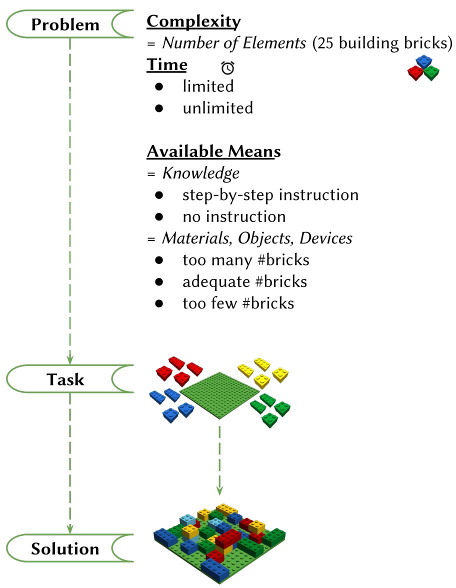



**Figure 1: Problem space: As defined by De Velde [9], we chart our problem space as a set of problems, tasks and solutions. We set out different characteristics for problems and map them to LEGO assembly tasks with predefined solutions (= goals).**

# Go for GOLD: Investigating User Behaviour in Goal-Oriented Tasks

**Sarah Prange[1,2], Daniel Buschek[2], Ken Pfeuffer[3], Lukas Mecke[1,2], Peter Ehrich[2], Jens Le[1,4], Florian Alt[3]**
[1]University of Applied Sciences Munich, Germany, {firstname.lastname}@hm.edu
[2]LMU Munich, Germany, {firstname.lastname}@ifi.lmu.de
[3]Bundeswehr University Munich, Germany, {firstname.lastname}@unibw.de
[4]Technical University of Munich, Germany, {firstname.lastname}@tum.de

## ABSTRACT

Building adaptive support systems requires a deep understanding of why users get stuck or face problems during a goal-oriented task and how they perceive such situations. To investigate this, we first chart a problem space, comprising different problem characteristics (complexity, time, available means, and consequences). Secondly, we map them to LEGO assembly tasks. We apply these in a lab study ($N = 22$) equipped with several tracking technologies (i.e., smartwatch sensors and an *OptiTrack* setup) to assess which problem characteristics lead to measurable consequences in user behaviour. Participants rated occurred problems after each task. With this work, we suggest first steps towards a) understanding user behaviour in problem situations and b) building upon this knowledge to inform the design of adaptive support systems. As a result, we provide the *GOLD* dataset (**G**oal-**O**riented **L**ego **D**ataset) for further analysis.

## KEYWORDS

User Behaviour; User Model; Goal-Oriented Interaction; Problems; Support Systems; Smartwatch; LEGO; OptiTrack; Motion Data


*CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK*

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3312949

**Complexity:** Number of involved LEGO building bricks, constant: each building (=goal) consisted of 25 bricks.
**Time:** time available for task, cf. condition (2).
**Available Means:** *Knowledge* represented by a (missing) paper-based step-by-step instruction, cf. condition (1). *Material* represented by the number of available bricks, cf. condition (3).
**Consequences:** We investigate consequences (i.e., resulting user behaviour) and user perception towards these problem characteristics.

**Sidebar 1: Mapping of problem characteristics (Fig. 1) to LEGO assembly tasks.**

(1) *(Without) Instruction*: LEGO tasks to be solved with or without a paper-based instruction (i.e., same style as people know them from LEGO) (referred to $I$ for instruction and $\neg I$ for no instruction).
(2) *(Without) Time Limit*: Participants have to solve the tasks undergoing a specific timelimit or not (referred to $T$ for time limit and $\neg T$ for no time limit). After a pilot test with three researches building single LEGO tasks independently, we choose a time limit of $02:30min$, which refers to the fastest of those and would lead to a stressful situation, hence problems, for the others.
(3) *Number of Building Bricks (normal / less / more)*: LEGO bricks available for the task (referred to as $\rightarrow B$ for normal, $\downarrow B$ for less, and $\uparrow B$ for more).

**Sidebar 2: Study conditions resulting from the mapping in sidebar 1.**

## INTRODUCTION

In daily life interaction, most tasks are goal-oriented (i.e., the desired solution is quite clear and can easily be tested). However, achieving this goal may oftentimes not be trivial due to potential problems. With this work, we obtain an understanding of user behaviour in problem situations to inform the design of, e.g., adaptive support systems. We propose (a) a problem space for everyday life interaction tasks that synthesises related work (Fig. 1), and (b) a mapping of problem characteristics to LEGO assembly tasks, resulting in a set of conditions which we tested in our exploratory lab study (sidebars 1, 2). We provide (c) a brief analysis on user perception towards these problems, and (d) a dataset comprising motion tracking data from two smartwatches and an *OptiTrack* setup.

## BACKGROUND & RELATED WORK

Several reasons for potential problems in interaction have been investigated. As an example, stress recognition using mobile devices and wearable sensors [8] as well as behavioural features [1] has been done. However, detecting stress (e.g., due to limited time) as a problem has not been subject to research so far. Furthermore, we assume that low proficiency and high cognitive workload may lead to problems in interaction and correlate to visible changes in behaviour. Metrics for this exist (e.g., NASA-TLX [4]) and cognitive workload during assembly tasks has been investigated [3, 5]. Lastly, adaptive support systems exist in various fields of HCI and for various target groups (e.g., in-situ support for workers doing manual assembly [2] or assistive systems in kitchens for people with cognitive impairments [6]). With this work, we inform future designs of such support systems by taking into account users' problems and needs through an analysis of their behaviour in and perception towards respective situations.

## PROBLEM CHARACTERISTICS

Frequent daily life tasks (cooking, interaction with devices, or manual assembly) mostly follow a certain, predefined goal (cf. Norman's goal-oriented "gulf of execution" [7]), whereas achieving this goal may oftentimes be non-trivial. The following characteristics describe such problems (cf. Fig. 1).

### Complexity

We define *complexity* as the number of elements involved in the desired action or task.

### Time

*Time* as a characteristic influences a problem solving task in several ways: (a) a given time limit can increase the difficulty and negatively influence the users' performance (possibly vice versa), (b) an occurrence of a problem itself can increase the time necessary for the task and vice versa.

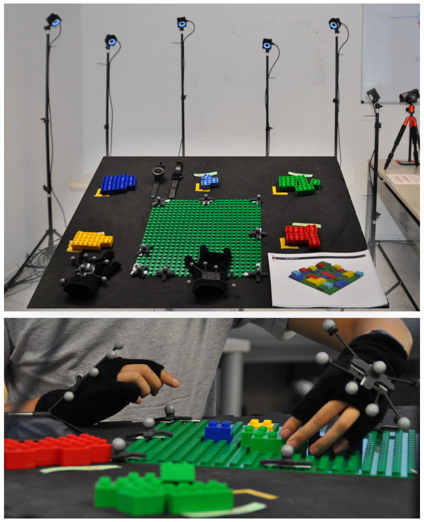


**Figure 2: Setup: We provided a clean setup at the beginning of each task (top). We used markers at both hands to track movements (bottom).**

**Table 1: Tasks: We combined our conditions and respective values to a total of 12 tasks (T0-T11).**

| ID | Conditions | | |
|---|---|---|---|
| T0 * | $I$ | $\neg T$ | $\rightarrow B$ |
| T1 | $I$ | $T$ | $\uparrow B$ |
| T2 | $I$ | $T$ | $\rightarrow B$ |
| T3 | $I$ | $T$ | $\downarrow B$ |
| T4 | $I$ | $\neg T$ | $\uparrow B$ |
| T5 | $I$ | $\neg T$ | $\downarrow B$ |
| T6 | $\neg I$ | $\neg T$ | $\uparrow B$ |
| T7 | $\neg I$ | $\neg T$ | $\rightarrow B$ |
| T8 | $\neg I$ | $\neg T$ | $\downarrow B$ |
| T9 | $\neg I$ | $T$ | $\uparrow B$ |
| T10 | $\neg I$ | $T$ | $\rightarrow B$ |
| T11 | $\neg I$ | $T$ | $\downarrow B$ |

* baseline

### Available Means

The solution of a task depends on the *available means* – if missing, a task may be impossible to solve.

*Knowledge.* On one hand, *knowledge* of the single required steps is an important mean to solve a task. If a user does not even know about these, a problem appears to be impossible to solve.

*Material.* On the other hand, the availability of *material* is an important mean to solve a task. If an essential part, object or device is not available, solving a task might hardly or not at all be possible.

### Consequences

Different characteristics of problems may result in different *consequences* for users. We assume that some of the illustrated characteristics may lead to more frustrating problems for users than others. Some problems may result in an instant need for solving (e.g., missing material), while some others may influence interaction continuously (e.g., limited time).

## STUDY DESIGN

We mapped the problem characteristics to LEGO assembly tasks (cf. sidebar 1). We conducted a between subjects lab study with the three variables instruction I, time T, and building bricks B.

### Conditions

We provided participants with LEGO assembly tasks along the problem characteristics as presented in Sidebar 2. We tested all combinations of these conditions, resulting in a total of 12 tasks (T0-T11, refer to Tab. 1), with T0 serving as baseline task (i.e., we assume an available instruction, no time limit, and adequate amount of building bricks to be the ideal, non-problematic case for a LEGO assembly task).

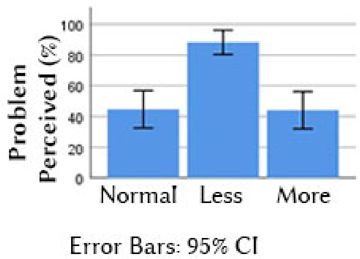
### Apparatus

*Tracking Technology.* To investigate user behaviour during the LEGO assembly tasks, we especially tracked users' hands while assembling the bricks. We applied Android smartwatches on both hands. In addition, we recorded high quality optical data via an *OptiTrack* system with six cameras, having markers on users' hands and the LEGO plate.

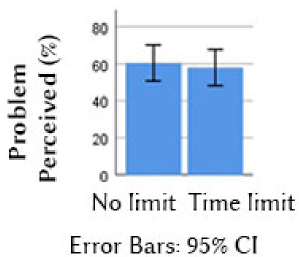
*Setup.* We provided the same setup at the beginning of each task, i.e. the LEGO bricks sorted by colour and piles of bricks as well as the instruction having the same distance to the LEGO ground plate. The arrangement was centred on a desk. With this clean setup, we wanted to avoid problems related to side effects such as searching (refer to Fig. 2).

### Procedure

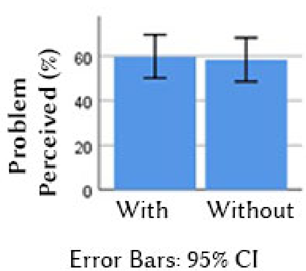
We invited participants to 60 minute sessions to our lab. All participants conducted the baseline task (T0) first, followed by the counterbalanced order of tasks T1-T11. Each task was concluded with a




**(a) Bricks**




**(b) Time**




**(c) Instruction**

**Figure 3: Perception of problems: Participants' rating (yes, no) on whether they perceived a problem during the task (mean).**

short questionnaire, comprising 5-point Likert items regarding perception, frustration and satisfaction. We recorded each session on video.

### Participants

We recruited 22 participants of which five were excluded due to missing or broken data (regarding one of the tracking technologies we applied). The remaining 17 participants were on average 23.4 years old (SD=4.37, MIN=18, MAX=32, 8 female). Most of them were students with diverse study subjects.

## PRELIMINARY RESULTS

### Qualitative Results

*Perception of Problems (Fig. 3).* When asked whether or not they perceived a problem during the task, our participants mostly perceived a problem for missing building bricks (refer to condition 3 and Fig. 3a)), while there is only small differences in the perception of a problem for (without) time limit (condition 2 and Fig. 3b)) as well as for (without) instruction (condition 1 Fig. 3c)).

*Frustration (Fig. 4).* Participants as expected stated time limit to be frustrating (Fig. 4b)).

*Solution to Problems (Fig. 5).* Participants were mostly content about their own solution, except for missing building bricks (refer to Fig. 5a)).

### The GOLD Dataset

As a result of the study, we present the *GOLD* dataset (**G**oal-**O**riented **L**ego **D**ataset), comprising, besides a timestamp, the following features:

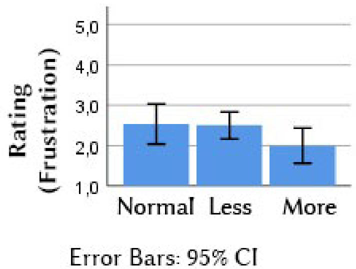
- Smartwatch sensor data
  gyroscope | linear acceleration | accelerometer | gravity | heart rate | game rotation vector
- OptiTrack data
  LEGO plane: rotation x, y, z, w | position x, y, z | mean marker error
  hand (left, right): rotation x, y, z, w | position x, y, z | mean marker error

The dataset includes this set of features for each participant for each hand and each task with respective conditions. We provide the dataset for further analysis.

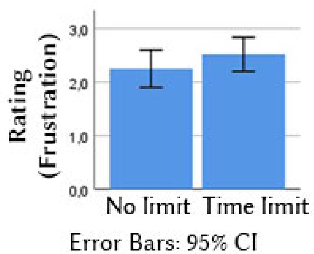
## LESSONS LEARNT & DISCUSSION

### Study Limitations

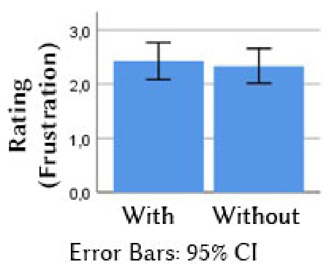
Our study sample is biased towards young people and students. This may influence our participants' opinion towards a) the tasks and b) the tracking technology we applied. In addition, participants were required to wear smartwatches and markers at both hands during the study, which may have influenced their interaction behaviour. Furthermore, our setup may have nudged users to a certain




**(a) Bricks**




**(b) Time**




**(c) Instruction**

**Figure 4: Frustration: Participants' rating (5-point Likert) on whether they were frustrated due to the problem (mean).**

behaviour (i.e., position of chair, desk, LEGO material). Lastly, participants were "under observation" (w.r.t. six *OptiTrack* cameras), which may have caused additional influence on their behaviour.

### Evaluating the GOLD Dataset

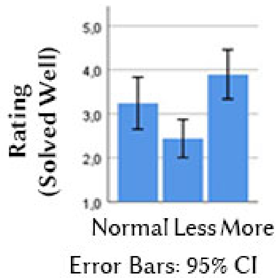
Applying tracking technologies requires synchronisation and polishing before analysis. In addition, data loss may occur due to technical problems or occlusion of tracking markers. For an adaptive support system, it is particularly relevant to know *when* a problem occurs (ideally beforehand). According to our observations, users were especially hesitating when searching for missing bricks. However, other characteristics may not allow to detect the problem at a certain moment in time (e.g. hesitation), while being recognisable in a longer time period (e.g., analysing the task as a whole and recognising hectic interaction). The GOLD dataset allows to work on classifications w.r.t. our (given) problem conditions. However, in daily life, data may include "real" problems, which are unknown beforehand. This might need a remarkably high effort on data collection and training time to recognise these.

### Problem Situations & Characteristics

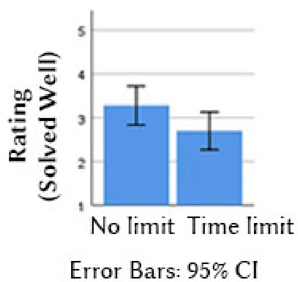
We intentionally focused on specific problem characteristics. However, this is not exhaustive. As an example, detectability may be another important problem characteristic. While some problems lead users to get stuck immediately, they may realise other problems after some time only. Moreover, "real" problems (in contrast to the ones we created artificially) in daily life tasks may be in a smaller (e.g., hesitation, thinking, searching) or larger (e.g., not knowing at all how to start) scope. Moreover, the difficulty of a task as well as the approach to solve a task are highly personal (i.e., our tasks may have been more or less difficult for particular participants, hence leading to different strategies) and may lead to side effects showing in user behaviour, which are not part of our evaluation so far.

### Applying Problem Detection in Everyday Life

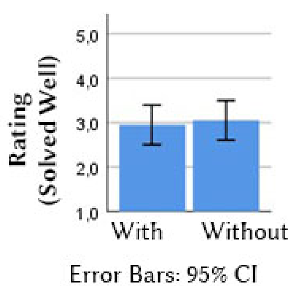
Problem detection and respective support could be applied in very frequent daily tasks, e.g. cooking. This action usually aims at preparing an enjoyable dish (*goal*). Meals with many ingredients may be more difficult to prepare than meals with fewer ingredients (*complexity*). *Time* is a limiting factor for a cooking task, e.g. due to arriving guests, due to parallel tasks (e.g. preparing sauce while cooking noodles) or due to increasing hunger. Lastly, the *available means* are crucial for success in a cooking task. *Knowledge* about sub-tasks (e.g., blanch vegetables) may be necessary. Missing ingredients or kitchen utensils (*material*) can lead to further problems. An ideal, adaptive support system recognises situations where users get stuck while cooking (e.g., due to a *complex* recipe, *time* running out, missing *knowledge* or *material*) and can provide them with help and guidance in the moment.




**(a) Bricks**




**(b) Time**




**(c) Instruction**

**Figure 5: Solving: Participants' rating (5-point Likert) on whether they solved the problem well (mean).**

## CONCLUSION & OUTLOOK

We lay the foundations for understanding user behaviour in problem situations. We hope to spark discussion on a) how to detect problems in interaction and b) developing adaptive support systems for such use cases. We provide the GOLD dataset for further analysis to the community.

## PROJECT RESSOURCES

The GOLD dataset can be requested from the first author, sarah.prange@hm.edu and is available at https://www.unibw.de/usable-security-and-privacy/downloads/datasets.

## ACKNOWLEDGMENTS

We thank all our participants for their time and effort to take part in our study. This research was supported by the German Research Foundation (DFG), Grant No. AL 1899/2-1.

## REFERENCES

[1] Davide Carneiro, José Carlos Castillo, Paulo Novais, Antonio Fernández-Caballero, and José Neves. 2012. Multimodal behavioral analysis for non-invasive stress detection. *Expert Systems with Applications* 39, 18 (2012), 13376 – 13389. https://doi.org/10.1016/j.eswa.2012.05.065

[2] Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality?: A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17)*. ACM, New York, NY, USA, 222–229. https://doi.org/10.1145/3056540.3056548

[3] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive Worker Assistance: Comparing the Effects of In-situ Projection, Head-mounted Displays, Tablet, and Paper Instructions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 934–939. https://doi.org/10.1145/2971648.2971706

[4] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139 – 183. https://doi.org/10.1016/S0166-4115(08)62386-9

[5] Thomas Kosch, Romina Kettner, Markus Funk, and Albrecht Schmidt. 2016. Comparing Tactile, Auditory, and Visual Assembly Error-Feedback for Workers with Cognitive Impairments. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16)*. ACM, New York, NY, USA, 53–60. https://doi.org/10.1145/2982142.2982157

[6] Thomas Kosch, Paweł W. Woźniak, Erin Brady, and Albrecht Schmidt. 2018. Smart Kitchens for People with Cognitive Impairments: A Qualitative Study of Design Requirements. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 271, 12 pages. https://doi.org/10.1145/3173574.3173845

[7] Don Norman. 2016. *The design of everyday things.* Verlag Franz Vahlen GmbH.

[8] A. Sano and R. W. Picard. 2013. Stress Recognition Using Wearable Sensors and Mobile Phones. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. 671–676. https://doi.org/10.1109/ACII.2013.117

[9] Walter Van de Velde. 1988. Inference Structure As a Basis for Problem Solving. In *Proceedings of the 8th European Conference on Artificial Intelligence (ECAI'88)*. Pitman Publishing, Inc., Marshfield, MA, USA, 202–207. http://dl.acm.org/citation.cfm?id=3069907.3069953