Evaluating Levels of Automation in Human–Robot Collaboration at Different Workload Levels

Dana Gutman, Samuel Olatunji* and Yael Edan

Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O. Box 653, Be‘er Sheva 8410501, Israel; Danagut@post.bgu.ac.il (D.G.); yael@bgu.ac.il (Y.E.)

* Correspondence: olatunji@post.bgu.ac.il

Abstract: This study explored how levels of automation (LOA) influence human robot collaboration when operating at different levels of workload. Two LOA modes were designed, implemented, and evaluated in an experimental collaborative assembly task setup for four levels of workload composed of a secondary task and task complexity. A user study conducted involving 80 participants was assessed through two constructs especially designed for the evaluation (quality of task execution and usability) and user preferences regarding the LOA modes. Results revealed that the quality of task execution and usability was better at high LOA for low workload. Most of participants also preferred high LOA when the workload increases. However, when complexity existed within the workload, most of the participants preferred the low LOA. The results reveal the benefits of high and low LOA in different workload situations. This study provides insights related to shared control designs and reveals the importance of considering different levels of workload as influenced by secondary tasks and task complexity when designing LOA in human–robot collaborations.

Keywords: human–robot collaboration; assembly task; user studies; user preferences; quality of task execution; usability

1. Introduction

Human–robot collaboration (HRC) involves one or more humans working with one or more robots to accomplish a certain task or a specific goal [1]. Significant research has focused on interaction aspects for designing robotic systems for use by or with humans [2–6]. This research, which focuses on factors that affect HRC [1,7] at different levels of automation, specifically evaluates the influence of workload.

The level of automation (LOA) of the system, defined as the degree to which the robot and the human are involved in the collaborative task [8–11], influences the characteristics of the dynamics of the collaboration, the behavior of the robots, actions to be taken, as well as autonomy of the human in the collaboration [12,13]. Workload addresses the actual and perceived amount of work that the human operator experiences as related to the effort invested in the task [14,15]. It can be described in terms of the elements that constitute the cost of accomplishing the goal for the human operator in the HRC [16]. These elements could be task-related (such as mental, temporal, and physical demands [17], operator-related (such as skill, strategy, experience [18]) or machine-related (such as poorly designed controls, feedback, inappropriate, or inadequate automation [15]). Workload consequences could be reflected in the stress, fatigue or frustration experienced by the human operator [16], depletion of attentional, cognitive or response resources [15], as well as in performance changes [19]. Workload can also be influenced by task complexity as characterized in terms of the stimuli involved in the task for inputs, as well as the behavioral requirements the human operator should emit in order to achieve a specific level of performance [20]. It could depend on the objective complexity derived from the task properties and on the subjective complexity which is influenced by the human operator’s perception [21]. The task properties include the component complexity—number of distinct...
actions that the human operator must execute or number of informational cues that should be processed (e.g., the number and type of subtasks to be managed, [22]); coordinative complexity—nature of relationships between task inputs and task products, the strength of these relationships as well as the sequencing of inputs (e.g., timing, frequency, intensity and location requirements [23]), and dynamic complexity—changes in the states of the environment which the human operator should adapt to [20,24].

The influence of LOA on HRC has been intensively investigated [25]. However, there are limited studies that investigated factors influencing workload in relation to the design of LOA modes suitable for different HRC collaboration contexts [26]. Moreover, research has revealed that the alignment between manufacturing strategy and automation decisions are often ad hoc in nature [27]. The current study therefore aims to examine the influence of different levels of workload when operating at different levels of automation (LOA) in a human–robot collaborative system. This is important when introducing robotics in real life situations.

To evaluate the overall performance and interaction in such HRC contexts, many different measures are commonly applied for the assessment [22,28–30]. However, by evaluating each measure separately, a holistic evaluation is lacking. We therefore specially designed two constructs that compile different evaluation measures. These constructs are useful in assessing the preferences, performance, and perception of the users regarding various aspects of the collaboration with the robot as required in a user-centered design [31–33]. The constructs are quality of task (QoT) execution (the user’s performance aspects) and usability (performance aspects along with other user perception aspects such as perceived ease of use). Additionally, user preferences were evaluated.

We design, implement and evaluate LOA modes in a user study involving 80 participants working at different workload conditions. Section 2 presents the study hypotheses, system design, LOA modes, task, and experimental evaluations of the design. Section 3 is devoted to the experimental results. Discussion is presented in Section 4 while Conclusions and suggestions for future work are discussed in the last section.

2. Materials and Methods

2.1. Experimental System

The experimental system included a 4 degree of freedom DOBOT Magician robotic arm (https://www.dobot.cc/dobot-magician/product-overview.html, accessed on 30 May 2021) equipped with a suction gripper, user interface (presented on a computer), cubes to be assembled and the human operator (Figure 1). The DOBOT Magician (135 mm high, 158 mm wide with a 320 mm radius and 500 g payload) connects to the computer through a USB connection and was programmed for the two LOA modes using the Python programming language.

Figure 1. The experimental system.
The HRC assembly task simulates a work scenario where participants are expected to assemble blocks made from cubes brought to them by a robot according to a configuration presented to them through a user interface. The task was performed in two LOA modes, at four workload levels. The workload levels, detailed below, are composed of different combinations of a secondary task and task complexity.

The user communicates with the robot through a user interface implemented on a GUI screen (Figure 2). This was designed to be friendly to promote ease of use as the human interacts with the robot through the GUI [34–36]. The configuration to be assembled is displayed on the GUI screen when starting the task. The robot brings the cubes in a sequence one after another from a predetermined place according to the specific LOA the robot is operating in. The robot releases the cube when it reaches the front of the participant. The participants are expected to assemble the cubes when received from the robot and place these cubes in a marked area on the desk in front of them.

**Figure 2.** The GUI screen.

### 2.2. Design of the Experimental Conditions

#### 2.2.1. Levels of Automation (LOA) Modes

The automation design focuses on the decision and action aspects of the overall process taken either by the robot or the user. This specifies the degree of control the user or robot in the decision of action(s) to be taken and the execution of the actions. It is conditioned in two levels for this study:

(a) Low LOA—the user has autonomy to select the type and order of cubes. The robot supports the user by bringing the type of cube the user selected via the user interface.

(b) High LOA—the robot has autonomy to bring the specific type of cube and in the order preprogrammed in its operation. The user simply demands for a cube through the user interface and the robot brings the type of cube suitable for the specific configuration assembled.

#### 2.2.2. Levels of Workload

The workload design focuses mainly on the physical and cognitive workload induced through the selection of the right cubes to assemble in the minimum possible time. This is the main task. Workload is increased in two ways: through a secondary task and by increasing task complexity.

The secondary task influence was depicted through an off-the-shelf well known cognitive game, the “RUSH HOUR” (https://www.thinkfun.com/products/rush-hour/, accessed on 30 May 2021) thinking game (Figure 3). It involves arranging toy cars in a way to get a specific car out of a gridlock. There are tabs at each stage showing how to arrange the cars and finding a way to get the required red car out at different stages.
In the main task, where cubes are assembled, the default setting is that the cubes for the assembly differ only by color. The users are required to assemble the cubes to match particular configurations characterized by differences in color pattern (Figure 4a).

The task complexity influence was depicted by introducing the cubes for the assembly that differ in color and in numbers on a particular side (Figure 4b). The users are required to assemble the cubes in color patterns as done in the low task complexity condition, but in addition, they must ensure that the specific numbers on a particular color of cubes match the required configuration per time. The task complexity is increased by the additional information cue (presence of numbers) and their spatial consideration (position of the number in the configuration). It represents component and coordinative task complexity induced through the number and type of sub-actions to be performed while selecting the right cubes and assembling along with the coordination of the actions in the secondary task.

Four levels of workload were designed using these factors:

(a) Low workload (LWL)—the users perform only the main task, assembling the cubes but with reference to the numbers on the cubes. It depicts the LWL level with increased task complexity (or high workload without secondary task).

(b) Medium workload 1 (MWL1)—the users perform only the main task of assembling the cubes but with reference to the numbers on the cubes. It depicts the MWL1 level with increased task complexity (or high workload without secondary task).

(c) Medium workload 2 (MWL2)—the users perform the main task of assembling (without references to the numbers on the cubes) simultaneously with the secondary task. It depicts the MWL2 level with increased task complexity (or high workload without secondary task).

(d) High workload (HWL)—the users perform the main task of assembling the cubes (with reference to the numbers on the cubes) along with a secondary task. This combines both secondary task and increased task complexity.
2.3. Experimental Design

The experimental design includes two independent variables: LOA and levels of workload. A between-within participant experimental design was conducted with the LOA as the within variable while level of workload was the between variable. Four groups were designed depicting the different levels of workload. Each participant was randomly assigned to one of the four groups and experienced both LOA modes (Table 1).

Table 1. Experimental design.

| Workload         | Low Workload                                      | Medium Workload 1                                    | Medium Workload 2                                    | High Workload                                   |
|------------------|--------------------------------------------------|-----------------------------------------------------|-----------------------------------------------------|-------------------------------------------------|
| Low LOA          | Condition 1a: The user chooses via a GUI screen which color of cube the robot will bring him. The user concentrates only on the main task, without reference to the numbers written on the cubes. | Condition 2a: The user chooses via a GUI screen which color of cube the robot will bring him. The user concentrates only on the main task, which has increased complexity (through the numbers written on the cubes). | Condition 3a: The user chooses via a GUI screen which color of cube the robot will bring him. The user performs a main + secondary task simultaneously, without reference to the numbers written on the cubes. | Condition 4a: The user chooses via a GUI screen which color of cube the robot will bring him. The user concentrates on performing a main + secondary task simultaneously, with an increased task complexity (must refer to the numbers written on the cubes). |
| High LOA         | Condition 1b: The robot brings the cubes to the user in a predefined order. The user concentrates only on the main task, without reference to the numbers written on the cubes. | Condition 2b: The robot brings the cubes to the user in a predefined order. The user concentrates only on the main task, which has increased complexity (through the numbers written on the cubes). | Condition 3b: The robot brings the cubes to the user in a predefined order. The user concentrates on performing a main + secondary task simultaneously, without reference to the numbers written on the cubes. | Condition 4b: The robot brings the cubes to the user in a predefined order. The user concentrates on performing a main + secondary task simultaneously, with increased task complexity (must refer to the numbers written on the cubes). |

2.4. Study Hypotheses

The model for the study (Figure 5) and the hypotheses describing the proposed connection between the constructs, user preferences and the study variables (LOA and levels of workload) along for the rationale for the hypotheses are presented as follows:

We suspect that at all workload levels, high LOA will enable the users to perform efficiently and effectively since the high LOA involves the robot carrying out most aspects of the main task which would likely improve performance [37]. Therefore, we propose:

**Hypothesis 1.** Quality of task (QoT) execution will be higher with high LOA than with low LOA for all workload levels.

Several meta-studies conducted regarding levels of automation [38], ref. [39] seem to suggest that the workload experienced by users is influenced by the LOA of the system, particularly in situations of routine performance. This does not discontinue the effect of task complexity but seems to point to the effect level of workload may have in low task complexity. Since a major component of usability is the users’ perception of the system use [40] along with effectiveness and efficiency, which high LOA will likely increase, we posit:
Hypothesis 2. Usability will be higher with high LOA than with low LOA for all workload levels.

![Figure 5. Model for the study and hypotheses.](image)

Research has revealed that as automation increases, workload is expected to decrease, particularly if the automation is properly designed and does not provide new challenges and tasks related to monitoring or other forms of engagement [39]. Moreover, in the design of adjustable robot autonomy in human–robot systems, research shows that as task complexity increases, robot effectiveness is likely to reduce if the robot is operating at higher autonomy [41]. Users seem to intuitively understand that autonomous systems could encounter difficulties in more complex situations with high uncertainty [42]. Therefore, in terms of user preferences, we propose:

Hypothesis 3. Participants will prefer high LOA to low LOA for high workload and low LOA to high LOA when task complexity is increased.

2.5. Participants

Eighty undergraduate industrial engineering third year students (44 females, 36 males, mean age = 26, SD = 1.4) participated in the study. All students had experience with both computers and robots. Participation was voluntary and every participant received compensation in the form of a bonus point contributing to a credit in an academic course. The participants completed a preliminary questionnaire which included demographics questions for the participants and the negative attitudes towards robots scale (NARS) [43].

The NARS results revealed that 21.06% of the participants had a negative attitude towards situations and interactions with robots while 63.65% were neutral about it. 26.58% had highly negative attitudes towards the social influence of robots, 47.61% had a low attitude and 25.81% were neutral about it. 65.82% had a highly negative attitude towards the concept of robots having emotions, 8.87% were indifferent about it while 25.31% had a low negative attitude towards it.

2.6. Experimental Procedure

Explanation was provided to the participants noting the robot would operate differently in the two trials. To avoid bias, the details of each trial in terms of LOA was not explained to them. They were told that a post-trial and final questionnaire will be provided to express their observations, assessments, and preferences. Then, the participant
experienced two experimental trials in which they collaborated with the robot to assemble the configuration that appeared during the GUI in a specific LOA (high/low) in random order. After each trial, they completed a post-trial questionnaire regarding their experience with the robot. At the end of the two trials, each participant completed a final questionnaire where they indicated their preferred level of automation. The experimental design and protocol were approved by the departmental ethical committee.

2.7. Dependent Variables

2.7.1. Objective Measures

**Effectiveness:** Accuracy of the robot during the task—calculated from the number of times the robot erred in bringing the cubes (e.g., failed to catch a cube, brought an incorrect cube). These are system errors to portray the context of a system whose performance may not be absolutely optimum at all times.

*Performance in the secondary task* was measured as the number of stages they passed in the secondary task (for the participants that experienced the higher workload).

**Efficiency:** Total time (in seconds) that it took the participant to complete the task for each trial. In the higher level of automation, the total time was constant since depended on robot motions only.

2.7.2. Subjective Measures

The subjective measures were collected through questionnaires that included questions regarding the participants’ experience with the robot. The post-trial questionnaire was prepared as a 5-point Likert scale ranging from “1 = strongly disagree” to “5 = strongly agree” through which participants were expected to express their experience and assessments. The questionnaire included NASA-TLX questions [17] to assess perceived workload in relation to the system efficiency. The raw NASA-TLX scores were added without the weights to provide an estimate of the overall workload (RTLX aggregation technique). The post-trial questionnaire also included questions from the technology acceptance model (TAM) to assess perceived ease of use [44]. The final questionnaire assessed user preferences regarding LOA modes and their perceptions as they collaborate with the robot at specific LOA modes.

2.7.3. Constructs

The dependent variables were defined through two constructs: QoT execution and usability. These constructs were derived from the objective and subjective measures explained above (mapping is provided in Figure 6). They were adapted to the context of human–robot collaboration from the ISO 9241-151 guideline [40,45] as follows:

**Figure 6.** Mapping of the measures into constructs for assessment. (O—objective measures; S—subjective measures).
Quality of task (QoT) execution. The extent to which specific goals in a task are accomplished to a specified degree of accuracy for a specified time period [46]. This construct involves effectiveness and efficiency of the collaboration. Effectiveness of the collaboration was evaluated by the accuracy and completeness of the task which the human and robot cooperate to execute. The efficiency of the collaboration depends on resources such as time and human effort spent to achieve the required goal [47].

Usability. The extent to which the robotic system can be used to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use (adapted from [40]). This construct, in this study, is composed of effectiveness, efficiency in addition to satisfaction derived from the perceived ease of use, perceived workload and perceived reliability of the system. All these variables could affect the degree to which the human operator believes that working with the robot will be free of difficulty or great effort. This is an adaption from [44] in the information technology domain to the context of HRC. They constitute the user’s perception regarding use of the system and is essential to ensure that the human can successfully team up with the robot to achieve such collaboration [35]. A negative user perception could lead to disuse of the support the robot can provide in the collaboration [48]. In the current study, the usability construct was comprised of the QoT measures, along with other user perceptions on ease of use, workload, and reliability.

2.8. Analysis

A generalized linear mixed model (GLMM) was applied to analyze the data with the LOA, and workload as independent variables. To combine variables for the constructs, multivariate analyses of variance (MANOVA) was used. The analyses considered all the constituent variables within constructs and combined them into a composite variable. Tukey’s honestly significant difference (Tukey’s HSD) test were used as the post-hoc test for multiple comparison. The tests were designed as two-tailed with a significance level of 0.05. The items in the user preferences questionnaire were analyzed using ANOVA to assess the effect of workload on their preferences for the LOA mode they experienced.

3. Results

Results of the assessments using the constructs (QoT execution and usability), details of the user preference regarding the LOA modes and a comparison within the workload groups are presented below.

3.1. QoT Execution

The interaction of LOA and workload had significant effect (F (3, 152) = 5.198, \( p = 0.002 \)) on the QoT execution. The QoT execution was higher at the high LOA when the workload was low compared to other LOA-workload combinations, confirming H1. LOA (F (3, 150) = 45.15, \( p < 0.001 \)) and workload (F (3, 152) = 18.725, \( p < 0.001 \)) were also significant as main effects on the QoT execution. The high LOA produced better QoT execution compared to the low LOA. Best results were obtained for low workload as expected. When the workload is high, the high LOA also produced a better QoT execution compared to the low LOA. Details of the constituent variables in the QoT execution (effectiveness and efficiency) are presented below:

3.1.1. Effectiveness

The interaction of LOA and workload did not have a significant effect on accuracy (F (3, 152) = 0.512, \( p = 0.675 \)) and neither did the LOA (F (1, 152) = 1.024, \( p = 0.313 \)) and workload (F (3, 152) = 0.376, \( p = 0.77 \)) as main effects. Workload level however, had a significant effect on the performance in the secondary task (F (1, 32) = 4.23, \( p < 0.001 \)) with MWL2 (M = 2.02, SD = 1.239) resulting in better performance compared to HWL (M = 1.93, SD = 1.047). All of the participants who did the secondary task finished the first stage of the game. The majority (71/80) reached the second stage of the game, 56/80 reached the third stage while only 10/80 reached the fourth stage.
3.1.2. Efficiency

The interaction of LOA and workload had a significant effect on completion time \((F(3, 152) = 4.838, p = 0.003)\). At high LOA and LWL, participants completed the task at shorter time compared to the other combinations. LOA also had significant effect on the completion time \((F(1, 152) = 136.565, p < 0.001)\) with the high LOA \((M = 87.3, SD = 0)\) having lower completion time compared to the low LOA \((M = 107.945, SD = 16.547)\) as expected, even though the users had the option to stop the robot’s operation at any point in the high LOA mode, thereby increasing the completion time. Workload also had significant effect on the completion time \((F(3, 152) = 4.838, p = 0.004)\) with the LWL \((M = 94.62, SD = 9.028)\) having less completion time compared to the HWL \((M = 103.158, SD = 23.924)\). Higher task complexity \((MWL1, M = 96.449, SD = 12.766)\) resulted in less completion time compared to the workload caused by the secondary task \((MWL2, M = 96.595, SD = 11.241)\).

3.2. Usability

The interaction of LOA and workload on usability was not significant \((F(18.137) = 1.615, p = 0.064)\). However, the main effects of LOA \((F(18, 135) = 7.768, p < 0.001)\) and level of workload \((F(18, 137) = 11.905, p < 0.001)\) was significant. At high LOA, the usability was higher \((M = 4.36, SD = 0.83)\) compared to the low LOA \((M = 4.31, SD = 0.773)\), in agreement with H2. At LWL \((M = 4.37, SD = 0.633)\), usability was higher compared to HWL \((M = 4.25, SD = 0.742)\). Higher usability was obtained when task complexity increased \((MWL1, M = 4.45, SD = 0.959)\) as compared to when there was a secondary task \((MWL2, M = 4.29, SD = 0.835)\).

There was no difference in the workload groups in terms of the perceived ease of use. However, workload level significantly influenced perceived workload as measured through the aggregated raw NASA-TLX scores \((F(3, 152) = 11.767, p < 0.001)\), with the HWL \((M = 14.6, SD = 4.337)\) resulting in higher perceived workload compared to the LWL \((M = 12.58, 3.796)\) as expected. Between the medium workload groups, MWL2 \((M = 15.33, SD = 3.318)\) resulted in higher perceived workload compared to MWL1 \((M = 11.18, SD = 2.123)\).

Workload also had significant effect \((F(3, 152) = 3.646, p = 0.014)\) on perceived reliability as assessed through the questionnaire. The reliability was perceived as higher by the participants who experienced the LWL \((M = 4.53, SD = 0.687)\) compared to the HWL \((M = 4.5, SD = 0.555)\). Between the medium workload levels, MWL1 \((4.63, SD = 0.628)\) resulted in higher perceived reliability compared to MWL2 \((M = 4.19, SD = 0.634)\).

3.3. User Preferences

A one-way ANOVA revealed that there was a significant difference between workload groups \((F(3, 76) = 9.276, p < 0.001)\). When comparing LWL and HWL, high LOA was preferred. However, when comparing between MWL1 and MWL2, low LOA was preferred for the MWL1 (confirming H3). More details regarding user preferences for the LOA modes between the workload groups are depicted in Figure 7.

3.4. Comparison between Workload Groups for Different LOA Modes

Multiple comparison made between the different workload groups with details on each LOA mode for groups that were significantly different are presented in Table 2. Results revealed that at low LOA: QoT execution is higher when workload is lower; usability is higher when a secondary task is involved, and user preference tended towards low LOA when complexity increases. However, at high LOA: QoT execution was the same for all workload types except when complexity is involved; usability was higher when a secondary task is involved, and user preference tended towards high LOA when a secondary task is involved.
3.3. User Preferences
A one-way ANOVA revealed that there was a significant difference between workload groups ($F(3, 76) = 9.276, p < 0.001$). When comparing LWL and HWL, high LOA was preferred. However, when comparing between MWL1 and MWL2, low LOA was preferred for the MWL1 (confirming H3). More details regarding user preferences for the LOA modes between the workload groups are depicted in Figure 7.

![Figure 7. LOA preference for the different workload levels.](image)

### Table 2. Comparison of assessment (with $p$-values) within the workload groups *

| Groups         | QoT Execution | Usability | User Preferences |
|----------------|---------------|-----------|------------------|
| LWL | MWL1         | 0.858     | 0.297            | 0.038 * Low LOA > High LOA |
| LWL | MWL2         | 0.88      | 0.03 *           | Low LOA: Low < MWL2 High LOA: Low < MWL2 0.089 |
| LWL | High         | 0.004 *   | 0.059            | 0.956 |
| MWL1 | MWL2        | 0.1       | 0 < 0.001 *      | Low LOA: MWL1 < MWL2 High LOA: MWL1 < MWL2 0 < 0.001 * Low LOA < High LOA |
| MWL1 | HWL          | 0.042 *   | Low LOA: MWL1 > HWL High LOA: MWL1 < HWL 0 < 0.001 * Low LOA < MWL1 < HWL 0.008 * Low LOA > High LOA |
| MWL2 | HWL          | 0.033 *   | Low LOA: MWL2 > HWL High LOA: MWL2 = HWL 0.782 0.242 |

* green depicts comparison with statistical significance; similar trends are marked with identical colors.

### 4. Discussion

The main influences and interacting influences of LOA in HRC in an assembly task context, considering different levels of workload is summarized in Table 3.

#### Table 3. Summary of findings.

| Metrics          | Constituent Measures                               | Significant Effects                                                                 | Finding                                                                 |
|------------------|----------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| QoT execution    | Efficiency; effectiveness                          | LOA ($p < 0.001$); workloads ($p < 0.001$); LOA*workload ($p = 0.002$)              | LOA and workload had significant effect on the QoT execution. The QoT execution was higher at the high LOA. |
| Usability        | QoT execution measures; perceived ease of use,     | LOA ($p < 0.001$); Workload ($p < 0.001$)                                         | The usability was higher at high LOA. The workload had more influence on the constituent variables, with the LWL resulting in higher usability. |
| User preferences | User choices regarding LOA modes                   | Workload ($p < 0.001$)                                                            | Most of the participants preferred the high LOA for both LWL and HWL. In the medium workload levels, the low LOA was preferred for the MWL1 where some task complexity was involved. |
4.1. Influence of LOA

In HWL situations, where additional resources are needed to complete the task in the least possible time and with minimal effort, high LOA is preferred. This corresponds with the observations made in the meta-analyses conducted in [38,39] where several automation-related data were analyzed. It also agrees with the characteristics of the suggested line of solution in workload demands amidst multiple resources as elaborated upon in [37]. However, in cases where complexity is involved, as seen in the results for the LOA preference of participants in the medium workload category, a low LOA can be considered. Most participants seem to prefer a low LOA when the task complexity is high. This confirms H3, and is also in agreement with previous studies where it was stated that a higher LOA may not always give a positive outcome in situations where uncertainties, and higher probabilities of failure exist [38,39]. In high complex tasks where high component and coordinative complexity increases the probabilities of failure [23,49], humans usually have a higher potential to better manage unknown or unexpected situations [50,51]. This reinforces the significance of evaluating LOA modes alongside different workload situations as emphasized in [52] for various contexts and causes of workload. It also calls for further assessments using these constructs.

4.2. Workload Considerations

Workload had significant influence on most of the measures. The significant effects were seen in effectiveness and efficiency leading to reduced QoT execution in situations where the workload was high. This is consistent with the literature highlighting the contribution of task-related demands (such as mental, temporal, and physical demands, including complexity demands involved in the HRC task) to workload, which could negatively influence resources available to complete task at hand [15].

The medium workload category more clearly reflects some of the differences in additional workload which can be induced by a secondary task or task complexity. Secondary task inclusion (depicted in MWL2) seems to produce a higher perception of workload compared to complexity in the task (depicted through MWL1). This could explain the reason why most users preferred the high LOA (which autonomously executes more aspects of the task) compared to the low LOA for MWL2. The LOA option seems to provide more mental space for the users to execute other tasks, particularly when the automation functioned well, as suggested in [38,39].

This difference in the medium workload category also brings into prominence the relevance of task complexity, specifically the influence reflected through the perceived reliability where MWL1 (reflecting higher complexity) condition was perceived more reliable compared to MWL2 (reflecting secondary task influence). This could be a result of higher uncertainty and failure probabilities which complexity induces as elaborated in [53,54]. It is therefore understandable that users preferred low LOA to the high LOA in this level of workload (where the task complexity exists) where they seem to have an increased sense of control over the operation [55]. This enables them to better manage the higher uncertainties in this condition (through the low LOA) compared to relying on the robot (through the high LOA). The results reveal that both objective and subjective complexity considerations as noted in [21] should be considered along with the suitable LOA modes for such HRC assembly tasks. This consequently affects the QoT execution and usability of the system.

4.3. Limitations

Evaluation was performed with users who had experience with computers and robots. We expect these results to be amplified with users who have experience in real industrial setting. We are also cognizant of potential differences in the subjective assessment of the students in comparison to professionals in an industrial setting since this plays a role in the perception of the users working alongside a robot in a work setting [56]. We therefore consider the results obtained with caution, with the perception that these could
be relatively equivalent to assessment with novice operators and different from expert or professional assessments.

The LOA and levels of workload design is simplified for research purpose and not fully representative of the degree of automation, workload levels demanded in more industrial settings. The results obtained, therefore, serve as building blocks and insights for further developments where more detailed automation, workload and complexity conditions are tested in sample industrial settings. Some other social aspects of interacting (such as verbal [57] and non-verbal communication methods [58]) with the robot for the collaborative work were not explicitly investigated in this study. However, further research should also investigate the interplay of the socio-technical aspects of the collaboration while also considering economic and societal issues to understand fuller dimensions of improved HRC in industry [56,59,60].

5. Conclusions

This paper presented the influence of LOA on a human robot collaborative assembly task considering different workload levels. The user study yielded valuable insights into participants’ preferences and influence of LOA and workload. The study also introduced two constructs for the evaluation: quality of task (QoT) execution and usability. The evaluation obtained through these constructs highlighted their potential for use in HRI studies. The study has served to provide support tools to further align manufacturing strategies and automation decisions putting into consideration level of workload to further improve productivity.

The QoT execution construct also pointed to the significance of combining efficiency and effectiveness together as a single variable. It revealed the influence of the LOA and workload in the extent to which goal of the task was accomplished under specified degree of accuracy and duration of the task. The usability construct was significant in revealing the combined effect of QoT execution and user perceptions of the ease of use, workload, and system reliability. The interactive effect of LOA and levels of workload on this construct pointed to the added value which user perceptions contribute when combined with the QoT measure.

We recommend a high LOA to support the user when the workload is high. A high LOA could reduce the stress or pressure of additional secondary tasks which the robot could support in. This was observed in the outcome of the user preferences which tended towards higher LOA when the workload was high. It also agrees with the observations of [38] in their meta-analyses considering the influence of LOA on workload. High LOA, when designed effectively, helps to extend the capabilities of the user to attend to other tasks concurrently as noted by [42,61]. However, lower LOA is helpful when high task complexities are involved, for which failure performance may occur as also noted in [39]. An adaptive LOA design that takes these outcomes into consideration is therefore recommended for further investigation.

There may be significant differences in the influence of these variables when observed in other settings, with different forms of robots, tasks and robot feedback modalities [62] and with the perception of different users as emphasized in [63]. Future work should evaluate different forms of increased workload. The workload design can be fine-tuned to portray distinct types of workload demands such as physical, cognitive and temporal demands during the task. Evaluation should also be conducted with other forms of tasks e.g., with a mobile robot delivering items and with other populations. Ongoing research is aimed at performing studies with older adults for daily living tasks and for non-professional users, putting into consideration the influence of demographics on the changes automation brings [64]. LOA has proven to influence performance for older adults [12]. We expect the effect of the levels of workload to amplify with them. The change of preferences and the differences in the reaction and performance of the older adults should be examined with different LOA options for different workload levels.
Author Contributions: Conceptualization, D.G., S.O. and Y.E.; methodology, D.G., S.O. and Y.E.; software, D.G.; validation, D.G., S.O. and Y.E.; formal analysis, D.G., S.O. and Y.E.; investigation, D.G., S.O. and Y.E.; resources, Y.E.; data curation, D.G.; writing—original draft preparation, D.G.; writing—review and editing, D.G., S.O. and Y.E.; visualization, D.G., S.O. and Y.E.; supervision, S.O. and Y.E.; project administration, Y.E.; funding acquisition, Y.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the EU funded Innovative Training Network (ITN) in the Marie Skłodowska-Curie People Programme (Horizon2020): SOCRATES (Social Cognitive Robotics in a European Society training research network), grant agreement number 721619. Partial support was provided by Ben-Gurion University of the Negev through the Agricultural, Biological and Cognitive Robotics Initiative, the Marcus Endowment Fund, and the W. Gunther Plaut Chair in Manufacturing Engineering.

Institutional Review Board Statement: This study was approved by the ethical committee of the Department of Industrial Engineering and Management at Ben-Gurion University of the Negev.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data supporting the reported results can be found at https://github.com/samuelolatunji/LOA-WorkloadLevels_Analyses.git (accessed on 9 August 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bauer, A.; Wollherr, D.; Buss, M.D. Human-Robot collaboration: A survey. Int. J. Humanoid Robot. 2008, 5, 47–66. [CrossRef]
2. Goodrich, M.A.; Schultz, A.C. Human-Robot Interaction: A Survey; Now Publishers Inc: Hanover, MA, USA, 2007.
3. Hentout, A.; Aouache, M.; Maoudj, A.; Akli, I. Human-robot interaction in industrial collaborative robotics: A literature review of the decade 2008–2017. Adv. Robot. 2019, 33, 764–799. [CrossRef]
4. Bonarini, A. Communication in human-robot interaction. Curr. Robot. Rep. 2020, 1, 279–285. [CrossRef]
5. Mohebbi, A. Human-Robot interaction in rehabilitation and assistance: A review. Curr. Robot. Rep. 2020, 1, 131–144. [CrossRef]
6. Prati, E.; Peruzzini, M.; Pellicciari, M.; Raffaeli, R. How to include user experience in the design of human-robot interaction. Robot. Comput. Integr. Manuf. 2021, 68, 102072. [CrossRef]
7. Johnson, G.I.; Wilson, J.R. Future directions and research issues for ergonomics and advanced manufacturing technology (AMT). Appl. Ergon. 1988, 19, 3–8. [CrossRef]
8. Kaber, D.; Endsley, M.R. Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. Am. Inst. Chem. Eng. 1997, 16, 126–131. [CrossRef]
9. Kaber, D.; Endsley, M.R. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics 1999, 42, 462–492. [CrossRef]
10. Lindström, V.; Winroth, M.; Stahre, J. Levels of automation in manufacturing. Int. J. Ergon. Hum. Factors 2008, 30, 1–29.
11. Shi, J.; Jimmerson, G.; Pearson, T.; Menassa, R. Levels of human and robot collaboration for automotive manufacturing. In Proceedings of the Workshop on Performance Metrics for Intelligent Systems, College Park, MD, USA, 20–22 March 2012; pp. 95–100. [CrossRef]
12. Olatunji, S.; Markfed, N.; Gutman, D.; Givati, S.; Sarne-Fleischman, V.; Oron-Gilad, T.; Edan, Y. Improving the interaction of older adults with a socially assistive table setting robot. In Lecture Notes of Artificial Intelligence, Proceedings of the 11th International Conference on Social Robotics, Madrid, Spain, 26–29 November 2019; Springer: Berlin/Heidelberg, Germany, 2019; Volume 11876, pp. 568–577.
13. Wang, W.; Chen, Y.; Li, R.; Jia, Y. Learning and comfort in human-robot interaction: A review. Appl. Sci. 2019, 9, 5152. [CrossRef]
14. Xu, J.; Anders, S.; Pruttinan, A.; France, D.; Lau, N.; Adams, A.J.; Weigner, M.B. Human performance measures for the evaluation of process control human-system interfaces in high-fidelity simulations. Appl. Ergon. 2018, 73, 151–165. [CrossRef] [PubMed]
15. Hart, S.G.; Wickens, C.D. Workload assessment and prediction. In Manprint; Springer: Amsterdam, The Netherlands, 1990; pp. 257–296.
16. Hart, S.G. NASA-Task load index (NASA-TLX); 20 years later. Proc. Hum. Factors Ergon. Soc. 2006, 50, 904–908. [CrossRef]
17. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv. Psychol. 1988, 52, 139–183. [CrossRef]
18. Hillburn, D.; Jorna, P.G. Workload and air traffic control. In Human Factors in Transportation. Stress, Workload, and Fatigue; Hancock, P.A., Desmond, P.A., Eds.; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 2001; pp. 384–394.
19. Yeh, Y.-Y.; Wickens, C.D. Dissociation of performance and subjective measures of workload. Hum. Factors J. Hum. Factors Ergon. Soc. 1988, 30, 111–120. [CrossRef]
20. Wood, R.E. Task complexity: Definition of the construct. Organ. Behav. Hum. Decis. Process 1986, 37, 60–82. [CrossRef]
47. Baraglia, M.; Cakmak, Y.; Nagai, R.; Rao, R.; Asada, M. Initiative in robot assistance during collaborative task execution. In Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 67–74. [CrossRef]
48. Parasuraman, R.; Riley, V. Humans and automation: Use, misuse, disuse, abuse. *J. Hum. Factors Ergon. Soc.* 1997, 39, 230–253. [CrossRef]
49. Hærem, T.; Pentland, B.T.; Miller, K.D. Task complexity: Extending a core concept. *Acad. Manag. Rev.* 2015, 40, 446–460. [CrossRef]
50. Monostori, L.; Vánca, J.; Kumara, S.R.T. Agent-based systems for manufacturing. *CIRP Ann.* 2006, 55, 697–720. [CrossRef]
51. Wang, X.V.; Kemeny, Z.; Vánca, J.; Wang, L. Human-Robot collaborative assembly in cyber-physical production: Classification framework and implementation. *CIRP Ann.* 2017, 66, 5–8. [CrossRef]
52. Kaber, D.B. Issues in human-automation interaction modeling: Presumptive aspects of frameworks of types and levels of automation. *J. Cogn. Eng. Decis. Mak.* 2018, 12, 7–24. [CrossRef]
53. Murthy, D.N.P. Confiabilidade e garantia de produto: Visão geral e pesquisas futuras. Product reliability and warranty: An overview and future research. *SciELO Braz.* 2007, 17, 426–434.
54. Niu, J.; Geng, H.; Zhang, Y.; Du, X. Relationship between automation trust and operator performance for the novice and expert in spacecraft rendezvous and docking (RVD). *Appl. Ergon.* 2017, 71, 1–8. [CrossRef] [PubMed]
55. Kaber, D.B.; Endsley, M.R. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theor. Issues Ergon. Sci.* 2004, 5, 113–153. [CrossRef]
56. Brynjolfsson, E.; McAfee, A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*; WW. Norton & Company: New York, NY, USA; London, UK, 2016; pp. 15–25.
57. Giuliani, M.; Lenz, C.; Müller, T.; Rickert, M.; Knoll, A. Design principles for safety in human-robot interaction. *Int. J. Soc. Robot.* 2010, 2, 253–274. [CrossRef]
58. Lenz, A.; Skachek, S.; Hamann, K.; Steinwender, J.; Pipe, A.G.; Melhuish, C. The BERT2 infrastructure: An integrated system for the study of human-robot interaction. In Proceedings of the 10th IEEE-RAS International Conference on Humanoid Robots, Nashville, TN, USA, 6–8 December 2010; pp. 346–351. [CrossRef]
59. Moniz, A.B.; Krings, B.J. Robots Working with humans or humans working with robots? Searching for social dimensions in new human-robot interaction in industry. *Societies* 2016, 6, 23. [CrossRef]
60. Moniz, A.B. *Human Resource Management and Technological Challenges*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 123–131. [CrossRef]
61. Gutman, D. Levels of Automation in Human-Robot Collaboration. Master’s Thesis, Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer Sheva, Israel, 2020.
62. Markfeld, N. Feedback for Human-Robot Collaboration. Master’s Thesis, Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer Sheva, Israel, 2020.
63. Broadbent, E. Interactions with robots: The truths we reveal about ourselves. *Annu. Rev. Psychol.* 2017, 68, 627–652. [CrossRef]
64. Acemoglu, D.; Restrepo, P. Demographics and Robots. 2018. Available online: http://www.nber.org/papers/w24421 (accessed on 9 August 2021).