An Android App for Posture Analysis Using OWAS

Christian Lins[1][a], Franziska Minh-Khai Dang Quang[2], Rica Schulze[3], Stefanie Lins[4], Andreas Hein[2][b] and Sebastian Fudickar[3][c]

[1]*Department of Computer Science, Hamburg University of Applied Sciences, 20099 Hamburg, Germany*
[2]*Assistance Systems and Medical Device Technology, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany*
[3]*MOVE Junior Research Group, Institute of Medical Informatics, University of Luebeck, Germany*
[4]*University of Hagen, 58097 Hagen, Germany*

Keywords: Smartphone App, Postural Assessment, Musculoskeletal Disorders, non-Neutral Postures, UEQ.

Abstract: In this paper the APA, an App for Posture Analysis, that incorporates the OWAS method (Ovako Working Posture Analysis System) for assessing potential harmful postures of physically hard working employees, is presented and evaluated. The app is intended as a tool for occupational safety experts in assessing the individual postures of workers to identify and prevent harmful working situations in regards to musculoskeletal hazards. For this, APA incorporates a digitized assessment sheet for the OWAS method and timing support that helps occupational safety experts structure and simplify the assessment workflow. To investigate whether the protocol sheet can be replaced with the app, a study was conducted in which the inter-rater reliability among the app and the paper protocol sheet (as control group) was evaluated. In addition, the usability of the app was determined via the User Experience Questionnaire (UEQ). In the study, the app achieved higher inter-rater reliability for watching a video recording of postures than the control group. The chi-square test revealed differences in the use of the app and paper only for leg postures. The UEQ indicated overall above average results, which indicates a sufficient usability, which was also confirmed by free-textual comments of the 13 study participants. The results suggest that it would be possible to replace the paper protocol sheet in the future with an app.

1 INTRODUCTION

Musculoskeletal disorders are a major occupational health problem, especially in industrialized countries. Since poor working postures – repeated often or over a long period of time – are a major risk for musculoskeletal disorders (Hoy et al., 2010; Amell and Kumar, 2001; Karwowski and Marras, 1998; Matsui et al., 1997), the postures of employees during work shifts should be observed regularly to detect and prevent posture damage early. It is important to identify patterns of work-related musculoskeletal disorders, symptoms, and their risk factors early in the workplace. In doing so, appropriate work postures not only counteract the problems, but also result in greater control of work performance and a reduction in the number of workplace accidents. Individual worker strain and corresponding risks are typically assessed by human observers via questionnaire-based assessment methods such as the Ovako Working Posture Analysis System (OWAS) (Karhu et al., 1981). It is a practical method for identifying and assessing poor working postures. OWAS is known to be easy-to-use and reliable in terms of high inter-rater reliability (Lins et al., 2021). In OWAS a human observer rates the posture of an individual in intervals (e.g. every few minutes) using three categories: arms, back, and legs. For each of these categories 3 to 7 different part-postures can be coded (arms: code 1-3, back: code 1-4, legs: code 1-7). The codes obtained can be used to determine both individual holding combinations and their proportion of the total time. In both cases, values are classified into four action classes (from 1: posture not harmful to 4: posture very harmful) indicating the severity of the posture and the corresponding need of improvement. These action classes for postu-

[a] https://orcid.org/0000-0003-3714-0069
[b] https://orcid.org/0000-0001-8846-2282
[c] https://orcid.org/0000-0002-3553-5131

Lins, C., Quang, F., Schulze, R., Lins, S., Hein, A. and Fudickar, S.
An Android App for Posture Analysis Using OWAS.
DOI: 10.5220/0011668300003414
In *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023) - Volume 5: HEALTHINF*, pages 307-313
ISBN: 978-989-758-631-6; ISSN: 2184-4305

ral combination are specified in accordance with risk assessments for musculoskeletal disorders. If a code is assigned action class 4, the posture should be corrected immediately (Karwowski and Marras, 1998). The second option gives an action class based on the relationship of the posture of the back, legs, and arms for an observation period. This refers to the partial codes and their proportion of time. The highest action class of the 4 partial codes indicates how dangerous the postures were for the worker. In this regard, the method should be easy enough to use so that untrained people can apply it. It should provide unambiguous answers, even if it leads to overconfirmation. The OWAS method fulfills these criteria. See (Lins et al., 2021; Lins and Hein, 2022) for a detailed discussion of OWAS.

A digitized assessment holds advantages over paper-based approaches: the time interval of the observation can be controlled, the assessment results can be easily processed and summarized, and potential coding errors can be avoided. Additionally, it might be possible to assess several subjects at once by carefully timing the different observations, a task that is very error-prone if using paper sheets. In some countries such as Germany companies might be required to legally document the workplace assessment, which can be automated using digital tools. This is possibly a useful intermediate step to the assessment with smart workwear, e.g. (Lins and Hein, 2022). As a result, we propose APA, an Android app which guides occupational health experts and untrained personnel through an OWAS observation session. This app is intended to replace the paper-based questionnaire and to save time both in entering the information and in evaluating it. In addition, it is evaluated whether the inter-rater reliability of the app is comparable to the one of the paper-based version. Furthermore, the user experience of untrained observers using the app is investigated. To summarize the contents of this paper:

- A newly developed app is presented as a tool for classifying postures using the OWAS method.
- The app is tested with users to determine inter-rater reliability of the OWAS method when using the app.
- The observation results are compared to the usual pen-and-paper application of OWAS.
- The User Experience Questionaire (UEQ) is used to examine the design of the app for strengths and weaknesses in use.

The remainder of the paper is organized as follows: first, the app and the evaluation study and its analysis are described (Section 2), followed by a presentation of the results in Section 3. A discussion of the results is provided in Section 4, and the paper concludes with Section 5.

2 METHODS

The APA App and the study design and the statistical analysis are presented in the following subsections.

2.1 Android App for Posture Analysis (APA)

APA's main aim is to enable its users (occupational safety experts and physiotherapists as well as laypersons) to assess the posture of the observed workers via the OWAS method. The app should support the user to record and document postures while observed workers move in work environments. APA was primarily developed for a 1:1 observation, i.e. one observer observes only one worker, although a future extension to more than one worker is possible. The APA app (see Figure 1) was developed via a two-step iterative human-centered design process. Initially, a click prototype was created and was discussed with an occupational safety expert resulting in a simplification of the user interface and a focus on a one-by-one assessment instead of the originally intended support of multiple parallel observations. In a second iteration the user interface (UI) was implemented applying the *8 Golden Rules* of Shneiderman (Shneiderman, 2002). Herein, we evaluate the resulting version of the second design iteration, integrating already the insights of these expert interviews. Once started, APA asks for the observer name and the observed worker's name and workplace (see Figure 1b). When all entries have been made, a new observation session can be started. The observation interval is set to 30 seconds in accordance with (Brandl et al., 2017) in order to leave sufficient room for the risk assessment, while collecting sufficient samples for an OWAS classification. Per observation for this session, the user interface first opens with input options for the body posture, which can be seen in Figure 1c). The observer is expected to enter individual postures for back, arms, and legs as well as the weight load. These values are initially set to code 1 for the first observation and then always set to the last used partial code. Hereby, observers can use shortcuts for subsequent session-observations, in case the posture does not change.

Once all codes are set correctly, the user can confirm the entry and then enter a waiting screen. Here the user can end the session or move on to the next observation. A special feature is the timer, which provides an acoustic signal after 30 seconds to remind the observer to continue with the observations. In addition, the timer time can be tracked visually with a bar that fills up.

The APA provides basic information about the specific OWAS postures used for assessment (Figure 1a) and allows the export of the assessment results (Figure 1d). The app implementation is based on Android, Java, SQLite database, the Room and the Hilt frameworks for depencency injection and the MPAndroidChart API for results visualization. Test-driven development is implemented via JUnit and Mockito.

2.2 Study Design

To evaluate the APA app and compare it with the standard paper assessment sheet, a study was conducted. Inclusion criteria for participants was the availability and experience in using an Android smartphone with a supported Android version (5 to 10). Due to COVID-19 restrictions, the study was conducted in one on one meetings via video calls. Participants were sent a file with a brief overview of the OWAS posture categories and general participant information in advance. Initially, participants were introduced to the study aims, the general study procedure and were shown two introductory videos. The groups were introduced to the protocol-guideline and the groups specific methodology (app or paper). Concluding open questions have been answered. Within the main study, participants were shown a video of a man adopting work related human postures. The man models postures evenly over all OWAS codes. All participants were shown the same videos and were requested to rate the posture every 30 seconds according to OWAS.

The study was approved by the Institutional Review Board of the University of Oldenburg (Drs. 24/2017).

2.3 Statistical Analysis

2.3.1 Inter-Rater Reliability

Since no direct comparison with a ground truth is possible with the OWAS observation method, the test-retest reliability with multiple observers (inter-rater reliability) is determined between the two logging methods (app and paper). Inter-rater reliability indicates the degree of agreement when a subject is assessed by multiple observers. It is usually expressed by a kappa value ($\kappa < 0$: strong disagreement to $\kappa = 1.0$: perfect agreement). The calculation is made according to Fleiss (Fleiss, 1971).

2.3.2 Test for Independence

The observations (dependent variable) of the two logging groups are tested for their independence from the respective logging method (app or paper). For this purpose, the χ^2-test for independence is applied. The null hypothesis H_0 is: the observations are stochastically independent of the logging method. The alternative hypothesis H_1: the observations are stochastically dependent on the logging method.

2.4 Usability Test

The User Experience Questionnaire (UEQ) is used to evaluate the usability of the app. The UEQ was also supplemented with free text fields to collect further impressions of the participants about the app. To compare the quality of the user experience of the app, the UEQ benchmark is used, which contains the aggregated data from various evaluation studies of diverse products. Through the benchmark, the relative quality of the app compared to other products is estimated. (Laugwitz et al., 2008)

3 RESULTS

3.1 Study Cohort

13 participants with no prior experience in coding postures were randomized into two groups: The intervention group using the APA ($n = 7$) and the control-group using the established paper-protocol ($n = 6$). The data of one rater from the APA group were excluded from the evaluation due to technical problems. Two raters from the pen-and-paper group were also excluded, because they completely misunderstood the OWAS method.

3.2 Statistical Analysis

Table 1 summarizes the inter-rater reliabilities for the APA use and the control condition per body-part.

Table 1: Inter-rater reliability values (per body-part category) for both logging method groups.

		Legs	Arms	Back
APA logging	$\kappa =$	0.448	1.000	0.383
Paper logging	$\kappa =$	0.417	0.657	0.249

Table 2 shows the χ^2-test values (number of ratings $N = 191$). According to the test values the H_1 hypothesis can be provisionally accepted for the legs category (stochastical dependence between logging

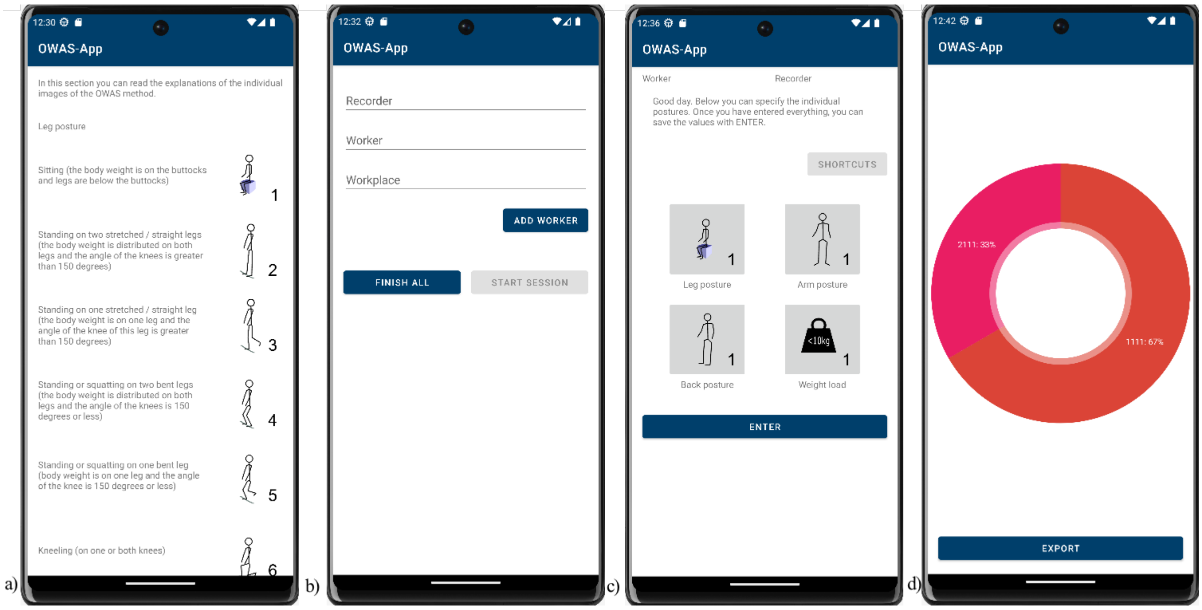

Figure 1: User Interface (UI) for Android of the APA App. The UI can be as well experienced as a screencast video (see https://youtu.be/BtpdeiYyRGU).

method and rating results), for both arms and back the H_0 must be kept.

Table 2: χ^2-test values (per body-part category) for both logging method groups ($N = 191$), *: H_1 provisionally accepted ($V = 0.342$).

	Legs	Arms	Back
χ^2	22.349*	0.448	4.762
p	< 0.001*	0.799	0.190

3.3 User Experience Questionnaire

Figure 2 summarizes the averaged answers to the individual questions of the questionnaire. The sub-items can be summarized based on coloring into the six main scales covered by the UEQ: Attractiveness can be used to determine whether the participant likes the product or not. Transparent indicates how easy it is for the user to become familiar with the product. Efficiency describes whether the tasks can be completed without unnecessary effort. Controllability indicates whether the user has the impression of being in control at all times. Stimulation asks whether the user feels motivated to use the product and originality indicates how interesting the product is and whether it arouses the interest of the user. The scale ranges from -3 ”horribly bad” to +3 ”extremely good”. The range of values between -0.8 and 0.8 encompasses a neutral rating, values >0.8 represent a positive rating, and values <-0.8 represent a negative rating. (Schrepp et al., 2017)

Figure 3 indicates the ”user experience” quality of the product. In this diagram, above average indicates that 25% of the results were better than the app, and 50% were worse. Accordingly, good means that 10% were better and 75% were worse (Schrepp et al., 2017). The results show that the app achieves good results for Perspicuity, Efficiency, Dependability and it achieves above average results in all remaining ones.

Considering the additional questionnaire-items provided room for further free-text comments:

- The app users indicated being comfortable with the RGB Videos.
- While many participants described the app as appealing and praised individual elements such as the timer and the intuitive user interface, one participant stated that she required an introduction to use the app. In addition, one user stated that users do not need the shortcuts, as users found the search in them slower compared to the already very fast normal input option.

4 DISCUSSION

Inter-Rater Reliability. The κ-values summarized in Table 1 show an overall increased inter-rater reliability for the app compared to the pen-and-paper logging group. The app-raters achieved a significantly increased reliability for rating the arms and the back

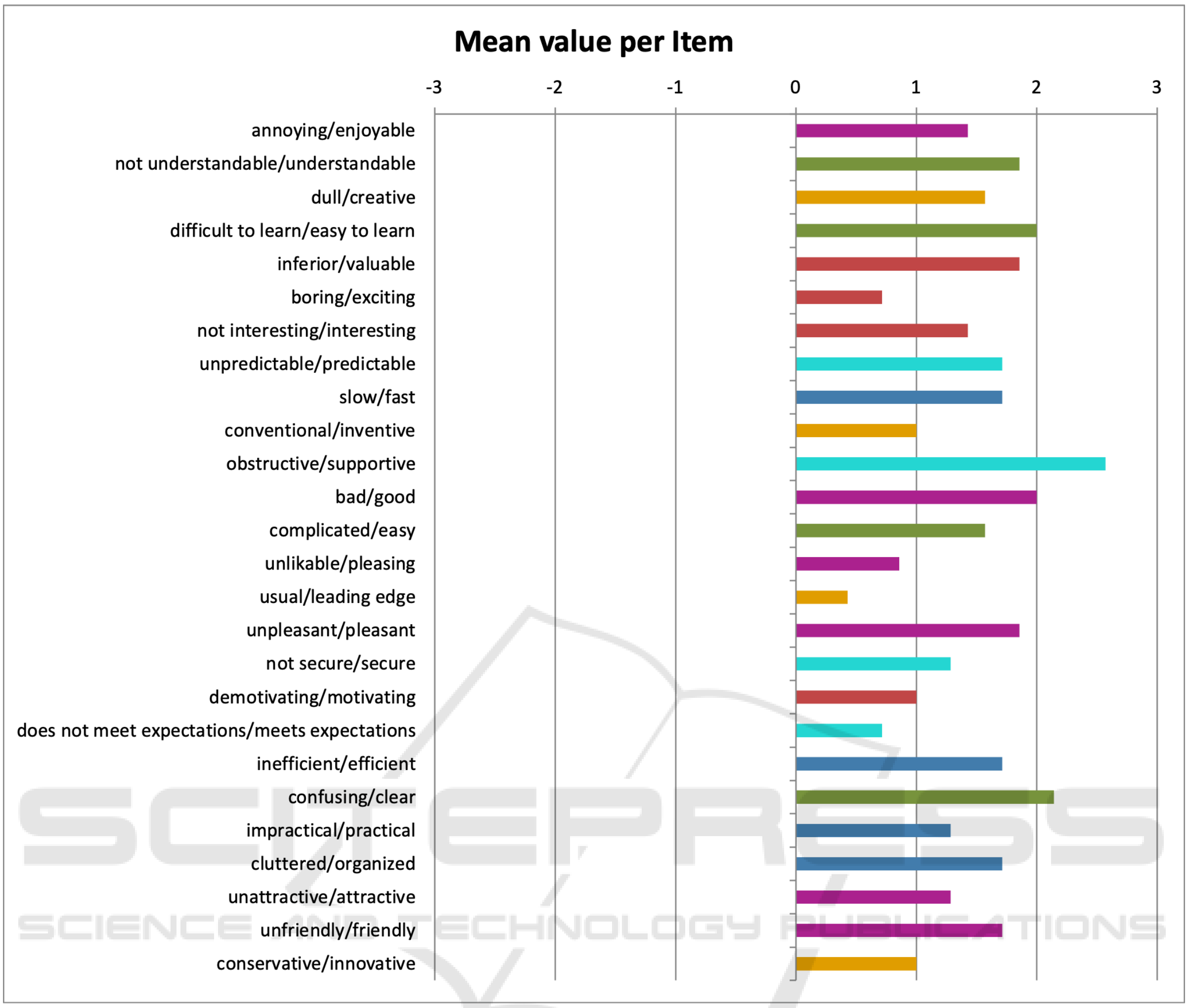

Figure 2: Mean answers to the individual questions of the standardized Usability Questionnaire UEQ (each color is assigned to an already given range: Attractiveness: purple; Perspicuity: green; Novelty: orange; Stimulation: red; Dependability: azure; Efficiency: dark blue);.

postures, with a perfect reliability for the arms when using the APA app.

Test for Independence. The χ^2 test for independence revealed no dependence of the observational results on the logging method, at least for the arms and back categories. For the legs, the test shows a dependence of the results on the logging method (medium effect size according to (Cohen, 2013)). This could be due to the small number of participants or the larger selection of legs poses, but perhaps also due to the interface design of the app. Smaller devices may require scrolling the interfaces, so perhaps observers were more inclined to connect to the icons visible from the beginning. This potential source of error should be further investigated in future studies (e.g., with professionals) and eliminated if necessary.

Overall, the results of the arms, legs, and back categories as a whole are inconclusive. The different results are probably due to the larger number of posture classes in the legs category (7 vs. 3 or 4). A larger number of subjects is necessary for a clearer conclusion.

Usability of App. Based on the data obtained from the UEQ questionnaire, it is possible to determine which areas need improvement. For the app, it would be the areas of Originality and Stimulation, which are not conspicuously bad with 1.0 and 1.25, but lower compared to the other areas. However, the following can be said about the originality area: The app is not designed to be novel or innovative. Moreover, the use case is very specific, while the participants came from fields where the OWAS method was not known.

Table 3: Results of the User Experience Questionnaire for the App user group.

UEQ Scales	Mean (SD)
Attractiveness (Overall impression of the product. Do users like or dislike it?)	1.524 (0.33)
Perspicuity (Is it easy to get familiar with the product and to learn how to use it?)	1.893 (0.60)
Efficiency (Can users solve their tasks without unnecessary effort? Does it react fast?)	1.607 (0.77)
Dependability (Does the user feel in control of the interaction? Is it secure and predictable?)	1.571 (0.66)
Stimulation (Is it exciting and motivating to use the product? Is it fun to use?)	1.250 (0.54)
Novelty (Is the design of the product creative? Does it catch the interest of users?)	1.0 (0.52)

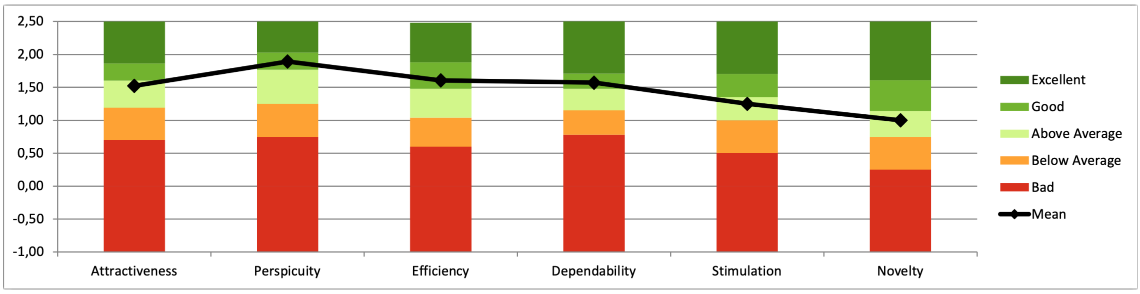

Figure 3: UEQ - Benchmark-Diagram for the App.

None of the participants had to deal with body posture analysis before. Therefore, it can be assumed that the paper protocol sheet was not known and therefore no comparison could be made. Thus, it is not surprising that the app scored lower on originality. As described above, the stimulation area expresses how high the motivation is to use the product. This score could have been better as well. However, considering that the test subjects had to watch a video with repeated 30-second pauses in which they had to wait, the rather low value is not surprising. Moreover, the APA app is designed for a work environment in which the observers would be busy alongside the observations.

General Considerations. In further studies the results should be confirmed in a work environment, to finally confirm the applicability of the APA app. In addition, it might be worthwhile to investigate which effect the level of training brings to the inter-rater reliability, as some of the common errors seem to be well addressed by the app. In order to support such research questions and enable other groups to benefit from the APA app, we decided to make the app available as open source software. Thus, the source code of the app can be found at https://github.com/fudickar-lab/owas-app.

5 CONCLUSION

We presented an app as a support tool for posture analysis using the OWAS method. The app was tested for reliability with human subjects and the usability of the app was evaluated. The results show that the app approaches the usual pen-and-paper logging method in terms of reliability and has higher inter-rater reliability, at least in our cohort. It is suggested to consider digital tools such as apps more for the analysis of movements and postures. In the future, the app could be extended to include other assessment methods apart from OWAS (e.g. REBA, RULA, etc.), and in addition, different observation strategies could be implemented and evaluated, e.g., with different sampling frequencies. It may also be worthwile to integrated a semi-automatic classification via the integrated camera of the smartphone.

ACKNOWLEDGEMENTS

The author(s) would like to thank all study participants for their contribution.

REFERENCES

Amell, T. and Kumar, S. (2001). Work-related musculoskeletal disorders: design as a prevention strategy. A review. *Journal of occupational rehabilitation*, 11(4):255–265.

Brandl, C., Mertens, A., and Schlick, C. M. (2017). Effect of sampling interval on the reliability of ergonomic analysis using the Ovako working posture analysing system (OWAS). *International Journal of Industrial Ergonomics*, 57:68–73.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.

Hoy, D., Brooks, P., Blyth, F., and Buchbinder, R. (2010). The Epidemiology of low back pain. *Best Practice and Research: Clinical Rheumatology*, 24(6):769–781.

Karhu, O., Härkönen, R., Sorvali, P., and Vepsäläinen, P. (1981). Observing working postures in industry: Examples of OWAS application. *Applied Ergonomics*, 12(1):13–17.

Karwowski, W. and Marras, W. S. (1998). *The Occupational Ergonomics Handbook -*. CRC Press, Boca Raton, Fla.

Laugwitz, B., Held, T., and Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *Lecture Notes in Computer Science*, pages 63–76. Springer Berlin Heidelberg.

Lins, C., Fudickar, S., and Hein, A. (2021). OWAS inter-rater reliability. *Applied Ergonomics*, 93:103357.

Lins, C. and Hein, A. (2022). Classification of body postures using smart workwear. *BMC Musculoskeletal Disorders*, 23(921).

Matsui, H., Maeda, A., Tsuji, H., and Naruse, Y. (1997). Risk indicators of low back pain among workers in Japan. Association of familial and physical factors with low back pain. *Spine*, 22(11):1242–7; discussion 1248.

Schrepp, M., Hinderks, A., and Thomaschewski, J. (2017). Construction of a benchmark for the user experience questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(4):40.

Shneiderman, B. (2002). *User Interface Design. - Effektive Interaktion zwischen Mensch und Maschine. Leitfaden für intelligentes Schnittstellendesign. Was Programmierer und Designer über den Anwender wissen müssen.* mitp-Verlag, Frechen.