UNIVERSITY OF LEEDS

This is a repository copy of *Eliciting Ergonomic User-Defined Gestures for Virtual Reality: A Pilot Study*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/217394/

Version: Accepted Version

Article:

Sanders, N., Sener, E. orcid.org/0000-0001-5159-0870 and Chen, K.B. (2023) Eliciting Ergonomic User-Defined Gestures for Virtual Reality: A Pilot Study. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 67 (1). pp. 1510-1515. ISSN 1071-1813

https://doi.org/10.1177/21695067231192530

© 2023 Human Factors and Ergonomics Society. This is an author produced version of an article published in Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Eliciting Ergonomic User-Defined Gestures for Virtual Reality: A Pilot Study

Nathan Sanders[1], Elif Şener[2], and Karen B. Chen[1]
[1]Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University
[2]Department of Graphic and Experience Design, North Carolina State University

People are increasingly using virtual reality (VR) for work. As a result of extended use, fatigue and musculoskeletal disorders affecting the upper arms and shoulders are already becoming common among VR users. This pilot study presented a "virtual working area" (VWA) to reduce the risk of fatigue resulting from using gestures obtained in gesture elicitation studies, and explored how the distance to the user interface (UI) interacted with different functions (select, scroll) during a mock reading task. Results showed that keeping the hands within the VWA had the potential to reduce Rapid Upper-Body Limb Assessment (RULA) and Borg CR10 scores at clinically significant levels. Scores were worse when the UI was far away and for the select function, suggesting the design of virtual UIs can play a role in eliciting naturalistic yet ergonomic interactions. The results also provide effect sizes and variance estimates to plan future work.

Introduction

It is anticipated that more and more people will use Virtual Reality (VR) to perform day-to-day tasks and work-related activities beyond entertainment and social interactions (Stevens et al., 2003). One task that is ubiquitous in our daily lives is reading. Not much is known about the long-term physical effects of reading in VR every day over long periods, but in other domains VR-specific muskuloskeletal disorderse (MSDs) are already starting to manifest. Gorilla-arm syndrome, for example, is characterized by chronic shoulder pain caused by prolonged extension of the arms without support (Boring et al., 2009). It seems likely that the risk of fatigue and subsequent MSDs will only be exacerbated by increased VR use, and will also be a hindrance to the adoption of this evolving technology.

Although the use of controllers is an option in the current state of VR, gesture recognition systems are gaining popularity as a more naturalistic interaction method (Bowman et al., 2012). Gesture elicitation studies (GESs) have also gained popularity because the resulting user-generated gestures may be more reflective of end users' behavior and preferences (Villarreal-Narvaez et al., 2020). However, the problem of fatigue remains because GES are prone to performance bias. Performance bias refers to the fact that fatigue and physical discomfort from postural stress, static exertions, and repetitive motions tend to occur after extended periods of time, and therefore typically do not have time to manifest during GES studies (Ruiz & Vogel, 2015; Uva et al., 2019). Thus, in the minds of participants there is not much fatigue to minimize, and so the resulting gestures reflect high-performance in a short-term laboratory environment. As a result, the gestures elicited in a typical GES may cause fatigue if used for long periods.

During a long-term reading activity, fatigue can be a function of the organization of the user interfaces (UIs) and the task users trying to accomplish (Bowman & McMahan, 2007). One example of UI organization is the distance to the UI. UIs within arm's reach can be easily manipulated via natural sensorimotor contingencies (i.e. selection by finger collision) (Bowman & McMahan, 2007; Jerald, 2016). By contrast, UIs that are placed further away may require symbolic or more complex gestures which demand a range of movements (e.g. ray-gun metaphors) where users may be forced to extend and stabilize their arms in order to make precise selections, leading to fatigue (Boring et al., 2009).

VR applications can produce highly complex interaction scenarios depending on the user's task (H. Wu et al., 2019). In GESs, gestures are associated with *referents* or *functions*. Functions can be grouped into two classes: Canonical manipulation functions (e.g. select, translate, scale) and abstract functions (e.g. scroll, highlight, copy) (A. S. Williams et al., 2020). Canonical manipulation functions, as the name suggests, may elicit gestures which rely on direct manipulations of UI elements, while abstract metaphors may elicit gestures which draw on previous experience with other technologies. This in turn may affect the position of the hands and arms when the gestures are executed, thus affecting fatigue.

Although there are many ways of measuring physical exertion in the occupational biomechanics repertoire, only a small subset of these have been applied to GESs. These include subjective psycho-physical ratings such as the Borg CR10 (N. Williams, 2017) and objective kinematics-based scoring systems such as the Rapid Upper Limb Assessment

(RULA)(McAtamney & Corlett, 2004). The Borg CR10 and RULA, assessed after a short number of gesture repetitions, have both been shown to be useful for predicting discomfort and fatigue resulting from actual prolonged and repetitive gesture use (Son et al., 2017). Strategies for mitigating performance bias include the use of covert kinaesthetic priming (Hoff et al., 2016) and soft constraints such as attaching weights to participants' wrists (Ruiz & Vogel, 2015). There is an idea from physical workstation design that to our knowledge has not yet been applied to gesture elicitation, but which may prove effective in mitigating performance bias – the idea of a *normal working area*. It comprises the intersection of a horizontal plane, such as a worktable or bench, with a zone described by "a comfortable sweeping movement of the upper limb, about the shoulder, with the elbow flexed 90 degrees or a little less" (Pheasant & Haselgrave, 2006). The present study extends this concept from the surface of a physical workbench to a 3D zone in VR.

This study presents a Virtual Working Area (VWA) as an easy-to-use constraint for GESs which encourages participants to keep their arms in an ergonomically sound position as they produce gestures. The primary goal is to evaluate the viability of the VWA as a tool to reduce performance bias in gesture production, and to reduce the risk of fatigue associated with the resulting gestures should they be used repetitively during long-term VR use. We took a user-centered and mixed-methods approach, using objective measures of upper body posture, subjective measures of exertion, and qualitative data from think-alouds and interviews. The secondary goal was to explore how UI Distance interacted with different Functions, to better understand how naturalistic yet ergonomic interactions can be supported through VR design considerations. As a pilot study, the overarching goal was to inform the design of a future experiment by providing descriptive statistics and error estimates with regard to minimal clinically important differences (MCIDs) using 80% confidence intervals (Lee et al., 2014).

Methods

Participants

Thirteen (13) participants (*M* = 27.2 years, *SD* = 4.9 years, 7 female) were recruited by distributing flyers at the university campus. They were required to have functional hands and arms, normal or corrected-to-normal vision, and the ability to speak. Nine people (69.2%) had used VR at least once before; one of these (7.7%) reported using VR "often"; the other four (30.8%) had never used VR. The study was approved by the university IRB.

Experimental design and procedure

The physical environment was a place where VR might realistically be used for intellectual work – a library office. Participants were seated in an office chair while they wore a head-mounted display (HMD) and entered a virtual workspace. Its (non-interactive) UI comprised three screens, each containing *lorem ipsum* text. The hypothetical scenario was reading, where someone might execute different gesture-based functions to select a virtual window and scroll through the text it contains.

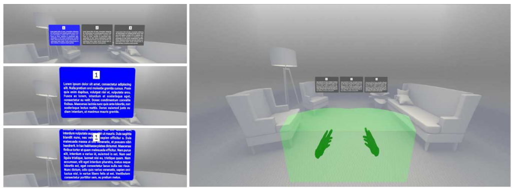

Figure 1. *Focus-Select (top left, middle) and Scroll (left middle, bottom) and a first-person view of the environment with the VWA (right)*

The experiment was a 2x2x2 randomized incomplete block design. The factorial design space comprised two levels of Virtual Working Area (off, on); two levels of UI Distance (50 cm, 2 m); and two levels of Function (focus-select, scroll). The design was blocked by participant, with each experiencing three randomized treatment combinations. Four gestures were elicited in each treatment combination, totalling twelve observations per participant. Upper limb kinematics, a RULA score, and a Borg CR10 rating were recorded in each observation. Linear mixed-effect models (LMMs) were fit using JMP Pro 16 and lme4 1.1-31 in R 4.2.0. The models included the maximal random effects structure justified by the design: by-subject intercepts and by-subject slopes for UI Distance.

The procedure was as follows. Participants were briefed, then fitted with the HMD, and introduced to the gesture elicitation process via a scripted demonstration by the experimenters. Audio and visual recording began, and the gesture elicitation commenced. For each treatment combination, the Function was presented as a looping non-interactive animation. Participants were prompted to name each gesture they produced and to think aloud as they developed it. If the VWA was on, they were reminded to keep their hands inside. Once a gesture was finalized, they performed it twenty times at one repetition per second, then provided a Borg CR10 rating. After they had elicited four gestures, they took a two-minute break to rest their arms. This process was repeated for the two remaining treatment combinations. Participants were allowed to re-use or adapt gestures for the same Function in different Distance or VWA factor levels (see Figure 2). There was a semi-structured interview after the runs were complete. The entire process lasted about one hour.

Figure 2. *A participant eliciting variants of the same gesture (named "point and click") with the VWA off (left) and on (right).*

Independent Variables

Virtual Working Area. This was a nominal treatment factor with two levels: Off and On. The geometry of the VWA is fully described in Pheasant and Haselgrave (2006). Its native two-dimensional shape was extruded 15 cm above and below a horizontal plane at the 50%ile seated elbow height (Pheasant & Haselgrave, 2006), appearing as a translucent green bubble aligned flush against the sternum (Figure 1). Hand tracking was used so that participants could see whether their hands were inside the VWA, and to aid in its initial positioning and alignment.

UI Distance. This was a nominal treatment factor with two levels: Near (0.5 m) and Far (2.0 m). The UI windows were centered on a line angled 6 degrees down from the approximate location of the participant's head. The far panels were larger so they and their text appeared the same angular size despite the increase in distance.

Function. This was a nominal treatment factor with two levels: Focus-Select and Scroll. Focus-select was a canonical manipulation function signified by a color-change in the selected window, followed by a button-press effect and the disappearance of the other windows as the selected window moved to the foreground-center. Scroll was an abstract function signified by text moving upwards (scrolling down) and stopping with a short bounce (Figure 1).

Dependent Variables

Upper-Limb Kinematics. The locations of bony landmarks were estimated from video recordings taken in the sagittal plane. The positions of T8 and C7 determined the line of the thorax (G. Wu et al., 2005). The acromion and lateral epicondyle of the humerus determined the line of the upper arm (G. Wu et al., 2005). Upper arm flexion was defined as the angle between the upper arm and thorax. The lateral epicondyle and radial/ulnar styloids were used to determine the line of the lower arm (G. Wu et al., 2005). Lower arm flexion was defined as the angle between the lower arm and the line of the humerus extending out of the elbow. Head/Neck flexion was the angle between true horizontal and a line from the tragus to a mark on the side of the HMD approximating the Frankfurt line (Ankrum & Nemeth, 2000).

RULA Scores. The measurements described above, along with rough estimates of wrist motion and shoulder abduction, were used to compute Rapid Upper Limb Assessment (RULA) scores (McAtamney & Corlett, 2004). Scores ranged from 1 ("posture acceptable") to 7 ("changes required immediately"). Because the difference between "changes may be required" (3) and "acceptable" (2) is 1 point, the MCID was taken to be 1.

Borg CR10 Scores. Participants provided a rating of 1 ("No exertion at all") to 10 ("Maximal") for each gesture they invented after performing the gesture for twenty repetitions. They were provided with a reference sheet in VR which showed the numerical scores and the semantic anchors. The MCID was taken to be 1 point (Ries, 2005).

Qualitative Data. Qualitative (verbal) data was acquired during the gesture elicitation process in a think-aloud and after the session in a semi-structured interview. Memos were recorded after each participant to document emerging impressions that may form the basis of themes in a future analysis (Nowell et al., 2017). These memos formed the basis of the interpretation of the results.

Materials and Apparatus

The virtual environment and animations were created with Unity (2019.4), Blender (3.4.1), and Fusion 360 (2.0.15509) and rendered on a Meta Quest 2 with a Dell Precision 5810 workstation and an NVIDIA GeForce RTX 3080 Ti. Participants were seated in a plastic chair without armrests with a seat height of 43 cm. They were filmed using a GoPro HERO10 Black positioned on a level tripod.

Results

Upper-Limb Kinematics

Results of a MANOVA of upper arm flexion, lower arm flexion, and head/neck flexion yielded a significant main effect of VWA ($F_{3,134}$ = 172.2, p < 0.01), with follow-up univariate LMMs showing a decrease in upper arm flexion, an increase in lower arm flexion, and an increase in head/neck flexion when the VWA was used. There were also main effects of UI Distance ($F_{3,134}$ = 1.69, p = 0.17) and Function ($F_{3,134}$ = 8.92, p < 0.01), with greater upper arm flexion and less lower arm flexion for far UI distances and Focus-Select, respectively.

The UI Distance x Function interaction was also significant ($F_{3,134}$ = 2.29, p = 0.08), with less upper-arm flexion and greater lower-arm flexion for Scroll, but only when the UI was near. There was also a UI Distance x Function x NWA interaction ($F_{3,134}$ = 4.02, p < 0.01), indicating that the above interaction only occurred when the VWA was off (Figure 3).

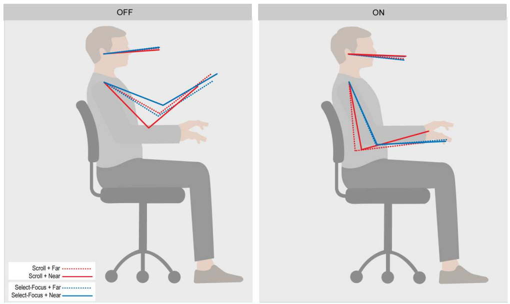

Figure 3. *Effects of VWA, UI Distance, and Function on upper limb kinematics. The T8-C7 line is drawn vertically.*

RULA Scores

Results of a LMM (Figure 4) yielded a main effect of VWA ($F_{1,57.8} = 68.54$, $p < 0.01$), showing that the average RULA score was lower when the VWA was on.

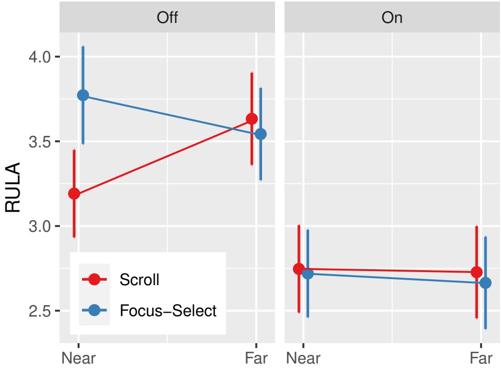

Figure 4. *Effects of VWA, UI Distance, and Function on RULA scores. Means are shown with 80% confidence intervals.*

Table 1

Means and REML variance components for RULA scores. Levels not connected by the same letter are significantly different.

Dist	Func.	VWA	Mean	80% CI		Sig. Diff.
Near	Scroll	Off	3.19	3.01	3.38	A
Far	Scroll	Off	3.63	3.44	3.83	B
Near	F.-Sel.	Off	3.77	3.56	4.98	B
Far	F.-Sel.	Off	3.54	3.35	3.74	B
Near	Scroll	On	2.75	2.56	2.93	C
Far	Scroll	On	2.73	2.53	2.92	C
Near	F.-Sel.	On	2.72	2.53	2.91	C
Far	F.-Sel.	On	2.66	2.47	2.86	C

Error Component	Var	80% CI		% Tot.
Subject	0.03	-0.01	0.06	8.44
Dist|Subject	0.00	-0.03	0.03	0.00
Residual	0.29	0.25	0.34	91.56

Its interaction with Function was also significant ($F_{1,66.58} = 1.97$, $p = 0.16$), as was the UI Distance x Function interaction ($F_{1,55.7} = 3.18$, $p = 0.08$), as well as the three-way interaction ($F_{1,59.58} = 2.42$, $p = 0.13$), suggesting that the average RULA score increased with UI distance for Scroll, but only when the VWA was off (Figure 4, Table 1).

Borg CR10 Scores

Results of a LMM (Figure 5) yielded a main effect for VWA ($F_{1,111.2} = 12.90$, $p < 0.01$). The average score was lower when the VWA was used. Its interaction with Distance was also significant ($F_{1,132.2} = 6.33$, $p = 0.01$), so that the average score increased with UI distance but only when the VWA was off. The main effect of Function was also significant ($F_{1,134.5} = 16.4$, $p < 0.01$); the average score was higher for Focus-Select compared to Scroll. Its interaction with Distance was also significant ($F_{1,108.9} = 5.91$, $p = 0.02$); the average score increased with UI distance for Focus-Select but decreased with distance for Scroll (Figure 5, Table 2).

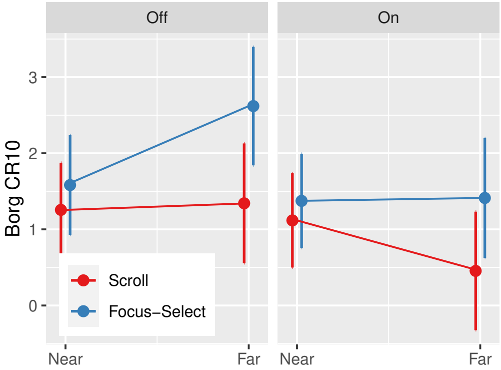

Figure 5. *Effects of VWA, UI Distance, and Function on Borg CR10 scores. Means are shown with 80% confidence intervals.*

Table 2

Means and REML variance components for Borg CR10 scores. Levels not connected by the same letter are significantly different.

Dist.	Func.	VWA	Mean	80% CI		Sig. Diff.
Near	Scroll	Off	1.26	0.83	1.68	A
Far	Scroll	Off	1.34	0.81	1.88	A
Near	F.-Sel.	Off	1.58	1.14	2.03	A
Far	F.-Sel.	Off	2.62	2.09	3.15	B
Near	Scroll	On	1.12	0.69	1.54	A
Far	Scroll	On	0.46	-0.07	0.98	C
Near	F.-Sel.	On	1.37	0.95	1.80	A
Far	F.-Sel.	On	1.41	0.88	1.95	A

Error Component	Var.	80% CI		% Tot.
Subject	0.73	0.22	1.25	45.15
Dist|Subject	0.29	0.06	0.53	17.20
Residual	0.60	0.51	0.71	36.85

Discussion

The first goal of the study was to evaluate the viability of a "Virtual Working Area" (VWA) to reduce the risk of long-term fatigue of gestures. Results showed that performing gestures within the VWA reduced both RULA and Borg scores at clinically significant levels (> 1 point), thus reducing the risk of long-term fatigue. The greatest improvements occurred for both functions at far UI distances, and for Focus-Select at near distances. Some people liked the VWA because it improved posture and economy of motion. But despite the benefits, the VWA produced several unforeseen consequences which make it non-viable in its current form, chiefly an increase in neck flexion. Participants stated that they looked down to get visual feedback of the hands for several reasons: to make sure that their hands were in the VWA, that the hands were being motion-tracked correctly, because sometimes the gesture metaphor required it (e.g. using their hand as a screen), and to establish a line of sight from their hand to the UI (when the VWA was off). Some participants did not like the VWA because they thought it was less ergonomic and too restrictive, both physically and creatively. It demanded greater attention to keep their hands in the VWA, and it induced greater wrist and finger motion to compensate for the lack of shoulder and elbow freedom as participants strove to produce big gestures they assumed the recognition system would require.

The second goal was to explore how UI Distance as a design consideration interacted with Functions to affect behavior and the risk of fatigue. The results of upper arm kinematics and RULA scores suggest that UI Distance and Function interacted when the VWA was off. One participant explained it like this: Scroll can be executed without needing to touch the UI, so she dropped her arm slightly when the UI was near, but she wanted to reach out and touch the UI for Focus-Select because direct manipulation felt more natural and intuitive for that function. In addition, the far UI distance made her instinctively reach out to it for both functions (see Figure 3).

The findings and limitations present a number of directions for future work. Themes regarding visual feedback of the hands may speak to the design requirements of future ergonomic gesture systems in general (Jerald, 2016, Ch. 27.2.5). The connection between canonical and abstract functions, intuitive interactions, UI distances, and fatigue are certainly worth exploring. Subjective instruments like the Borg CR10 would benefit from a body-part-specific component, as well as within-subject repeated-measures designs to account for large individual differences (see Table 2). We hope that a future study can embrace user-centered approaches to develop novel ergonomic gesture systems that promote freedom and wellbeing, rather than the totalitarian imposition of rigid interaction methods inherited from designers and legacy technologies. Such an approach should think critically regarding current research methods used in human factors and design research, and adopt one that opens up spaces for debate and facilitates multi-disciplinary collaboration beyond insular perspectives.

References

Ankrum, D., & Nemeth, K. (2000). Head and Neck Posture at Computer Workstations – What's Neutral? *Proceedings of the 14th Triennial Congress of the International Ergonomics Association*, *5*, 565–568.

Boring, S., Jurmu, M., & Butz, A. (2009). Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group - Design: Open 24/7, OZCHI '09*, *411*, 161–168.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, *40*(7), 36–43.

Bowman, D. A., McMahan, R. P., & Ragan, E. D. (2012). Questioning naturalism in 3d user interfaces. *Communications of the ACM*, *55*(9), 78–88.

Hoff, L., Hornecker, E., & Bertel, S. (2016). Modifying Gesture Elicitation: Do Kinaesthetic Priming and Increased Production Reduce Legacy Bias? *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, 86–91.

Jerald, J. (2016). *The VR Book: Human Centered Design for Virtual Reality*. Association for Computing Machinery.

Lee, E. C., Whitehead, A. L., Jacques, R. M., & Julious, S. A. (2014). The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Medical Research Methodology*, *14*(1), 41.

McAtamney, L., & Corlett, N. (2004). Rapid upper limb assessment (rula). In *Handbook of human factors and ergonomics methods* (pp. 86–96). CRC Press.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1), 1–13.

Pheasant, S., & Haselgrave, C. (2006). *Bodyspace: anthropometry, ergonomics, and the design of work* (3rd). CRC Press.

Ries, A. L. (2005). Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD: Journal of Chronic Obstructive Pulmonary Disease*, *2*(1), 105–110.

Ruiz, J., & Vogel, D. (2015). Soft-Constraints to Reduce Legacy and Performance Bias to Elicit Whole-body Gestures with Low Arm Fatigue. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3347–3350.

Son, M., Jung, J., & Park, W. (2017). Evaluating the utility of two gestural discomfort evaluation methods. *PLoS ONE*, *12*(4).

Stevens, R., Papka, M., & Disz, T. (2003). Prototyping the workspaces of the future. *IEEE Internet Computing*, *7*(4), 51–58.

Uva, A. E., Fiorentino, M., Manghisi, V. M., Boccaccio, A., Debernardis, S., Gattullo, M., & Monno, G. (2019). A User-Centered Framework for Designing Midair Gesture Interfaces. *IEEE Transactions on Human-Machine Systems2*, 1–9.

Villarreal-Narvaez, S., Vanderdonckt, J., Vatavu, R. D., & Wobbrock, J. O. (2020). A systematic review of gesture elicitation studies: What can we learn from 216 studies? *DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 855–872.

Williams, A. S., Garcia, J., & Ortega, F. (2020). Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Transactions on Visualization and Computer Graphics*, *26*(12), 3479–3489.

Williams, N. (2017). The Borg Rating of Perceived Exertion (RPE) scale. *Occupational Medicine*, *67*(5), 404–405.

Wu, G., van der Helm, F. C., (DirkJan) Veeger, H., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., Werner, F. W., & Buchholz, B. (2005). ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. *Journal of Biomechanics*, *38*(5), 981–992.

Wu, H., Wang, Y., Qiu, J., Liu, J., & Zhang, X. (2019). User-defined gesture interaction for immersive VR shopping applications. *Behaviour & Information Technology*, *38*(7), 726–741.