Understanding Differences in Human-Robot Teaming Dynamics between Deaf/Hard of Hearing and Hearing Individuals

A'di Dust
adust@macalester.edu
Macalester College
St. Paul, MN, USA

Carola Gonzalez-Lebron*
gonzalezlebronc1@udayton.edu
The University of Dayton
Dayton, OH, USA

Shannon Connell
sdcnai@rit.edu
National Technical Institute for the Deaf
Rochester, NY, USA

Saurav Singh
ss3337@rit.edu
Rochester Institute of Technology
Rochester, NY, USA

Reynold Bailey
rjbvcs@rit.edu
Rochester Institute of Technology
Rochester, NY, USA

Cecillia Ovesdotter Alm
coagla@rit.edu
Rochester Institute of Technology
Rochester, NY, USA

Jamison Heard
jrheee@rit.edu
Rochester Institute of Technology
Rochester, NY, USA

ABSTRACT

With the development of industry 4.0, more collaborative robots are being implemented in manufacturing environments. Hence, research in human-robot interaction (HRI) and human-cobot interaction (HCI) is gaining traction. However, the design of how cobots interact with humans has typically focused on the general able-bodied population, and these interactions are sometimes ineffective for specific groups of users. This study's goal is to identify interactive differences between hearing and deaf and hard of hearing individuals when interacting with cobots. Understanding these differences may promote inclusiveness by detecting ineffective interactions, reasoning why an interaction failed, and adapting the framework's interaction strategy appropriately.

CCS CONCEPTS

• **Human-centered computing → Interaction paradigms**; • **Computing methodologies** → *Artificial intelligence.*

KEYWORDS

Human-Robot Interaction, cobot, deaf, hard of hearing

ACM Reference Format:
A'di Dust, Carola Gonzalez-Lebron, Shannon Connell, Saurav Singh, Reynold Bailey, Cecillia Ovesdotter Alm, and Jamison Heard. 2023. Understanding Differences in Human-Robot Teaming Dynamics between Deaf/Hard of Hearing and Hearing Individuals. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23 Companion), March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3568294.3580146

*Both authors contributed equally to this research.

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden

ACM ISBN 978-1-4503-9970-8/23/03…$15.00
https://doi.org/10.1145/3568294.3580146

1 INTRODUCTION

Collaborative robots (cobots) that share space and interact with a human teammate [20] are becoming more prevalent in manufacturing as we move towards Industry 4.0 (the next industrial revolution using the internet of things) [16]. Cobots are able to increase productivity and efficiency while lowering production costs [4]. Moreover, cobots can improve working conditions for employees by taking on repetitive tasks such as assembly or heavy labor tasks [18]. However, it remains unclear which factors facilitate or hinder the successful integration of industrial cobots [15].

Trust directly influences a person's willingness to interact with a robot [6]; therefore, it is necessary to understand trust in cobot settings [26]. Without the appropriate level of trust, humans may disuse (under-reliance), or misuse (over-reliance) the robot [26]. Furthermore, studies suggest that the use of robots can improve performance and time-to-completion while increasing usability and productivity [24, 28]. Regarding perceived workload, we have yet to fully understand how and if this is affected by the use of cobots. However, some studies suggest that there is no negative impact on humans' perceived workload when collaborating with robots [10, 11]. More extensive and robust research is needed to understand how to best consider all these teaming factors when designing effective human-robot collaboration (HRC) teaming paradigms.

Prior research works seek to understand these teaming factors by running human-subject studies with over-represented individuals, such as hearing individuals. However, these factors and their respective meanings may diverge with varying populations with different cultural backgrounds. For example, a human continuously glancing at a robot's end-effector may be an indication of mistrust. However, a deaf individual may need to visually check the robot to maintain an appropriate awareness level, while a hearing individual can rely on the auditory perception of sounds to recognize the robot's current position. Broadening representation by including deaf and hard of hearing (D/HoH) individuals within experimental paradigms may provide more insight into how human-robot teams function and how their dynamics change over time. Understanding these aspects is especially critical within the manufacturing domain, as this is the largest sector that deaf individuals work in (15.7% [5]). Approximately 18% of all manufacturing employees have hearing difficulties, and about 20% of noise-exposed tested manufacturing workers have a material hearing impairment [1]. This paper's goal is to examine if there are differences between D/HoH and hearing individuals in a simulated manufacturing environment using the Baxter Robot. We detail a mixed-designed study and subsequently present results. The results indicate that D/HoH individuals trusted the robot less, but had more fluent (quality of the interaction during a shared activity) collaborations [12]. Thus, there may be a dissociation between trust and team fluency for D/HoH individuals. Future human-robot collaboration research needs to incorporate D/HoH individuals within their experimental set-ups in order to properly understand the overall team dynamics.

2 RELATED WORK

Safety is a key aspect in human-robot collaboration. The safest way to implement cobots into manufacturing is still a subject of research, especially when we consider the D/HoH community. Efforts to mitigate safety concerns include the development of post-collision responses, collision prediction, and collision free trajectory mapping technologies [27]. For instance, post-collision responses are when a cobot stops once a collision is detected. Such strategies have already been included in commercial robots [17]. Another method that has been shown to increase safety is better human situational awareness and robot interpretability. Maurtua et al. [17] developed a system where hearing subjects were able to tap the robot to make it stop moving and point to an object they wanted the robot to grasp. LED lights were also added to the robot to communicate if the gesture had been acknowledged, which increased interpretability. Overall, subjects reported a positive perception of trust in these safety implementations.

Cobots communicating intent to humans has also been explored with mixed reality [21]. Participants were asked to identify colliding and non-colliding motions with 2D visualization, no visualization and a head-mounted display. Results showed that participants had 16% higher accuracy and 63% better response time when using head-mounted displays. Participants also perceived lower overall workload compared to when using 2D and no visualizations. It is important to note that interpretability of a robot has been linked to increased trust, leading to better fluency in collaboration.

Current research involving D/HoH individuals and manufacturing cobots is scarce, as research has focused primarily on sign-language communication frameworks [2, 19, 22, 23, 25]. For example, RASA is a popular Iranian social robot that has been researched as a sign language tutor [13]. The robot was taught by having hearing individuals sign using a data glove and imputing this data into the robot. Other researchers have explored other topics such as the use of robotic dogs as service dogs for D/HoH individuals [14]. There is a need to consider other research topics that involve the D/HoH community that do not necessarily involve sign language.

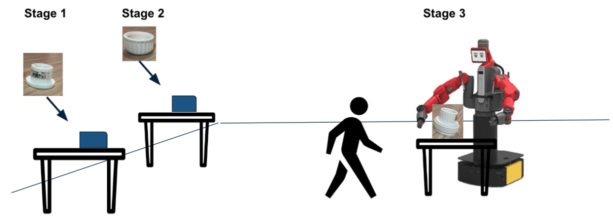

Figure 1: Physical layout for Task 1 with the Baxter Robot.

In response to this gap, our study focuses on cobots interacting with D/HoH individuals.

3 METHODOLOGY

The study's overall goal was to investigate how hearing status may impact various collaborative assembly tasks. The study used a mixed design, where hearing status (hearing vs. D/HoH) was the between-subjects variable and either the use of Baxter faces or normal vs. high cognitive load, depending on the task, was the task-specific within-subjects variable. Each participant completed four 5-minute trials with a Baxter Robot, where the first two trials were task 1 and the remaining two trials were task 2. Each trial corresponded to a task-specific manipulation, which were counterbalanced within a task. Participants were trained on a task prior to task completion.

Task 1 required participants to assemble a PVC piece with a robot (see Figure 1). The assembly was broken into three stages, which were physically separated from each other. Participants picked up a part in stage one, and walked over to stage two, which required inserting the previous part into the new part. Stage three was where participants physically collaborated in a joint-task with the robot, where the robot held a screw assembly up in the air on which the participant screwed the entire assembly together. The robot then dropped the piece, such that the participant could bring it into a bin at stage 1 and repeat the assembly process. The robot followed a fixed-timing script where it picked up an assembly piece, moved the piece in front of it, and opened the gripper after 5-seconds. The robot did not adapt to the participant. Thus, synchronization had a strong effect on task performance.

This task's physical separation was designed to hinder visual awareness, as the participant had to turn their back to the robot in stages 1 and 2. Hence, hearing individuals may have an advantage, as the auditory cues from the robot permitted some situational awareness. The robot also "failed" approximately 4-minutes into the trial, when it dropped the assembly piece prematurely on the ground. This designed failure was used to mimic failures in the real-world and determine how the participant may be impacted.

The task-specific within-subjects variable was the faces Baxter displayed, where one trial had the default Baxter face throughout its duration and the other trial had three faces corresponding to what Baxter was doing: handing a piece to the participant, picking up a piece, and when the pre-determined failure occurred (see Figure 2). These faces were inspired from Deaf culture and were hypothesized to allow a D/HoH individual to better understand what the robot was doing. For example, the robot dropping the piece on the floor, aimed to let the participant understand what happened more quickly.

Figure 2: The robot's faces. Left: when handing a piece. Middle: when picking up pieces. Right: when a failure occurred.

Task 2 focused on a complex assembly task using Lego pieces, where Baxter provided the pieces for each assembly stage to the participant seated in front of the robot. The overall assembly resembled a duck (see Figure 3) and assembly instructions were provided on a monitor. The task-specific within-subjects variable was cognitive load: normal vs. high. The normal load condition had the instructions show each individual assembly stage and the robot handed the correct piece to the human, while the instructions showed every other assembly stage in the high cognitive load condition. Participants had a separate bin in the high cognitive load condition from which they had to find the missing part. Similar to task 1, the robot ran a pre-programmed script in task 2, and it did not adapt to what the participant was doing.

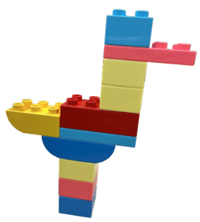

Figure 3: The completed Lego assembly for Task 2.

3.1 Participant Demographics

Twelve participants completed this Institutional Review Board-approved study. Due to sensor failures some participants had to be excluded. Seven hearing participants and two hard-of-hearing participants were analyzed. The age range for participants was 19-27 years old with a mean of 21.6 ± 2.8 years. Approximately 58% of participants identified as male, 33% identified as female, and 8% identified as non-binary. In terms of ethnicity, about 58% of the participants identified as white, 33% identified as Asian/Pacific Islander, 8% identified as Native or American Indian, and 8% identified as Hispanic or Latino. In addition, 75% of participants held a high school diploma, and 25% of participants a Bachelor's degree. Participants were able to communicate via a professional American Sign Language Professional, if requested.

3.2 Metrics

Physiological, behavioral, subjective, and performance metrics were collected throughout the experiment. A Zephyr Bioharness collected the physiological measures (i.e., heart rate, heart rate variability, and respiration rate), which were used to determine differences in cognitive load. The behavioral metrics consisted of eye-tracking-based metrics collected using Pupil Core eye-tracking glasses. The number of eye-gaze fixations and fixation durations on the robot's face, the end-effector, and the assembly pieces provided indication of a person's trust and comfort with the robot. The subjective metrics included the NASA-Task Load Index (TLX) [7], a post-trial survey, and a post-task survey. The participants completed the NASA-TLX survey after each trial to measure perceived workload. The post-trial survey was also completed, which asked participants to rate trust, comfort, and how helpful the robot was, using a Likert scale ranging from 1 (little to none) to 5 (a lot). The post-task survey asked participants what trial they preferred within a task and why. Lastly, the performance metrics consisted of productivity (number of pieces completed, time to assemble) and fluency (number of functional delays and time at each stage) metrics. Additionally, the overall interaction was encoded using the Behavioral Observational Research Interactive Software [3], which was used to encode the interaction over a trial (e.g., what assembly stage the participant was at, where was the participant looking). These encodings were then analyzed using a Needleman-Wunsch Similarity Score. The Needleman-Wunsch algorithm for creating similarity scores was originally created for the purpose of aligning protein sequences in biology and produces a score that reflects the minimum changes needed to align a sequence.

4 RESULTS

We analyzed behavioral, physiological, and survey-based metrics to understand the differences between D/HoH and hearing participants. Two D/HoH participants were used in this analysis; thus, the results should be interpreted with caution, and there is insufficient statistical power to determine significant differences. The behavioral data consisted primarily of fixations and functional delay information. Fixations are divided into robot head fixations (which account for the screen on the robot), end effector fixations, and piece fixations. Piece fixations are specifically fixations on pieces on the robot's board or in the robot's gripper. Figure 4 provides the median fixation durations by task and area type. The largest differences between D/HoH and hearing individuals occurred with head fixations, as D/HoH individuals fixated on the head approximately 0.5 seconds longer than hearing individuals during task 1. This trend was reversed for task 2, indicating that environment type (stationary vs. non-stationary) may impact what individuals with different hearing capabilities find important. Future work will systematically explore this variable.

D/HoH individuals fixated on the end-effector longer than hearing individuals during task 2, but similar fixation durations occurred during task 1. This may indicate that D/HoH individuals perceived safety as an issue, as the chance of collision was greater during task 2 due to the proximity to the robot. The two D/HoH individuals did rate trust (median=2.5) lower than hearing individuals (median=3.0) across all trials, which indicates trust may also have been an issue. The lower trust level may also be an indicator of safety concerns.

Team fluency was also investigated for Task 1 as the number of functional delays (participant was late or early to the robot station) and time at each station. The respective medians are provided in Figure 5. Overall, hearing individuals spent more time waiting on

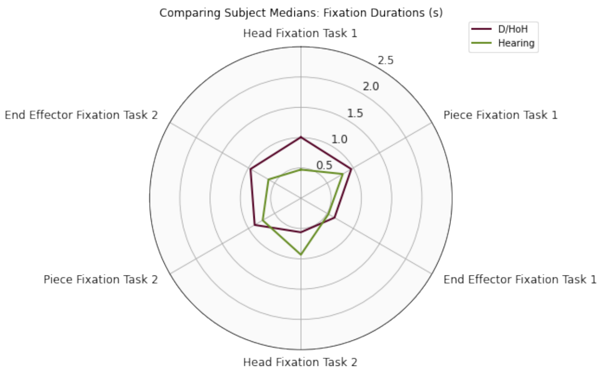

Figure 4: Radar chart comparing the duration of fixations for D/HoH and hearing individuals.

the robot than D/HoH individuals did, resulting in a larger number of functional delays. This may indicate that hearing individuals were not as synchronized with the robot as the D/HoH individuals were; yet to confirm this preliminary result, data from more subjects would be needed.

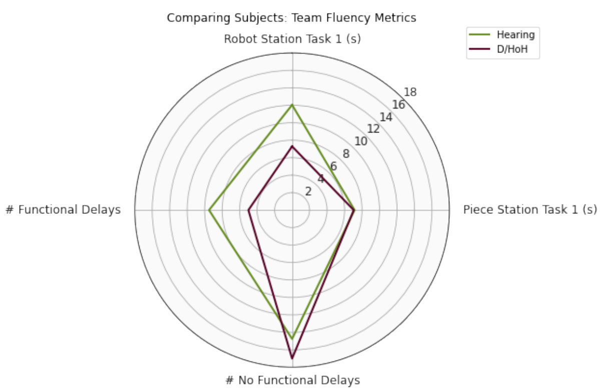

Figure 5: Radar chart comparing the team fluency measures for D/HoH and hearing individuals.

The Zephyr BioHarness collected physiological information including heart rate, heart rate variability, and respiration rate, which were all normalized from a baseline measurement to address potential individual differences. D/HoH individuals had larger mean heart-rate(hearing mean=90, D/HoH mean=109) and respiration values (hearing mean=21, D/HoH mean=22) and lower heart-rate variability values (hearing mean=74, D/HoH mean=33) than hearing individuals, indicating a higher overall workload value [8, 9]. However, the overall NASA-TLX values indicated small differences in perceived workload (hearing mean=45, D/HoH mean=34).

Lastly, time-series charts for the order of events were generated by BORIS, along with sequence similarity scores. Since we manually gathered time sequence behavioral data from gaze fixations of the participants, we found this sequence alignment method to be helpful in understanding the similarities between participants' behaviors. The BORIS software returns a score out of 100, where 100 represents identical sequences. The events are recorded as part of the sequence every second. Our analysis of the first task shows that D/HoH individuals had an average similarity of sequences at 56%, whereas hearing individuals had a similarity of sequences at 50%. While the difference is small, perhaps the D/HoH individuals behavior was somewhat more in synch with each other than for the hearing individuals.

5 DISCUSSION AND CONCLUSION

While the study's sample was small and not balanced equally between the two participant groups, some of the findings suggest the potential for differences in human-robot interactions between D/HoH individuals and hearing individuals, and deserve further study. These included behavior, subjective trust, and overall workload. Although hearing individuals perceived higher workload than D/HoH individuals, the physiological metrics indicate that the reverse was true. This discrepancy may be due to the small difference in median NASA-TLX overall workload ratings or potential confounding factors, such as trust and anxiety. The D/HoH individuals indicated lower overall trust in the robot than hearing individuals, which is supported by the D/HoH participants visually fixating more on the robot. This additional fixation time may have allowed D/HoH to have fewer functional delays and obtain higher team fluency. Conversely, the auditory cues may have resulted in higher time pressure on the hearing individuals, making them rush assembly processes that did not require the robot and then wait for the robot for another piece. Overall, there appears to be some preliminary evidence for distinct differences in the human-robot team dynamics between D/HoH and hearing individuals.

The results presented are limited by the number of participants. We could not meaningfully examine significant effects due to insufficient power. Additionally, this was a pilot study and we did not analyze/gather additional metrics that may impact a human-robot team. The robot also encountered multiple natural failures during the second task. These instances were recorded, but the timing and failure type may have impacted the collected data to some extent.

Continued work is focusing on engaging additional participants from the deaf community. Additionally, we plan to investigate higher-fidelity interaction enhancements with potential to improve the overall human-robot team. These enhancements will be created in conjunction with members of the deaf community in order to promote inclusivity in HRI.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Award No. IIS-1851591. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] 2021. Manufacturing Statistics - Occupational Hearing Loss. https://www.cdc.gov/niosh/topics/ohl/manufacturing.html
[2] Neha Baranwal, Avinash Kumar Singh, and Gora Chand Nandi. 2017. Development of a framework for human–robot interactions with indian sign language using possibility theory. *International Journal of Social Robotics* 9, 4 (2017), 563–574.

[3] Olivier Friard and Marco Gamba. 2016. boris : A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7, 11 (2016), 1325–1330. https://doi.org/10.1111/2041-210x.12584

[4] Rinat Galin, Roman Meshcheryakov, Saniya Kamesheva, and Anna Samoshina. 2020. Cobots and the benefits of their implementation in intelligent manufacturing. *IOP Conference Series: Materials Science and Engineering* 862, 3 (may 2020), 032075. https://doi.org/10.1088/1757-899x/862/3/032075

[5] Carrie Lou Garberoglio, Jeffrey Levi Palmer, Stephanie W Cawthon, and Adam Sales. 2019. *Deaf people and employment in the United States: 2019.* Technical Report. National Deaf Center on Postsecondary Outcomes.

[6] P. A. Hancock, Theresa T. Kessler, Alexandra D. Kaplan, John C. Brill, and James L. Szalma. 2020. Evolving Trust in Robots: Specification through sequential and comparative meta-analyses. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 63, 7 (2020), 1196–1229. https://doi.org/10.1177/0018720820922080

[7] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.

[8] J. Heard, C. E. Harriott, and J. A. Adams. 2017. A human workload assessment algorithm for collaborative human-machine teams. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 366–371.

[9] J. Heard, C. E. Harriott, and J. A. Adams. 2018. A Survey of Workload Assessment Algorithms. *IEEE Transactions on Human-Machine Systems* 48, 5 (2018), 434–451.

[10] Belal Hmedan, Dorilys Kilgus, Humbert Fiorino, Aurelie Landry, and Damien Pellier. 2022. Adapting Cobot Behavior to Human Task Ordering Variability for Assembly Tasks. 35 (2022).

[11] Belal Hmedan, Dorilys Kilgus, Humbert Fiorino, Aurelie Landry, and Damien Pellier. 2022. Adapting Cobot Behavior to Human Task Ordering Variability for Assembly Tasks. In *The International FLAIRS Conference Proceedings*, Vol. 35.

[12] Guy Hoffman. 2019. Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems* 49, 3 (2019), 209–218.

[13] Seyed Ramezan Hosseini, Alireza Taheri, Minoo Alemi, and Ali Meghdari. 2021. One-shot learning from demonstration approach toward a reciprocal sign language-based HRI. *International Journal of Social Robotics* (2021), 1–13.

[14] Kheng Lee Koay, Gabriella Lakatos, Dag Sverre Syrdal, Márta Gácsi, B Bereczky, Kerstin Dautenhahn, Adám Miklósi, and Michael L Walters. 2013. Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent. In *2013 IEEE symposium on artificial life (ALife)*. IEEE, 90–97.

[15] Tobias Kopp, Marco Baumgartner, and Steffen Kinkel. 2020. Success factors for introducing industrial human-robot interaction in practice: An empirically driven framework. *The International Journal of Advanced Manufacturing Technology* 112, 3-4 (2020), 685–704. https://doi.org/10.1007/s00170-020-06398-0

[16] Heiner Lasi, Peter Fettke, Hans-Georg Kemper, Thomas Feld, and Michael Hoffmann. 2014. Industry 4.0. *Business & information systems engineering* 6, 4 (2014), 239–242.

[17] Iñaki Maurtua, Aitor Ibarguren, Johan Kildal, Loreto Susperregi, and Basilio Sierra. 2017. Human–robot collaboration in Industrial Applications. *International Journal of Advanced Robotic Systems* 14, 4 (2017), 172988141771601. https://doi.org/10.1177/1729881417716010

[18] Christoph Mühlemeyer. 2020. Assessment and Design of Employees-Cobot-Interaction. In *Human Interaction and Emerging Technologies*, Tareq Ahram, Redha Taiar, Serge Colson, and Arnaud Choplin (Eds.). Springer International Publishing, Cham, 771–776.

[19] Anup Nandy, Soumik Mondal, Jay Shankar Prasad, Pavan Chakraborty, and GC Nandi. 2010. Recognizing & interpreting Indian sign language gesture for human robot interaction. In *2010 international conference on computer and communication technology (ICCCT)*. IEEE, 712–717.

[20] Alena Pauliková, Zdenka Gyurák Babel'ová, and Monika Ubárová. 2021. Analysis of the impact of human–cobot collaborative manufacturing implementation on the occupational health and safety and the quality requirements. *International Journal of Environmental Research and Public Health* 18, 4 (2021), 1927. https://doi.org/10.3390/ijerph18041927

[21] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2020. *Communicating Robot Arm Motion Intent Through Mixed Reality Head-Mounted Displays*. Springer International Publishing, Cham, 301–316. https://doi.org/10.1007/978-3-030-28619-4_26

[22] Dmitry Ryumin, Denis Ivanko, Alexandr Axyonov, Ildar Kagirov, Alexey Karpov, and Milos Zelezny. 2019. Human-robot interaction with smart shopping trolley using sign language: data collection. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 949–954.

[23] Brian Scassellati, Jake Brawer, Katherine Tsui, Setareh Nasihati Gilani, Melissa Malzkuhn, Barbara Manini, Adam Stone, Geo Kartheiser, Arcangelo Merla, Ari Shapiro, et al. 2018. Teaching language to deaf infants with a robot and a virtual human. In *Proceedings of the 2018 CHI Conference on human Factors in computing systems*. 1–13.

[24] Konrad Sowa, Aleksandra Przegalinska, and Leon Ciechanowski. 2021. Cobots in knowledge work: Human – AI collaboration in managerial professions. *Journal of Business Research* 125 (2021), 135–142. https://doi.org/10.1016/j.jbusres.2020.11.038

[25] Nazgul Tazhigaliyeva, Yerniyaz Nurgabulov, German I Parisi, and Anara Sandygulova. 2016. Slirs: Sign language interpreting system for human-robot interaction. In *2016 AAAI Fall Symposium Series*.

[26] Qiao Wang, Dikai Liu, Marc G. Carmichael, Stefano Aldini, and Chin-Teng Lin. 2022. Computational Model of Robot Trust in Human Co-Worker for Physical Human-Robot Collaboration. *IEEE Robotics and Automation Letters* 7, 2 (2022), 3146–3153. https://doi.org/10.1109/LRA.2022.3145957

[27] Pu Zheng, Pierre-Brice Wieber, Junaid Baber, and Olivier Aycard. 2022. Human Arm Motion Prediction for Collision Avoidance in a Shared Workspace. *Sensors* 22, 18 (2022). https://doi.org/10.3390/s22186951

[28] Étienne Fournier, Dorilys Kilgus, Aurélie Landry, Belal Hmedan, Damien Pellier, Humbert Fiorino, and Christine Jeoffrion. 2022. The Impacts of Human-Cobot Collaboration on Perceived Cognitive Load and Usability during an Industrial Task: An Exploratory Experiment. *IISE Transactions on Occupational Ergonomics and Human Factors* 10, 2 (2022), 83–90. https://doi.org/10.1080/24725838.2022.2072021 arXiv:https://doi.org/10.1080/24725838.2022.2072021 PMID: 35485174.