# Teleoperation for Learning by Demonstration: Data Glove versus Object Manipulation for Intuitive Robot Control

Kamil Kukliński*, Kerstin Fischer**, Ilka Marhenke**, Franziska Kirstein** and Maria V. aus der Wieschen**
*Mechanical Department, Bialystok University of Technology
**Institute for Design and Communication, University of Southern Denmark.

Dorthe Sølvason, Norbert Krüger and Thiusius Rajeeth Savarimuthu
Maersk McKinney-Moller Institute, University of Southern Denmark

*Abstract*— **Learning by demonstration is a useful technique to augment a robot's behavioral inventory, and teleoperation allows lay users to demonstrate novel behaviors intuitively to the robot. In this paper, we compare two modes of teleoperation of an industrial robot, the demonstration by means of a data glove and by means of a control object (peg). Experiments with 16 lay users, performing assembly task on the Cranfield benchmark objects, show that the control peg leads to more success, more efficient demonstration and fewer errors.**

***Keywords: Teleoperation, Control Device, Object Manipulation, Robot control, Learning by Demonstration.***

## I. INTRODUCTION

Learning by demonstration is a useful technique to augment a robot's behavioral inventory, especially for small or medium size production lines, where the production process needs to be adopted or modified more often. In these cases, learning by demonstration constitutes an alternative to programming performed by a skilled technician. Learning by demonstration, by means of which a robot is taught novel behaviors possibly by lay users [9], provides the robot with trajectories and actions that it can then re-apply automatically to new objects and situations. Learning by Demonstration was introduced as a natural and intuitive way of programming robots [2][6][8]. Pardowitz et al. [12] describes learning by demonstration as one way of implementing a user-friendly programing interface for robots with the aim "to solve the problems of skill and task transfer from human to robot, as a special way of knowledge transfer between man and machine".

Demonstrations from which the robot can learn can be carried out using different methods; a common method is a kinesthetic guidance by means of which the robot is moved manually through the relevant positions, which are recorded and stored for future use. This method can be problematic, for instance, if the objects manipulated are large, far apart, or dangerous to handle. Another common method for learning by demonstration is the use of a control panel by means of which individual joints of the robot can be manipulated. This method can be very slow and tiresome. Finally, a third method is to teleoperate the robot.

During robot teleoperation, the user of the robotic system manipulates the teleoperation device and hereby manipulates the robotic system. Teleoperation describes the "demonstration technique in which the teacher operates the robot learner platform and the robot's sensors record the execution" [1]. Teleoperation allows complex movements and very intuitive handling and is thus also suited for lay users.

The most common control mode for teleoperation is the use of a data glove by means of which the robot can be steered [7]. In the current study, we employ such a data glove for teleoperation, yet we compare its usability to another external teleoperation device.

We suggest that this novel method is more intuitive and avoids the pitfalls created by misleading associations arising in the context of the data glove. We therefore investigate teleoperation with the glove and the external manipulation device. We suggest, namely an object from the Cranfield set [14], which creates more accurate associations in lay users and thus leads to more efficient teleoperation.

## II. SYSTEM

The experiments have been performed on the MARVIN platform. This platform consists of two Universal Robots (UR5), of which only one is used in this work (see Fig. 1) mounted with a Schunk SDH 3-finger dexterous gripper (see Fig. 1 bottom right). The control software is based on the RobWork framework [15], which is a collection of C++ libraries for simulation and control of robot systems, including kinematic modeling, path planning and collision checking. Communication with hardware is enabled by using the Robot Operating System (ROS) [16].

The teleoperation system is based on the trakSTAR[1] electromagnetic motion tracking system, see Fig. 2, which consists of three parts: a transmitter, a number of sensors and a controller unit. The controller unit calculates the position and orientation (the 6D pose) of each sensor with respect to the transmitter position. The transmitter is installed in such a way that it covers the workspace of the robot in the MARVIN platform. A so-called 'Dead man switch' is connected to the system, which enables teleoperation when

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 269959, IntellAct.

Paper submitted to ICUMT, St.Petersburg, Russia.

[1] www.ascension-tech.com/

pressed. An overview of the software structure and hardware interfacing is shown in Fig. 4.

Two modes of teleoperation are implemented: the ‘data glove mode’ (see Fig. 3E) and ‘the control peg mode’ (see Fig. 3D). In both cases, a trakSTAR sensor (mounted either in the glove or in the object) tracks the 6D pose relative to the transmitter and this pose is then transformed into the robot on a 1:1 movement scale. The user holds the ‘Dead man Switch’ in his or her left hand while his right hand is either wearing the glove (see Fig. 3C) or holding the object (see Fig. 3A). When doing teleoperation, s/he pushes (and holds) the ‘Dead man Switch’ (see Fig. 3B) while producing with his or her right hand the desired robot movement.

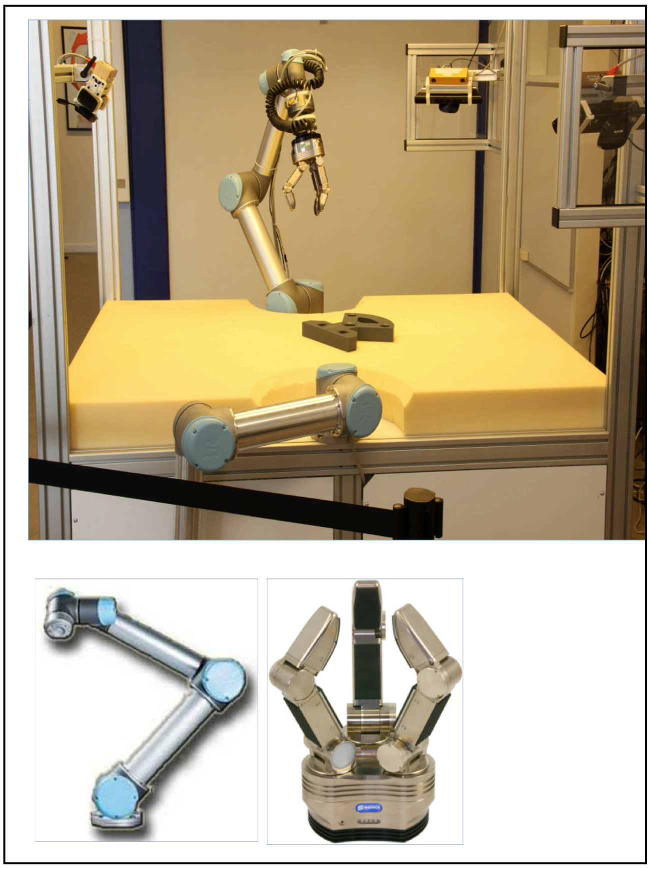

Fig. 1. Universal robot (UR5) and Schunk SDH hand on the MARVIN platform.

To sum up, two teleoperation devices are compared in this study: the data glove, where the motion sensor is integrated in the glove at about the back of the hand, and an object similar to those manipulated by the robot arm into which the sensors have been integrated (the control peg).

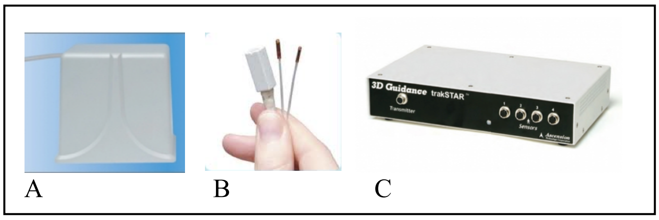


Fig. 2. trakSTAR system with transmitter (A), sensors (B) and control unit (C).

## III. Empirical Evaluation of the System

In these experiments, we compare the traditional approach to learning by demonstration via teleoperation, using the data glove, with the external device described above, the control peg.

### A. Methods

Since the user has the role of a teacher or a demonstrator, s/he needs to know how to transfer her/his knowledge to the robot. ”*What is good for the demonstrator may not necessarily be good for the imitator*” [4]. The user needs to have an understanding of the robot, its abilities and limitations. The quality of the robot’s performance is in fact highly influenced by the user [1].

Furthermore, much research done on robot learning by demonstration focuses on the robotic system, which however has been identified to be the main problem in the usability of products and interfaces [13]. Since, the central characteristic of learning by demonstration is that the user does not need to be an expert in robotics, it is essential to take her/him into account. Therefore, we apply a user-centered approach to learning by demonstration.

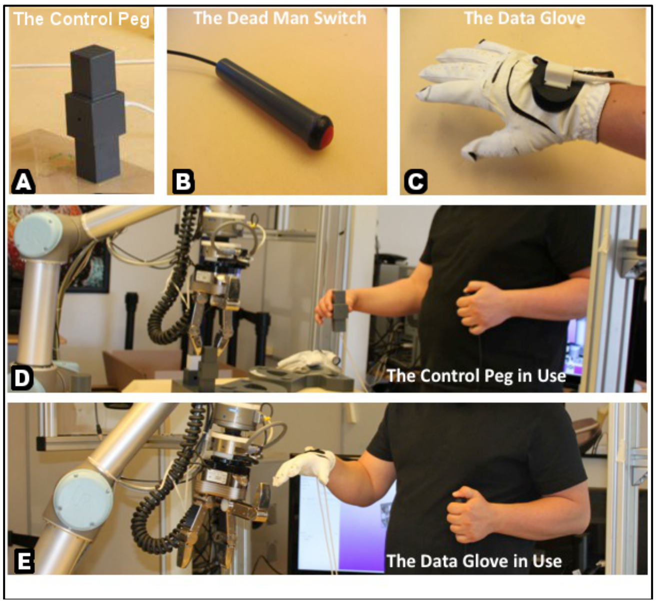


Fig. 3. A: Control peg. B: Dead man switch. C: Data glove. D: Control peg in use (middle). E: data glove in use (lower).

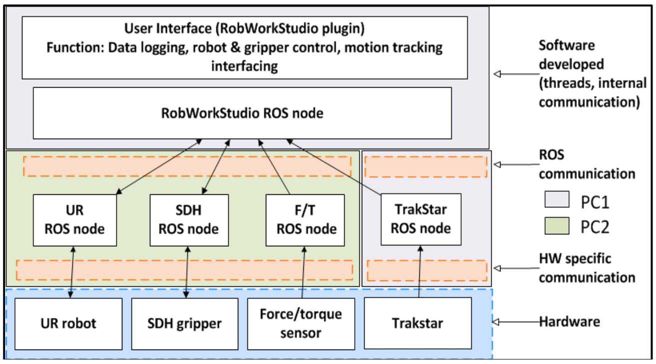


Fig. 4. Software structure and hardware interfacing. The lowest layer represents the hardware (blue) used in this project. The hardware is connected through ROS modules to the control software programmed in RobWork. This allows for a distributed system where multiple pc’s can be used to share the computational load.

### B. Tasks

The tasks that the participants had to perform comprise peg-in-hole tasks, which consists of moving objects from one place to another and placing them in a particular hole. Peg-in-hole is a well-known assembly process, which consists of different sub-tasks:

1. Moving the gripper towards a peg
2. Aligning the gripper with the peg and grasping it
3. Moving the gripper holding the peg towards the faceplate
4. Aligning the peg with the hole and placing it

The objects used for this peg-in-hole tasks are parts of the Cranfield Benchmark [14], an assembly set frequently used in robotic research. It consists of nine differently shaped objects and requires basic assembly actions. For this study, four pieces of the original benchmark were used, including the faceplate and three different objects that were to be placed in the faceplate (see Fig. 5).

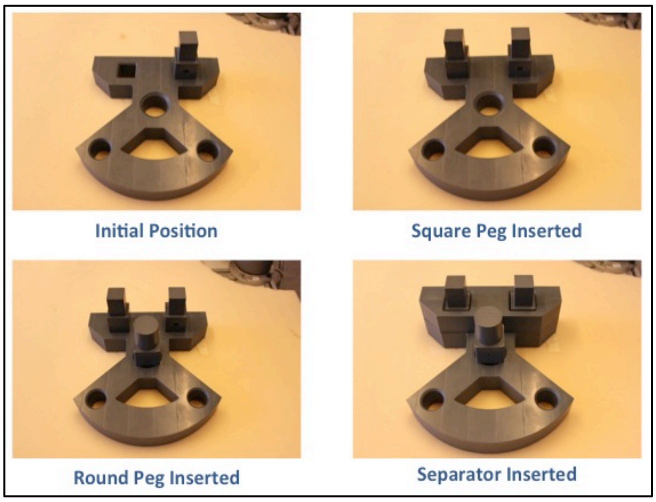


Fig. 5. 5 Objects of the Cranfield benchmark.

*C. Participants*

A within-subject approach was used, that is, one set of participants was used to test both control modes, the control peg and the data glove. 16 participants were recruited through announcements on social networks and the information boards in the university as well as by word-of-mouth. The only requirement was that participants had no prior experience with the MARVIN platform. All 16 participants were either students or employees of the University of Southern Denmark. One was female while 15 were male. They were between 19 and 36 years old, with an average age of 25.3.

*D. Procedure*

A laboratory set-up was used due to the fact that the MARVIN platform is installed in a robotic lab in the Engineering Faculty of the University of Southern Denmark in Odense.

This laboratory set-up allows a high level of control and stable variables. A room next to the robotic laboratory was used to welcome the participants before the tests, to have them fill out the consent form and questionnaire, and to show them the instruction video, see attached video. While watching the video the participants were left alone in order for them to focus on the instructions given.

After this, the participants (P) entered the actual laboratory where they were introduced to the three researchers (R) and to the robot supervisor. Afterwards they were briefly introduced to the task by one of the researchers and told which control mode they were supposed to begin with, the data glove or the control peg. Two camcorders were used to videotape the trials in case one of the cameras failed to record. The cameras were both directed at the participant and the robot, one from across the room, and one from a diagonal angle, see Fig. 6.

The order in which participants used the data glove and the control peg was randomly assigned and conditions counter–balanced.

*E. Think Aloud Technique*

The combination of methods used in this study resembles the reconciled theory and practice of thinking aloud [5]. The idea behind this technique, called "speech communication", was to obviate problems with the traditional think-aloud method. Boren et al. [5] describes the communication process as a dialogue between the participants and the experimenter, not as a monologue by the participant. The dialog allows the researcher to request further information about specific aspects of the system or product by directing the participants' attention towards it. Another benefit of giving the experimenters a more active role is that they can interfere whenever technical problems or breakdowns occur [5].

In this study, questions were asked in response to both verbal and non-verbal cues from the subjects. Whenever the participants gave a non-verbal cue like laughing, frowning, sighing, they were asked what exactly caused this reaction. Non-verbal cues are interesting to pay attention to since they often happen unconsciously and thus, reflect what the participants really feel or think [13]. During the task-solving process, the participants were not given any clues on how to do something.

Whenever the participants asked a question concerning the robotic system and its control, they were reminded of the instruction video and asked to find a solution themselves. Giving clues can prevent the participants from struggling and solving the situation themselves, which, on the other hand, can reveal unexpected issues or opportunities [13].

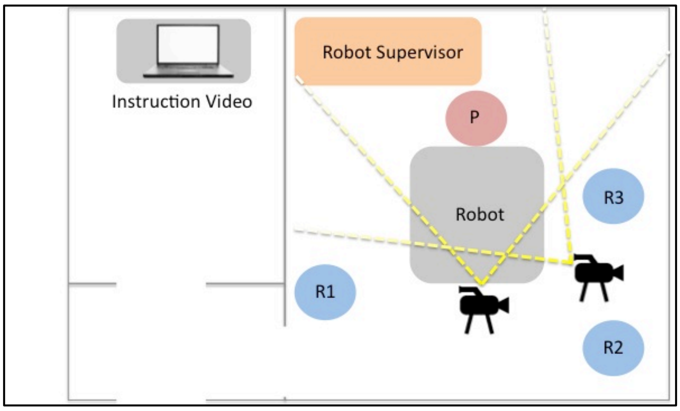


Fig. 6. Experimental personal: R1, R2, R3 are researchers and P is the participant.

To sum up, the empirical methods used in this study are a combination of the think-aloud technique, open interviews during and after the actual experiments, as well as a questionnaire before the experiment. The combination of empirical methods provides information about both user performance and user experience. This information is essential for the design process of an intuitive and user-friendly interface.

*F. Measures*

Task success was measured with respect to success, efficiency and the number of errors. Concerning *success*, the number of subtasks fulfilled in a time frame of 20 minutes was taken in account. Since the task consisted of placing three objects, two pegs and one separator (see Fig. 5), four categories of success can be distinguished: A participant did either not fulfill the task at all, placed one object correctly, placed two objects correctly, or fulfilled the task by placing all three objects correctly. Since the separator can only be placed when both the square peg and the round peg had been placed correctly before, fulfilling one third or two thirds of the task always means that participants have placed the square and/or the round peg.

Efficiency is the “quickness with which the user’s goal can be accomplished accurately and completely and is usually a measure of time” [13]. Thus, the efficiency is measured by looking at the time the participants need to fulfill the task completely. Finally, we investigate the number of *errors* occurring in the two control modes.

## IV. Results

### *A. Quantitative Data*

The sixteen participants completed the whole set of subtasks using the control peg, but only nine of the sixteen participants completed all subtasks using the data glove. The average completion rate is 77.08% for the glove where 9 participants completed all 3 tasks, 4 completed 2 out of the 3 tasks, 2 completed 1 out of 3 tasks and 1 completed none of the tasks within the 20 min.

The average completion rate is 100% for the peg; that is, all participants succeeded in placing all three items from the Cranfield set [14] correctly within the given 20 minutes timeframe. The difference is significant ($t= -3.738$, $p < .0005$), see Fig. 7.

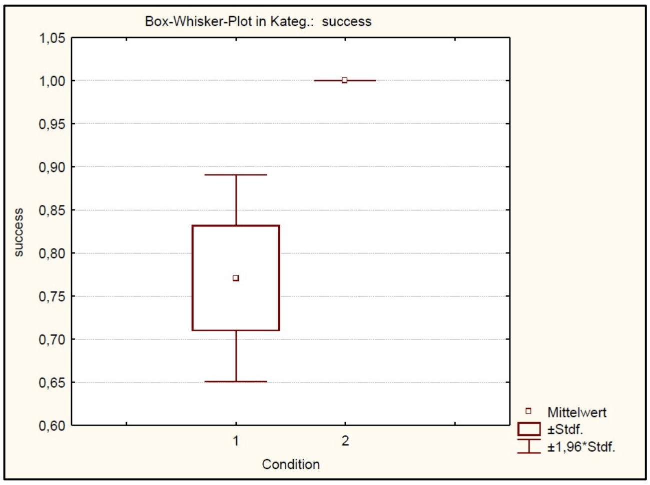


Fig. 7. Completion rate of data glove (condition 1) and control peg (condition 2).

Regarding efficiency, the mean completion time for the individual subtasks is 188.24 seconds (sd = 121.45) for the data glove and 144.52 seconds (sd = 84.15) for the control peg. These results also show that the standard deviation for the data glove is very high, indicating that users differ considerably in demonstration efficiency when using the data glove. The difference approaches significance: $t= 1.950$, $p= .0535$, see Fig. 8.

Furthermore, there is a higher training effect for the glove than for the control peg, such that it takes the user longer to insert the first peg into the hole than the second. In order to illustrate the order effect, Fig. 9 shows the time used for each task for those who started with the data glove (condition 1) in comparison with those who started with the control peg (condition 2). Thus in condition 1, tasks 1, 2 and 3, as well as in Condition 2, tasks 4, 5 and 6 were carried out with the data glove, while in Condition 1, tasks 4, 5 and 6 and in Condition 2, tasks 1, 2 and 3 were carried out using the control peg.

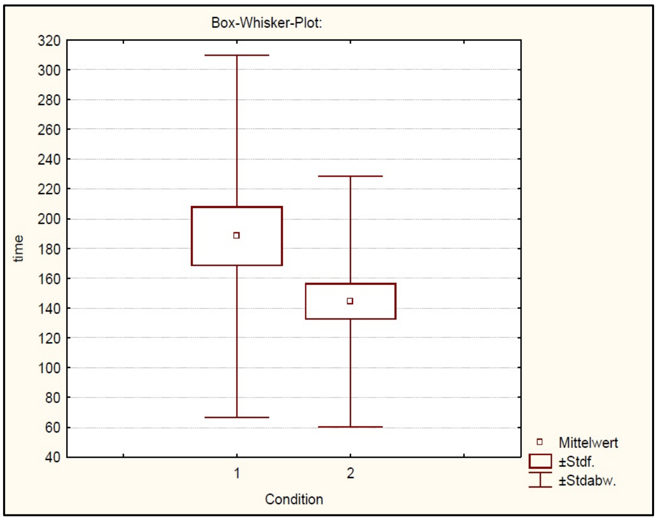


Fig. 8. Plot of completion time for the individual subtasks (condition 1 = data glove, condition 2 = control peg).

The training effect is more pronounced if users start with the data glove than when they start with the control peg. Especially interesting is that it takes participants the longest to get the peg into the hole if they have operated the control peg first and switch to the glove then (see Fig. 9). Thus, the fourth peg in condition 1, i.e. the first peg moved with the control peg after teleoperating with the glove, takes the longest (330.3 seconds on average).

In contrast, participants using the control peg only have problems with the more complex separator, yet the operating times are still much below those needed when using the data glove. The interaction between time, order and condition is highly significant ($F(11, 73)= 4.330$, $p < .0001$).

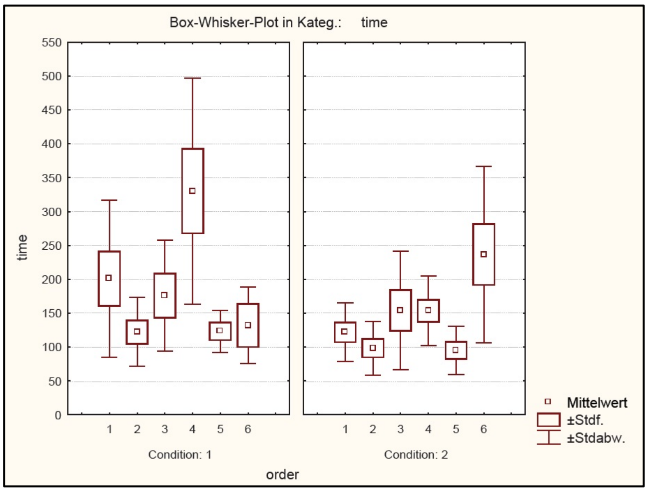


Fig. 9. Time used for each task for those who started with the data glove (condition 1) in comparison with those who started with the control peg (condition 2). Thus Condition 1, tasks 1, 2 and 3, as well as Condition 2, tasks 4, 5 and 6 were carried out with the data glove, while Condition 1, tasks 4, 5 and 6 and Condition 2, tasks 1, 2 and 3 were carried out using the control peg.

Finally, the frequency of different kinds of errors in the two control modes was investigated. Fig. 10 illustrates the overall number of errors that occurred during the user study for all 16 participants. In the data glove mode, almost twice as many errors occurred than in the mode in which participants used the control peg. Applying too much pressure was the main problem of the data glove, while the main problem of the control Peg mode were technical issues which were not caused by the user.

These results indicate that the control peg is more effective when it comes to the error rate than the data glove is.

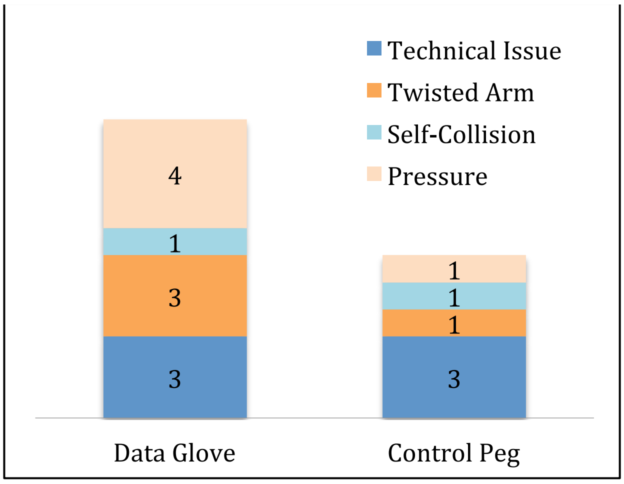


Fig. 10. Error types in the two control modes.

### B. Qualitative Data

We furthermore carried out a qualitative analysis of participants' statements during robot operation, which we recorded, transcribed and color-coded regarding positive (blue), negative (orange) and neutral statements (black) (see Table I).

TABLE I
THE STATEMENTS ABOUT THE CONTROL MODES

| Statements about glove | Statements about peg |
|---|---|
| "I can't seem to control it"; "I need to get a hang of it" | "It's easier than the glove"; "Small movements are a lot easier" |
| "I feel a little handicapped"; "I don't feel in control at all" | "I feel like the robot does exactly what I want it to do" |
| "It doesn't do what I want it to do"; "I thought it would be easier, but it is really hard to control" | "It's quite easy"; "I don't know what I'm doing" |
| "This is more natural, but peg is easier"; "It's illogical"; "Grasping is sweet" | "I move it like it's my own arm" |
| "It feels like an extension of my own arm, quite easy" | "It's quite easy"; "I'm not in doubt I'm in control" |
| "It's pretty hard to control"; "It's like an Xbox controller"; "I still haven't figured it out" | "It really does what I want it to do" |

The statements about the two control modes show that people generally do not have a good sense of control when teleoperating with the data glove, and that the apparent analogy with their own arm is in fact misleading.

## V. Discussion

The results show that teleoperation by means of the control peg is more efficient in terms of speed, success rate and number of errors than teleoperation by means of the data glove in this scenario.

Furthermore, while the demonstration by means of the data glove exhibits considerable sensitivity to training, the control peg allows more intuitive access from the start, at least for objects similar to the control object. That is, people needed longer for the separator than for the round and square pegs when using the control peg for demonstration. The separator also constituted a problem for users of the data glove, such that it took them slightly longer to insert it due to higher complexity of the problem yet the difference was not as pronounced for the users of the data glove.

However, the success rates are considerably higher for the demonstration using the control peg as well as the average demonstration times. Thus, the control peg seems to allow more efficient teleoperation than the traditional data glove for the tasks at hand.

The qualitative data provide an explanation why this may be the case: While the data glove suggests intuitive handling, it is misleading because of the seeming analogy between the human's and the robot's arm.

This may of course be different if the robot instructed is a humanoid [7], such that there is a higher correspondence between human and robot morphology and degrees of freedom [11]. However, in industrial scenarios, the control peg may be more suitable.

## VI. Conclusion and Future Work

The external teleoperation device proposed is superior to the traditional data glove in many ways, even though the glove seems to be very intuitive to handle. However, our data show that the glove provides users with a wrong sense of security, and that they are more efficient with the control peg.

These results give strong evidence for the potential of using an external device which provides intuitive feedback during the teleoperation process to the user. In our current research, we are developing such device which will facilitate learning by demonstration for robot assembly.

## References

[1] B. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems,* vol. 57, pp. 469-483, 2009.

[2] P. Bakker, and Y. Kuniyoshi, "Robot see, robot do: An overview of robot imitation," *AISB Workshop on Learning in Robotics and Animals*. pp. 3-11, 1996.

[3] C. M. Barnum, "Ready, Set…Test," *Usability Testing Essentials*, Burlington, MA: Elsevier, 2010.

[4] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," *Springer Handbook of Robotics*. pp. 1371-1394, 2008.

[5] M. Boren, and J. Ramey "Thinking aloud: Reconciling theory and practice," *IEEE Transactions on Professional Communication, 43 (3)*, pp. 261-278, 2000.

[6] S. Calinon, (2009), "A Probabilistic Approach,", *Robotic Programming by Demonstration,* Lausanne: EPFL Press, 2009.

[7] C. L. Campbell, R. A. Peters II, R. E. Bodenheimer, W. J. Bluethmann, E. Huber, and R. O. Ambrose, "Superpositioning of Behaviors Learned through Teleoperation," *IEEE Transactions on Robotics*, vol. 22, no. I, pp. 79-91, 2006.

[8] Y. Demiris, and G. Hayes, "Imitation as a dual-route process featuring predictive and learning components: a biologically-plausible computational model," *Imitation in Animals and Artifacts,* Cambridge, MA: MIT Press, pp. 327-361, 2002.

[9] H. Friedrich, R. Dillmann, and O. Rogalla, "Interactive robot programming based on human demonstration and advice," *Proceedings of the International Workshop on Sensor Based Intelligent Robots*. pp. 96-119, 1999.

[10] D. Liljekrans, "Investigating peg-in-hole strategies through teleoperation," Master's Thesis. University of Southern Denmark, Odense, 2012.

[11] C. Nehaniv, and K. Dautenhahn, “The Correspondence Problem in Social Learning: What Does it Mean for Behaviors to ‘Match’ anyway?,” *Perspectives on Imitation: From Cognitive Neuroscience to Social Science* (Royaumont Abbey, France, 2002).

[12] M. Pardowitz, S. Knoop, R. Dillmann, and R. D. Zöllner, “Incremental learning of tasks from user demonstrations, past experiences and vocal comments,” *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, 37 (2)* pp. 322-332.

[13] J. Rubin, and D. Chisnell, *Handbook of UsabilityTtesting – How to Plan, Design, and Conduct Effective Tests*. Indianapolis, IN: Wiley, 2008.

[14] K. Collins, A. Palmer and K. Rathmill, “The Development of a European Benchmark for the Comparison of Assembly,” Proc. First Robotics Conference, Brussels, 1984.

[15] L. P. Ellekilde, and J. A. Jorgensen, “RobWork: A Flexible Toolbox for Robotics Research and Education,” Robotics (ISR), 2010 41st International Symposium on and 2010 6th German Conference on Robotics (ROBOTIK), 2010.

[16] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Y. Ng, “ROS: an open-source Robot Operating System,” in Proc. Open-Source Software workshop of the International Conference on Robotics and Automation (ICRA), 2009.