm 9241/10 and this questionnaire would be assigned to the same interaction principle in terms of content, but without being the same in terms of items. As singular items represent specific functions, it is possible that a robotic system would score differently on two questionnaire if they contained different facets of the DIN EN ISO 9241-110. However, the results show that the factors do not correlate with their content wise counterpart with such consistency, that this explanation does not entirely suffice to explain these results. Once again underlying data quality could be a contributing factor.

To rule out a possible influence of previous experience with robotic systems both samples were compared on that facet. However, both groups had a similar distribution of experience, with the completely inexperienced being the dominant group.

The Cronbach's α of all factors in the CFA is low, including that of the questionnaires used for divergence and convergence. Almost all factors reduced in their Cronbach's α compared to the EFA results. While it is normal for a questionnaire to show variance in their Cronbach's α , e.g. ATI-S reports average values between .83-.92, factors like effort required to correct user errors dropped from .795 to .240. This is a reduction in internal consistency by a large margin for several of the factors. As the robotic system had not changed between the two assessments, comparable results were expected. The reduction in the Cronbach's α could suggest inconsistent perceptions of the robotic systems within the CFA sample, or irregularities in the underlying data.

Overall, the results of the EFA are indicating a clear factor structure. The significant contrast of the CFA results indicate the need for further studies to assess the quality of the questionnaire.

V. CONCLUSION

Varieties of factors contribute to how a human perceives interaction with a robot. The systems usability being one of them. Therefore, creating a survey instrument to assess the perceived usability of an interactive robotic system can help improve future interaction and robotic design. Established usability norms such as the DIN EN ISO 9241-110:2020-10 can function as a basis for such questionnaires. Survey instruments such as IsoMetrics [5] and IsoNorm 9241/10 [7] implement DIN EN ISO 9241-110 [4] in the form of questionnaires and thus enable users and experts to evaluate interactive systems. Currently no survey instrument specializes in usability of human robot interaction in the context of the norm.

The developed questionnaire shows an initial clear factor construct with high content validity, high inner-factorial consistency, with good explanation of variance. This result could not be replicated in a smaller sample and confirmatory factor analysis. Convergent and divergent constructs return a mixed picture in terms of validity, with limited explanation. How-

ever, the lack of correlation between content wise almost identical factors brings up the need to consider the underlying data quality of the CFA sample as a possible influence on the results. This is further supported by the wide span of Cronbach's α for all factors, both the new questionnaires and the IsoNorm 9241/10 and ATI-S. These low internal consistencies indicate a weak basis for the CFA, contributing to the insufficient model fit.

Nevertheless, the results of both surveys offer important indications of possible factorial structures that structurally deviate from DIN EN ISO 9241-110:2020-10, but are consistent with it in terms of content. The structure of the new principle of user engagement in particular was examined for the first time and forms two distinct factors. In addition, it can be observed that the factor of controllability splits into two separate constructs as well: controllability and customizability. The later was integrated into the prior in the newest form of the norm. The results however, suggest that customizability is a separate construct from controllability. Therefore, while the EFA did not replicate the exact factors of the DIN EN ISO 9241-110, it highlighted an underlying content focused structure.

In future work, the preliminary developed questionnaire of the EFA could be used in addition to other established survey instruments in order reassess the results of the CFA. This can help to further understand the structure of the questionnaire.

Future studies could also consider adjusting their method to better reflect a working situation. While video-based evaluation have many advantages, using the questionnaire after a person did directly interact with a robot could improve data quality.

Despite the mixed results, the interaction principles show promise in their capability to assess robotic systems for their usability in human-robot interaction. In addition to that, while reliable usability questionnaires exists, there are benefits in developing an instrument specialized for robotic systems, as it represents the capabilities and extend of the interaction more precisely.

ACKNOWLEDGMENT

This article was created as part of the SOPHIA project (which is covered by funding agreement no. 871237). This project is funded by the European Union's research and innovation program "Horizon 2020" (H2020-ICT-2019-2/ 2019-2023).

REFERENCES

- [1] International Federation of Robotics. "Executive Summary World Robotics 2021: Industrial Robots," 2021.
- [2] L. Onnasch, X. Maier and T. Jürgensohn, "Mensch-Roboter-Interaktion- Eine Taxonomie für alle Anwendungsfälle," Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. 2016.
- [3] L. Onnasch, and E. Roesler. "A taxonomy to structure and analyze human-robot interaction," in *International Journal of Social Robotics*, 13th ed. vol.4, 2021, pp.833-849

- [4] DIN EN ISO 9241-110:2020-10, "Ergonomie der Mensch-System-Interaktion - Teil 110: Interaktionsprinzipien (ISO 9241-110:2020); Deutsche Fassung EN ISO 9241-110:2020," Berlin: Beuth, 2020.
- [5] G. Gediga, K.C. Hamborg and I. Dünisch, "The IsoMetrics usability inventory: an operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems," in *Behaviour & Information technology*, 18th ed. vol.3, 1999, pp. 151-164.
- [6] P. Rosen, "Aufgabengestaltung in der Mensch-Roboter-Interaktion-Tätigkeitsspielräume bei teilautomatisierten Produktionsaufgaben," *Gesellschaft für Arbeitswissenschaft*. 64th. GfA-Frühjahrskongress, Februar. Frankfurt a. M, 2018.
- [7] J. Prümper, "Test It: IsoNorm 9241/10," in *HCI*, 1st ed. 1999, pp. 1028-1032.
- [8] J. Wessel, C. Attig, T. Franke, "ATI-S-An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies," in *Proc.of Mensch und Computer*, 2019, pp. 1028-1032.
- [9] IBM Corp. IBM SPSS Statistics for Windows, Version 27.0. Released 2020. Armonk, New York: IBM Corp
- [10] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. 2021. <https://www.R-project.org/>.
- [11] R Core Team, "foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase'," R package version 0.8-81. 2020. <https://CRAN.R-project.org/package=foreign>
- [12] A. Lishinski, "lavaanPlot: Path Diagrams for 'Lavaan' Models via 'DiagrammeR'," R package version 0.6.2. 2021. <https://CRAN.R-project.org/package=lavaanPlotM>.
- [13] Y. Rosseel, "lavaan: An R Package for Structural Equation Modeling," in *Journal of Statistical Software*, 48th ed. Vol. 2, 2012, pp. 1-36. <https://www.jstatsoft.org/v48/i02/>
- [14] H. Wickham and J. Bryan, "readxl: Read Excel Files." R package version 1.3.1. 2019.<https://CRAN.R-project.org/package=readxl>
- [15] J. Bortz, and N. Döring. "Quantitative Methoden der Datenerhebung." *Forschungsmethoden und Evaluation*. Springer, Berlin, Heidelberg, 2006. pp. 137-293.
- [16] H. Gatignon, „Reliability Alpha, Principle Component Analysis, and Exploratory Factor Analysis“. In *Statistical Analysis of Management Data*, Springer, New York, NY. 2010, pp. 29-57.
- [17] T. A. Brown, „Confirmatory factor analysis for applied research,“ Guilford publications. 2015.
- [18] I. Jin and M. Divitini, "Affinity for Technology and Teenagers' Learning Intentions," in *Proceedings of the 2020 ACM Conference on International Computing Education Research*, 2020, pp. 48-55.